

Finite Topology as Applied to Image Analysis

V. A. KOVALEVSKY

*Central Institute of Cybernetics and Information Processing, Kurstrasse 33, 1086 Berlin,
German Democratic Republic*

Received August 10, 1988; accepted September 7, 1988

The notion of a cellular complex which is well known in the topology is applied to describe the structure of images. It is shown that the topology of cellular complexes is the only possible topology of finite sets. Under this topology no contradictions or paradoxes arise when defining connected subsets and their boundaries. Ways of encoding images as cellular complexes are discussed. The process of image segmentation is considered as splitting (in the topological sense) a cellular complex into blocks of cells. The notion of a cell list is introduced as a precise and compact data structure for encoding segmented images. Some applications of this data structure to image analysis are demonstrated.

© 1989 Academic Press, Inc.

1. INTRODUCTION

The essential difference between an image and a set of gray values is due to the adjacency relations between the gray values in an image. These relations may be specified by means of certain 2-dimensional structures. The simplest and most well known is the structure of a matrix or a rectangular raster. A raster image is well suited for the human perception. However, many notions important for the image analysis, such as the connectivity of regions, the boundaries, the adjacency, etc. are represented in a raster image only implicitly. It is rather difficult to formalize these notions and to make them accessible for a computer. These notions are subjects of the topology.

The modern general topology is mainly considering sets of space elements or points having the property that in an arbitrarily small neighborhood of a point there are infinitely many other such points. An example of such a set is the continuum. The number of its elements is uncountable. A digitized image, to the contrary, has always a finite number of elements. Thus we need a consistent topology of finite sets. All the necessary knowledge about such a topology is contained in the topological literature but it is spread through many books and articles. It is rather difficult to gather and to comprehend this knowledge. The specialists for topology, however, have made no efforts to implement their knowledge into image processing, computer graphics, and related sciences.

For this reason the specialists on image processing were forced to look for their own solution of the problem. At the beginning of the seventies Rosenfeld [2] has suggested regarding an image as a graph whose nodes are the pixels and whose edges are linking the adjacent pixels to each other. It is called the neighborhood graph. This notion is widely used in image processing and computer graphics up to now. Thus the possibility has arisen to define certain topological notions for finite sets, e.g., that of the connectivity.

However, the attempts to develop a consistent topology of 2-dimensional images by means of graphs have failed due to the so-called connectivity paradox and some other contradictions [3, 4]. The connectivity paradox, e.g., consists in the following.

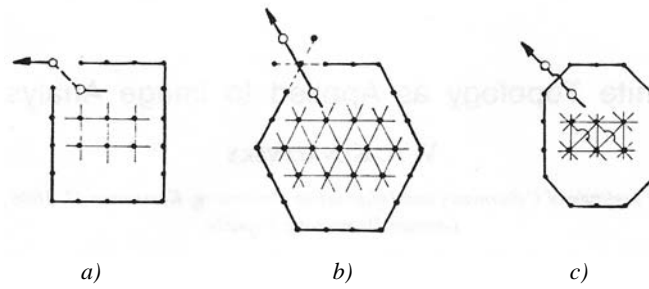


FIG. 1. Validity of the Jordan theorem under different neighborhood graphs

The well-known Jordan theorem claims that a simple closed curve in the Euclidean plane separates the remainder of the plane into two connected components (the inside and the outside of the curve). It may be easily shown that, when removing a point from the curve, the remainder of the plane becomes connected. When replacing the plane by the raster with a neighborhood graph on it, everyone expects to find a similar situation. The finite analog of a curve is a simple closed path: a sequence of adjacent graph nodes in which every node has exactly two adjacent ones (Fig. 1). The remainder of the raster (i.e., the set of the nodes not belonging to the path) is, however, not always separated into two connected components: it depends upon the kind of neighborhood graph. The graphs commonly used in image processing are the so-called 4- and 8-neighborhood graphs (Figs. 1a and 1c). In the case of the 8-neighborhood the inside of a simple closed path is not separated from its outside (Fig. 1c). In the case of a 4-neighborhood they remain separated even after removing a corner from the path (Fig. 1a). Both properties of a closed path are present only under the 6-neighborhood (Fig. 1b).

Especially unpleasant are the contradictions in the definitions of the boundaries. According to the general topology, the boundary of a region G relative to a space R is the set of all space elements e' such that every neighborhood of e' intersects both G and its complement $R \setminus G$. When transferring this definition directly onto a neighborhood graph one arrives at four different definitions: that of the inside and of the outside boundary, as well as that of the 4- and of the 8-boundary [3]. Some of these boundaries are shown in Figs. 2a and 2b. The figures show that the boundary defined under a 4-neighborhood is not connected under this neighborhood: it is 8-connected; the boundary under the 8-neighborhood is, however, 4-connected! Neither of them leads to a consistent definition of adjacent regions [3]. Moreover, a boundary so defined is itself a set of pixels and thus has a finite area, while according to our intuition it must be a 1-dimensional set having a zero area. Further, the boundary of a set and that of its complement are different, which is never the case in the general topology.

One more difficulty consists in the impossibility to transfer the most important topological notions such as that of an open subset, open neighborhood, continuity, topological mapping, and many others, onto a 2-dimensional structure described by a neighborhood graph.

Substantial steps to overcome the difficulties have been made in the 1970s; Rosenfeld [5] suggested using different neighborhoods for the background and the

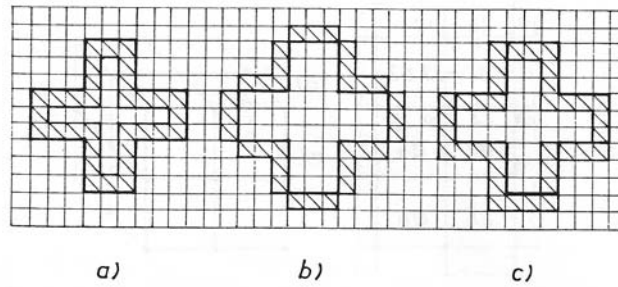


Fig. 2. The inside 8-boundary (a), the outside 4-boundary (b), and the extended boundary by Feng and Pavlidis (c).

objects. He also noticed that the 6-neighborhood does not lead to the connectivity paradox. Pavlidis and Feng [3, 6] introduced the notion of the "extended boundary" (Fig. 2c) which solved the region adjacency problem; Serra[7] suggested considering a mixed 4/8-neighborhood; Elliott and Srinivasan [8] proposed considering the so-called boundary elements, i.e., line segments separating two adjacent pixels from each other; Klette and Voss [9] introduced the notion of "meshes" corresponding to cycles in the neighborhood graph which uniquely define its imbedding into the plane.

All these attempts, however, turned out to be imperfect solutions: they have not led to a consistent topological theory. Moreover, after having watched the development of these efforts one gets the impression that there exists no way to a "self-made" or "substitute" topology at all. It is necessary to look for the solution of the problem in classical topology.

On the other hand, when analyzing the intuitive trends of the above-mentioned attempts, one may notice that many of them have tried to consider a digital plane as a structure consisting of *heterogeneous elements*, namely, of elements of different dimensions: 0-dimensional points, 1-dimensional line elements, and 2-dimensional area elements [10, 11]. Such a structure is well known in the topology as a cellular complex [1]. A consequent study of this conception has really led to a consistent theory [12] which is based exclusively on well-known and proven topological results. Amazing as it is, these results have never been used until now in the image processing and related fields. The theory is represented in the following sections.

2. THE TOPOLOGY OF CELLULAR COMPLEXES

This classical branch of the general topology was founded by the famous Riemann in the 19th century. It considers structures consisting of elements of different dimensions called *cells* [1]. The cells may be interpreted as faces of a polyhedron or simply considered as abstract elements of a set.

DEFINITION 1. An abstract *cellular complex* $C = (E, B, \dim)$ is a set E of abstract elements provided with an antisymmetric, irreflexive, and transitive binary relation $B \subset E \times E$ called the *bounding relation* (or the face relation) and with a *dimension* function $\dim: E \rightarrow I$ from E into the set I of the non-negative integers such that $\dim(e') < \dim(e'')$ for all pairs $(e', e'') \in B$.

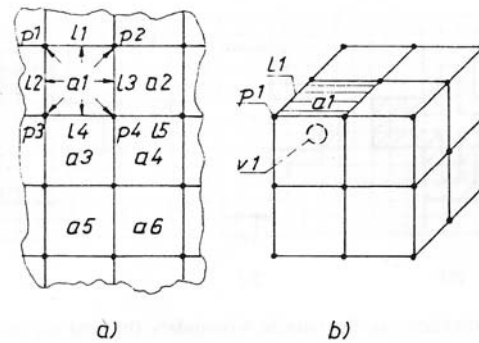


FIG. 3. Examples of a 2-dimensional (a) and a 3-dimensional (b) complexes.

The bounding relation B is a partial order in E . It indicates the ordered pairs (e', e'') of elements such that e' is said to bound e'' , or e' is a (proper) face of e'' which is denoted by $e' < e''$. "To bound x " means "to be a part of the boundary of x ". Thus a cell is said to bound another cell of higher dimension if it is a face of it in the same sense as e.g. a triangle may be a 2-dimensional face of a tetrahedron, a line segment may be a 1-dimensional face of a triangle, or of a tetrahedron, and a point may be a 0-dimensional face of a line segment, or of a triangle, or of a tetrahedron. (The relation B may be also chosen to be reflexive. In this case an element must be regarded as a non-proper face of itself and $(e', e'') \in B$ implies $\dim(e') \leq \dim(e'')$. We shall however regard B always as being irreflexive).

The function $\dim(\cdot)$ defines the dimension of each element. An element e' with $\dim(e')=d$ is called a d -dimensional element or a d -cell. A complex is called k -dimensional or a k -complex if the dimension of all its elements is less or equal to k . Examples of a 2-complex and a 3-complex are shown in Fig. 3.

The 2-dimensional elements are the area elements. In image processing they must be associated with the notion of pixels since a gray value assigned to a pixel originates from measuring the amount of energy radiated from an elementary area. The 0-dimensional and 1-dimensional elements are not represented in the hardware realizations of image memories and displays. This is a possible cause which has prevented the implementation of cellular complexes for image processing. We shall, however, show that there is no problem in encoding the elements of lower dimensions in the same way as it is being done for the pixels. The construction of the corresponding hardware is of course also possible but the necessity of it must yet be proved.

Consider now some notions of the theory of cellular complexes. A subcomplex $C' = (E', B', \dim')$ of a given complex $C = (E, B, \dim)$ is a complex whose set E' is a subset of E and the relation B' is the intersection of B with $E' \times E'$. This means that the bounding relation remains the same as that of C for every pair of elements contained in E' . The dimensions also do not change: $\dim'(e) = \dim(e)$ for all $e \in E'$. All this is important because it means that to define a subcomplex C' of $C = (E, B, \dim)$ it suffices to define the corresponding subset E' of the elements. Thus it is possible to speak of a subcomplex $E' \subset E$ while understanding the subcomplex $C' = (E', B', \dim')$. All subcomplexes of C may be regarded as subsets of E and thus it is possible to use the common formulae of the set theory to define

intersections, unions, and complements of subcomplexes of one and the same complex C .

The cellular complexes have to play an exclusively important role in image processing and computer graphics since they are the only means to consistently define the topology of finite sets. Before proving this we shall remind the reader of some fundamental topological notions.

A *topological space* is a pair (E, SY) consisting of a set E of abstract elements and a system $SY = \{S_1, S_2, S_3, \dots, S_i, \dots\}$ of subsets S_i of E . These subsets are called the *open subsets* of the space and must satisfy the following axioms:

- A1. The empty subset \emptyset and the set E belong to SY .
- A2. For every family F of subsets S_i belonging to SY the union of all subsets which are elements of F must also belong to SY .
- A3. If some subsets S_1 and S_2 belong to SY then the intersection $S_1 \cap S_2$ must also belong to SY .

A topological space has the separation property (more precisely T0-property) if it satisfies one more axiom:

- A4. For any two elements $e_1, e_2 \in E$ there exists in SY such an open subset S' that exactly one of the elements is in S' .

The importance of Axiom 4 becomes clearer if we note that according to this axiom there exists for any two space elements $e_1, e_2 \in E$ such a subdivision of the whole space into two complementary parts S' and $E \setminus S'$ that one of the said elements be in S' , the other in its complement $E \setminus S'$, and one of the parts be open.

A topological space is called a *finite* one if the set E contains a finite number of elements. It must be noticed that Axiom 2 and Axiom 3 are formulated in two different ways because Axiom 2 is valid for both finite and infinite families of F while Axiom 3 is valid only for the intersection of a *finite number* of subsets S_i . This remark is, however, of no importance for the case of a finite set E which is only under consideration in the present paper.

We have also cited Axiom 4 in its weakest version (called the "T0 axiom"). The stronger versions (corresponding to the T1 and T2 axioms) are of no importance for finite sets. Therefore we say that a space satisfying Axiom 4 has the separation property while it is common in the topological literature to use this term for T2 spaces called also Hausdorff spaces.

THEOREM. *Every finite topological space with the separation property is isomorphic to an abstract cellular complex according to Definition 1.*

The proof of the theorem may be constructed by citing some well-known topological results. Such a proof would be, however, comprehensible for specialists only. That is why we prefer to present here a rather self-contained proof whose understanding requires only common mathematical knowledge.

Proof. Any open subset containing a given element e' is called the *open neighborhood* of e' . The intersection of all open neighborhoods of e' (in a finite space) is according to Axiom 3 an open subset. We shall call it the smallest neighborhood of e' and denote by $SN(e')$.

Consider the smallest neighborhoods of all elements of E . Suppose, first, there exists an element e_1 whose smallest neighborhood $\text{SN}(e_1)$ contains at least one element e_2 different from e_1 :

$$e_2 \in \text{SN}(e_1) \quad \text{and} \quad e_2 \neq e_1. \quad (1)$$

$\text{SN}(e_1)$ is the intersection of all open neighborhoods of e_1 . Therefore e_2 belongs to all of them. According to Axiom 4 there must be such an open neighborhood of e_2 to which e_1 does not belong. Thus e_1 does not belong to $\text{SN}(e_2)$:

$$e_1 \notin \text{SN}(e_2). \quad (2)$$

Conditions (1) and (2) represent a binary relation B between the elements e_1 and e_2 . It is easy to see that this relation is an antisymmetric and irreflexive one. We are going to show that it is transitive.

Consider three elements e_1 , e_2 , and e_3 such that the relation B holds for the pairs (e_1, e_2) and (e_2, e_3) . This means:

$$e_2 \in \text{SN}(e_1) \quad \text{and} \quad e_1 \notin \text{SN}(e_2); \quad (3)$$

$$e_3 \in \text{SN}(e_2) \quad \text{and} \quad e_2 \notin \text{SN}(e_3). \quad (4)$$

Regard the intersection $S = \text{SN}(e_1) \cap \text{SN}(e_2)$. According to Axiom 3 it is an open subset and it contains e_2 . Thus it is an open neighborhood of e_2 . It cannot be a proper subset of $\text{SN}(e_2)$, since $\text{SN}(e_2)$ is by definition contained in every open neighborhood of e_2 . Hence $S = \text{SN}(e_2)$ and

$$\text{SN}(e_2) \subset \text{SN}(e_1). \quad (5)$$

It follows from (4) and (5) that $e_3 \in \text{SN}(e_1)$ and, according to Axiom 4, there exists an open subset S' such that $e_3 \in S'$ but $e_1 \notin S'$. The set $\text{SN}(e_3)$ is by definition the intersection of all open subsets containing e_3 . Therefore $\text{SN}(e_3) \subset S'$ and

$$e_1 \in S' \Rightarrow e_1 \notin \text{SN}(e_3).$$

Thus the relation B holds for the pair (e_1, e_3) which proves the transitivity of B .

According to Definition 1 any antisymmetric, irreflexive, and transitive relation may serve as the bounding relation in an abstract cellular complex. The dimension function $\text{dim}(\cdot)$ may be defined in any way providing:

$$\forall e_1, e_2 \in E, \quad (e_1, e_2) \in B \Rightarrow \text{dim}(e_1) < \text{dim}(e_2).$$

One possible way [13] of finding a concrete function $\text{dim}(\cdot)$ consists in counting the number of elements in the smallest open neighborhoods. The dimension of an element e' is then equal to

$$\text{dim}(e') = \max_e (\text{CARD}(\text{SN}(e)) - \text{CARD}(\text{SN}(e'))). \quad (6)$$

We have shown during the proof of the theorem that $e_1 < e_2$ implies

$SN(e_2) \subset SN(e_1)$ and hence $CARD(SN(e_2)) < CARD(SN(e_1))$. Therefore the dimensions defined by (6) satisfy the condition of Definition 1.

Thus we have shown that a finite topological space in which at least one smallest neighborhood consists of more than one element is isomorphic to an abstract cellular complex.

In the remaining case the condition $SN(e_1) = e_1$ holds for every $e_1 \in E$. Thus every element is an open subset. Such a topological space is called a *discrete space*. There exists in it no element e_2 satisfying the conditions (1) and (2) together with e_1 . The relation B is in this case empty. This corresponds to a cellular complex in which there are no elements bounding one another. It may be, e.g., a 0-dimensional complex consisting of 0-dimensional elements only. Therefore a finite topological space is in any case isomorphic to an abstract cellular complex. \square

The proven theorem is of great importance since it shows that the search for finite topological structures different from cellular complexes is *hopeless*.

It may be demonstrated also that every cellular complex by Definition 1 is a topological space. Really, we have shown during the proof of the theorem that the smallest neighborhood of a given element e_1 consists of e_1 itself and of all elements bounded by e_1 . This consideration indicates the way of defining all open subsets of a cellular complex:

DEFINITION 2. A subcomplex S of C is called *open* in C if for every element e' of S all elements of C which are bounded by e' are also contained in S .

E.g., $S_1 = \{a_1, l_3, a_2\}$ is open in the complex of Fig. 3a but $S_2 = \{l_4, p_4, l_5\}$ is not open in it since S_2 does not contain, e.g., the element a_1 which is bounded by l_4 .

The open subset $St(e', C)$ consisting of e' and all elements of C which are bounded by e' is called the *open star* of e' in C . It is exactly the same as the smallest neighborhood $SN(e')$ of e' in C . The notion of an open star is of great importance since open stars are the simplest open subsets in a complex. All other open subsets are unions of some open stars. Open stars of various elements in complexes of different dimensions are shown in Fig. 4.

It is easy to see that the open subsets defined in this way satisfy the axioms of the topology. Therefore any complex is a topological space. All notions and results of the general topology which are meaningful for finite sets may be transferred to complexes. Being apart of the general topology the theory of complexes is consistent. We may demonstrate now the ways in which all the paradoxes and contradictions of the neighborhood graphs disappear.

3. SOLUTION OF THE PARADOXES

It is well known that the so-called 4-neighborhood and the 8-neighborhood lead to the connectivity paradox [3, 4]. The 6-neighborhood, however, is free of the paradox. Hence it may seem that in a consistent finite topological space every 2-cell must have 6 adjacent 2-cells. But the classical topology is considering complexes whose 2-cells may have an arbitrary number of "neighbors" without running against contradictions. Hence the real cause of the paradox is not the number of neighbors but rather the fact that the neighborhood graphs are not in agreement with the axioms of the topology. The way in which the topology of complexes overcomes the paradox is the following [11, 12].

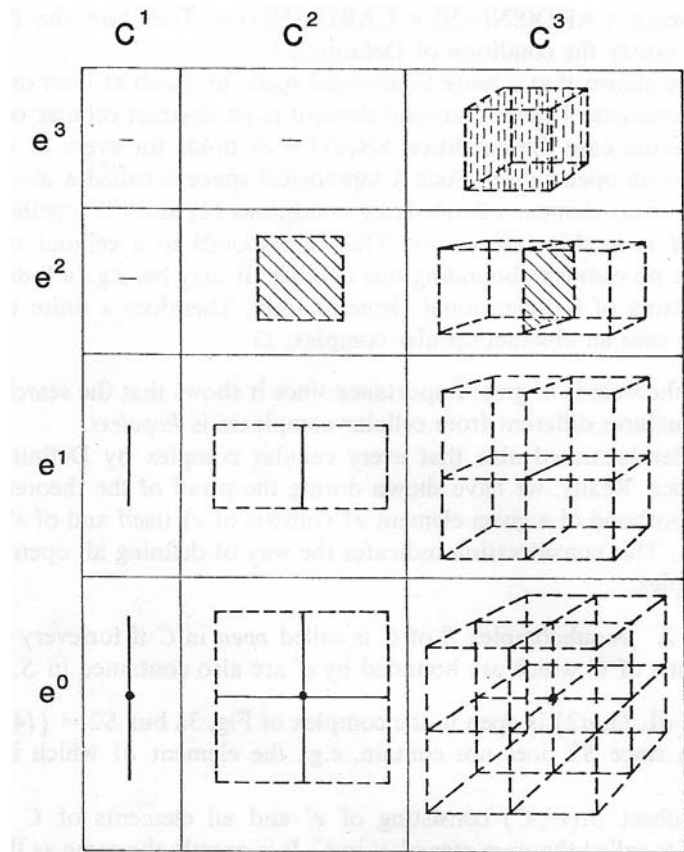


FIG. 4. The open stars $St(e^k, C^d)$ of various elements e^k in complexes C^d of dimensions $d = 1, 2,$ and 3 .

Consider the topological notion of connectivity as applied to cellular complexes:

DEFINITION 3. A sequence of elements of a subset S of a complex C beginning with e' and finishing at e'' is called a *path* in S from e' to e'' if for every two elements which are adjacent in the sequence one of them is bounding the other.

DEFINITION 4. A subset S is called *connected* if for any two elements e' and e'' of S there exists a path in S from e' to e'' .

Definition 4 corresponds to the notion of "arcwise connected" in the general topology. Let us apply this Definition to the simplest situation of Fig. 5 which was often used [3, 4] to demonstrate the essence of the connectivity paradox.

Consider the complex O of Fig. 5 which contains the 2-elements $A, B, C,$ and D . A subcomplex S containing A and D but not B and C is connected according to Definition 4 if and only if the 0-element P belongs to S since a path in S from A to D can only go through P . If $P \in S$ then it cannot belong to $O \setminus S$ and therefore $O \setminus S$ is not connected. Therefore when assigning the 0-element P either to S or to its complement no paradox arises. When, however, the membership of P in S

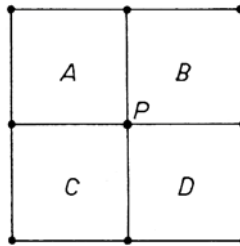


FIG. 5. To the connectivity paradox.

remains undefined then everything becomes possible: e.g., that S and $O \setminus S$ are both connected or both disconnected. Hence ignoring the 0-elements is the true cause of the paradox.

What is the exact cause of the paradox when using a neighborhood graph? To explain it we shall need two more topological notions:

DEFINITION 5. Two n -dimensional cellular complexes $C = (E, B, \dim)$ and $D = (E', B', \dim')$ are called *dual* to each other if there exists a one-to-one mapping from E to E' such that the bounding relation be reversed: if e_1 is mapped onto d_1 and e_2 is mapped onto d_2 then $e_1 < e_2 \Rightarrow d_2 < d_1$ and $\dim'(d_i) = n - \dim(e_i), i = 1, 2$.

DEFINITION 6. A k -dimensional subcomplex S of an n -dimensional cellular complex $C = (E, B, \dim)$ with $k < n$ containing all elements of E with dimensions less than or equal to k is called a *k-dimensional skeleton* of C .

It is important to notice that for each n -complex there exist exactly one dual n -complex and exactly one k -dimensional skeleton for each $k < n$.

A neighborhood graph (as any other graph) is a one-dimensional complex. It is easy to see that the 4-neighborhood graph NG_4 is a one-dimensional skeleton of the 2-dimensional complex dual to the complex CR representing the raster (Fig. 6a). A subgraph SG of NG_4 (e.g., Fig. 6b) is specified by a subset V of the vertices and those edges whose both end points are in V . Such a subgraph SG is specifying a dual sub-

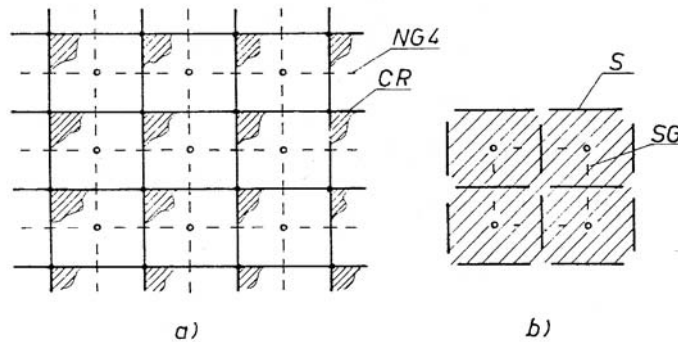


FIG. 6. Relation between the 2-complex representing the raster and the 4-neighborhood graph (explanations in the text).

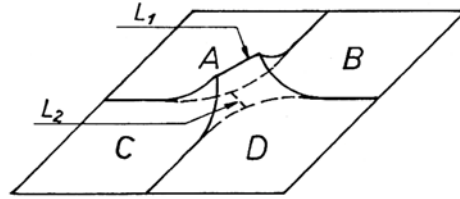


FIG. 7. The unreal topological structure corresponding to that of the 8-neighborhood graph.

complex S (Fig. 6b) of the raster CR which subcomplex consists of the pixels dual to the vertices of V and of the 1-elements dual to the edges of the subgraph SG . The 1-elements are each separating two pixels of S . The graph $NG4$ as well as its subgraph SG being 1-skeletons contain no 2-elements which are dual to the 0-elements of the raster CR . Hence the membership of the 0-elements in S remains undefined: they belong neither to S nor to its complement and therefore there are no means to represent the diagonal connections which are due to the 0-elements.

This drawback is compensated for in the 8-neighborhood graph $NG8$ by inserting in the 1-skeleton the diagonal edges which correspond to *no* 1-elements of the given complex CR . Their role is thought of as representing the adjacency of the 2-elements having a common 0-element. This representation is, however, not adequate: it represents the unreal topological situation shown in Fig. 7. It shows how the complex of Fig. 5 should be deformed to be in accordance with the 8-neighborhood graph: the 2-elements A and D must have a common 1-element $L1$ and the 2-elements B and C must have another common 1-element $L2$ which does not intersect $L1$. These two 1-elements are artificially inserted into the structure to compensate the absence of the 0-element P which is a common face of all four 2-elements A , B , C , and D . The structure of Fig. 7 is, however, topologically impossible: to represent it we were forced to "split" all the 1-faces of A , B , C , and D .

Why is the 6-neighborhood free of paradoxes? A 2-complex in which every 0-element bounds not more than three 1-elements has the peculiarity that any two 2-elements which have a common face have also a common 1-face (such 2-complexes are called *strongly connected*). Therefore there is no need to specify the membership of the 0-elements: the connectedness is completely defined by the membership of the 1-elements. This information may be correctly specified by a neighborhood graph. Thus a neighborhood graph is able to represent the connectivity of a 2-dimensional topological structure if and only if the corresponding subcomplex is a strongly connected one.

It is easy to see that the connectivity paradox disappears if every subset of the complex CR representing the raster is correctly defined. This means one should never define a subset S of CR by simply specifying all pixels belonging to it and saying nothing about the space elements of lower dimensions. A correct definition may be done by a *membership predicate* $M(S, e)$ according to which any space element e of *every dimension* is declared as belonging either to S or to its complement $C \setminus S$. Especially plausible is the case when N subsets S_i , $i = 1, 2, \dots, N$ compose a partition of the space CR . Then any membership predicate may be interpreted as coloring every space element of any dimension by one of N colors.

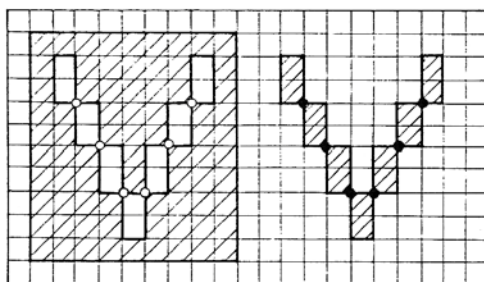


FIG. 8. Example of a consistent coding of connected black and white diagonals.

Consider a rather illustrative example of Fig. 8. Neither of the commonly used neighborhood relations makes it possible to define the boundaries of both V-shaped regions of Fig. 8 in such a way that the boundary of each V is connected and a tracing procedure is able to perform similar sequences of steps in both cases. It is, however, easily realizable if we define the 0-cells depicted by white circles as being members of the white V-shaped region and the 0-cells depicted by black circles as belonging to the dark V. Then each boundary must be traced in such a way that both the 0-cells and the 2-cells having the color of the chosen region remain always to the right-hand side of the direction of the tracing. No connectivity paradox arises in the case of the white V since the white pixels are connected to each other through the white 0-cells but the black pixels of the background are disconnected. A similar situation takes place for the black V.

Thus we see that the paradox vanishes if we cease forcing a neighborhood relation to pairs of space elements of *equivalent dimension*. The only consistent way of specifying the connectivity of some subsets of an image is that of representing the image as a cellular complex and defining the membership of all cells of any dimension in these subsets. Such a representation of images not only abolishes the contradictions, it also gives the possibility to adequately represent the structure of images containing fine details as it was demonstrated by the example of Fig. 8. The techniques for a practical realization of this method are discussed in the next section.

Consider now the way of overcoming the contradictions in defining the boundaries. The theory of complexes enables one to use for finite sets the classical definition of the boundary:

DEFINITION 7. The *boundary* of a subset $S \subset C$ relative to C is the subset $\text{Fr}(S, C)$ of all elements e' of C such that any open neighborhood of e' contains elements both of S and of its complement $C \setminus S$.

It is more convenient in practice to replace "any open neighborhood" by "an open star $St(e', C)$ " in Definition 7. Thus the subset S of Fig. 9a which consists of the elements a_1, a_2, a_3, l_1, l_2 , and p , has the boundary shown in Fig. 9b by heavy lines and dots. It is easy to see that the open stars of all boundary elements of Fig. 9b satisfy Definition 7.

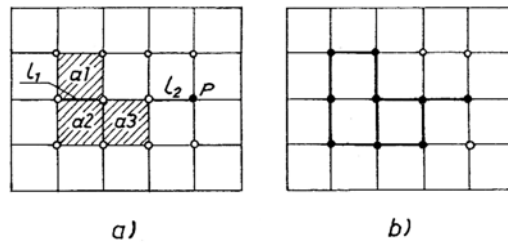


FIG. 9. Example of a subcomplex (a) and its boundary (b).

Under the topology of cellular complexes there exists a unique definition of the boundary: there is no need to distinguish between the inside and outside boundary as it was the case with the neighborhood graphs. The boundary does not depend on the kind of neighborhood since there is no such notion at all! The boundary has no end points as this was the case for the 4-boundary of Fig. 2b. The reader may easily prove this assertion when supposing a 0-cell to be an end point of the boundary of a subset S and considering all possible combinations of the membership in S of the four pixels bounded by this 0-cell.

The notion of adjacent regions which was profoundly investigated by Pavlidis [3] may be successfully replaced by that of incident ones:

DEFINITION 8. Two subcomplexes S_1 and S_2 of a complex C are called *incident* to each other if they do not intersect and there are two elements $e_1 \in S_1$ and $e_2 \in S_2$ such that one of them bounds the other.

Consider now the problem of the area of a boundary. A boundary of a 2-dimensional subset S'' of a 2-dimensional complex C'' (e.g., of an image) contains no 2-elements since the open star of such an element e'' consists of a single element (which is e'' itself) and hence it cannot cross both S'' and its complement $C'' \setminus S''$. Therefore such a boundary consists only of 1- and 0-dimensional elements, i.e., of line elements and points. It is a 1-dimensional subcomplex of C'' . It is natural to assign a non-zero area only to the 2-dimensional elements (the pixels). Then the area of a 1-dimensional complex is equal to zero, which is in correspondence with our intuition.

Finally the boundary according Definition 7 is one and the same for a subset S and for its complement $C \setminus S$, since the Definition 7 is symmetric with respect to S and $C \setminus S$.

We have demonstrated that the concept of cellular complexes being a consistent branch of the classical well-proved topology removes all topological paradoxes and contradictions from the theory of digitized images. It must be stressed here that the concept has many more advantages: it may be applied without any change to describe the topological structure of 3-dimensional and even n -dimensional images. This, however, is a topic for a possible separate article.

4. ENCODING IMAGES ON CELLULAR COMPLEXES

An image on a raster is defined by a mapping of raster elements onto a set of labels. The labels may be interpreted as gray values or simply as symbols of certain

subsets, e.g., the objects and the background or subsets of pixels all having a certain gray value. According to our concept the raster must be represented as a 2-dimensional cellular complex consisting of elements of three different dimensions: the 2-elements which are the pixels and the 1- and 0-elements which are the faces of the pixels.

Encoding an image means encoding the subset membership of all the elements. The encoding of the pixels must be realized in the traditional way: a memory word must be assigned to each pixel and the label of the subset to which the pixel belongs may be put into the assigned word. The encoding of the elements of lower dimensions may be generally realized in the same way. It should be noticed that in many cases a fraction of a memory word is sufficient to encode the membership of an element: if up to NS subsets must be distinguished in the image then $K \geq \log_2 NS$ bits per element are necessary. E.g., in the case of binary images only two subsets are to be distinguished and hence only one bit per element is needed. It may be shown by means of the well-known Euler theorem that the number of the 0-elements in the 2-dimensional raster (when neglecting the borders) is equal to the number NP of the pixels. The number of 1-elements is $2NP$. Therefore the total memory demand for encoding the subset membership of all elements of a 2-dimensional raster in the most general case is equal to $4NP \cdot \log_2 NS$. It is four times more than when encoding the pixels only.

Such a great memory volume is, however, rarely (if ever) needed in practice. There are many ways to use a priori knowledge about the images under consideration to obtain an economical coding of the set membership. E.g., it is quite natural to try introducing some kind of a *global face membership*, i.e., a rule specifying the set membership of the faces of every 2-cell as a function of the membership of the 2-cell itself. However, such a global face membership cannot be chosen arbitrarily. It must be representable as a membership predicate as defined above. E.g., it is not allowed to specify all faces of every 2-cell as belonging to the same set S as the 2-cell itself since one and the same element may be a face of two 2-cells, one belonging to S and the other to $C \setminus S$.

An example of a consistent global face membership may be constructed by assigning every 0- and 1-element to a uniquely defined 2-element, e.g., as shown in Fig. 10. The rule consists in declaring every 0- and 1-cell as belonging to the same set as the 2-cell uniquely defined by the assignment of Fig. 10. It is easy to see that such a global face membership is equivalent to declaring a 6-neighborhood of the 2-cells.

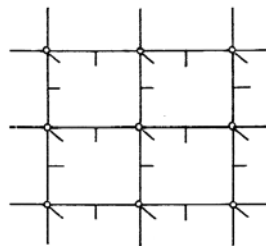


FIG. 10. One possible assignment of the 0- and 1-dimensional elements to the 2-dimensional ones.

The two other most often used neighborhoods may also be regarded as special cases of a global face membership: the 4-neighborhood specifies that the 0-cells belong to *no subset* (which is impossible). The 8-neighborhood corresponds to *doubled* 0-cells: every 0-cell exists in two copies, one belonging to the same subset as the pixel lying to the north-west of it and the other to the same subset as the one to the north-east. As we already know this may lead to contradictions if these two pixels belong to different subsets. Therefore both these global rules and the corresponding neighborhoods are contradictory. They should not be used any more.

Among the practically useful rules there is the following one which may be used if the labels of the pixels are elements of an ordered set as, e.g., that of the numbers:

THE MAXIMUM LABEL RULE. Every element e' of dimension 0 or 1 gets the label which is equal to the maximum of the labels of all pixels bounded by e' .

The effect of this rule may be demonstrated on the example of Fig. 5: when two pixels composing a diagonal pair in an open star of a point P (e.g., A and D in Fig. 5) have one and the same label (gray value) L which is different from the labels of the two other pixels, then the pair becomes connected if and only if L is greater than both the labels of the other diagonal pair (e.g., B and C in Fig. 5). The other diagonal pair becomes in this case disconnected. It is easy to see that Rosenfeld's proposal [5] to use the 4-neighborhood for the background (small labels) and the 8-neighborhood for the objects (large labels) is a particular case of the maximum label rule. The advantage of the latter is its applicability to multivalued (non-binary) images.

An important drawback of this very simple rule is that narrow slanting stripes which are only one pixel wide become disconnected if their gray level is smaller than that of the background. It is possible to introduce a *minimum label rule* but the same problem arises then for light stripes. A similar drawback is characteristic for the above-mentioned global rule corresponding to the 6-neighborhood: when using the assignment of Fig. 10, then narrow stripes slanted to the left are connected but those slanted to the right are disconnected. It is possible to change the assignment to make the stripes slanted to the right connected but those slanted to the left become disconnected. No one of these three rules may represent all narrow stripes in one and the same image (either dark or light, either those slanted to the left or to the right) as being connected. An example of such an image was shown in Fig. 8.

In the cases when the latter demand is of importance a more complex global face membership rule may be applied. This rule provides an approximate solution of the problem of minimizing the number of connected regions in the image:

THE EQUINALI-RULE. Every 1-element l' gets the label which is equal to the maximum of the labels of the two pixels bounded by l' . The label of a 0-element p' is defined as follows:

If the open star of p' contains exactly one diagonal pair of pixels with equivalent ("EQU") labels then p' gets this label.

If there are two such pairs but only one of them belongs to a narrow ("NA") stripe then the label of the latter pair is assigned to p' .

Else the maximum (or minimum) label rule must be applied and thus the lighter ("LI") (or darker) pair becomes connected.

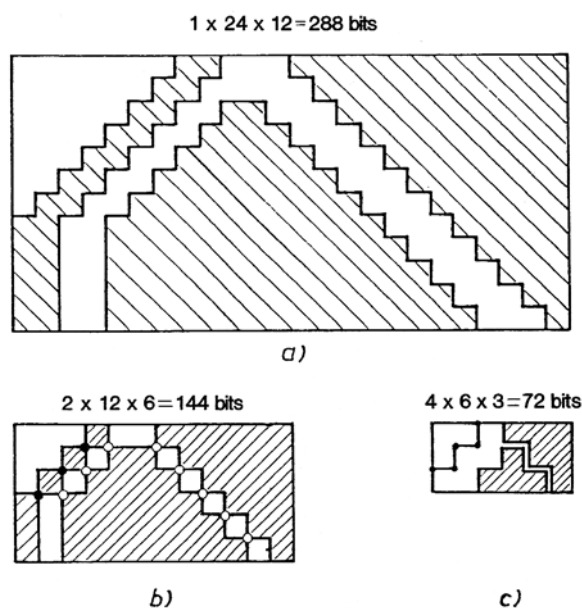


FIG. 11. An example of the necessity of explicitly encoding the face membership.

It is obvious that testing a neighborhood of the 0-element p' containing 2×2 pixels is not sufficient to decide whether a pixel pair belongs to a narrow stripe. The neighborhood must contain at least $4 \times 4 = 16$ pixels. The result of the application of this rule may be seen in Fig. 8.

Using a global face membership rule enables one to save memory space since only the pixels must be encoded explicitly. The membership of the pixels' faces is then specified implicitly and needs no memory space. However, it must be stressed here that all the existing global rules represent only a small fraction of all possible membership specifications of the pixels' faces. There are situations in which no global rule may be applied and the face membership must be encoded explicitly.

Consider the example of Fig. 11a which is a binary image with comparatively high resolution. There are no fine details in the image and hence it may be encoded by means of some global rule. This means that only the pixels must be encoded explicitly and one bit per pixel is sufficient to correctly specify the connectivity of the regions. The image contains 12 rows and 24 columns. Thus the total memory demand is 288 bits.

Fig. 11b represents the same image after being compressed by the factor 2 in both directions. The resolution became smaller and now not every global rule may correctly represent the connectivity of all black and white regions. The image may be, e.g., correctly encoded either by the EQUALI rule or by explicitly encoding the 0-elements. The 1-elements need not to be encoded explicitly: it suffices to define their membership as that of one of the incident pixels. Therefore when explicitly encoding the 0-elements one needs 2 bits per pixel: one for the pixel itself and one for the 0-element assigned to it according to the assignment rule of Fig. 10. The total memory demand is $2 \times 12 \times 6 = 144$ bits. It is less than before in spite of using 2 bits per pixel rather than 1 bit as before.

After having compressed the image once again, we are still able to correctly encode the connectivity of all regions when using the explicit encoding of both the 0- and 1-elements (Fig. 11c). Now 4 bits per pixel are needed, but the total memory demand becomes still smaller: $4 \times 6 \times 3 = 72$ bits.

One may conclude from this example that the necessity to encode four times more space elements when using the cellular complexes is not necessarily a drawback: under a high resolution a global face membership may be used and hence no additional memory as compared to the traditional encoding is necessary; under a low resolution more memory space per pixel is needed, but this enables one to correctly represent some fine details which would be impossible when using the traditional way of encoding an image.

5. THE CELL LIST DATA STRUCTURE

We have defined an image as a 2-dimensional cellular complex with a label assigned to each element. The labels originating from the gray values measured during the scanning of an optical image are mostly rather variable. This means that there are in an image rarely great connected regions consisting of pixels having a constant value. The term "region" is used here in accordance with the corresponding notion of the topology where a region is defined as an open connected 2-dimensional subcomplex.

On the other hand, there are often regions in an image which have a semantic meaning from the point of view of the applications, e.g., regions representing some objects of interest or parts of objects. To find such regions and to label all elements of such a region by the same symbol is the essence of the segmentation problem. We are, however, not going to discuss here the segmentation problem. Our aim is only to consider a possible way of encoding the results of the segmentation.

The procedure of image segmentation may be regarded from the point of view of cellular complexes as *splitting*: a cellular complex representing the raster with an image on it must be split into *blocks of cells* all possessing a certain predefined feature. The feature may be defined, e.g., by means of the "uniformity predicate" by Pavlidis [3].

The 2-dimensional blocks are the *regions* as defined above. The boundary of a region is a 1-dimensional subcomplex which may have *branching points* at the locations where three or more regions meet. The branching points are the 0-dimensional blocks. Parts of a boundary separated by branching points are the 1-dimensional blocks. They are also called boundary segments or *lines*. If the boundary of a region has no branching points then every connected boundary component is a line or a 1-block.

There exists a bounding relation between the blocks: the lines may be bounded by some branching points, the regions are bounded by some lines and, maybe, by some branching points. Thus the collection of these objects may be regarded as a new cellular complex whose 2-cells are the regions, the 1-cells are the lines, and the 0-cells are the branching points. These new cells are the blocks obtained as the result of splitting the original complex. We shall call such a new complex the *block complex* and its elements the *block cells*. An example of a block complex is shown in Fig. 12.

A block complex BC may be regarded as a generalization of the notion of a *region adjacency graph* [3, 5]. The latter is from the topological point of view the

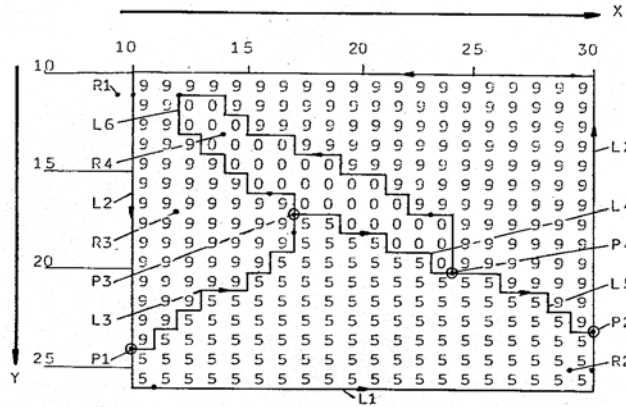


FIG. 12. Example of a block complex. R_i are the 2-cells, L_i are the 1-cells, and P_i are the 0-cells.

1-dimensional skeleton of the complex dual to BC . The block complex exactly represents the topology of the original image. This is its important advantage as compared to the region adjacency graph which does not represent the branching points.

A block complex may be described by a data structure called the *cell list*. It consists of one metric and three topological sublists. The topological sublists are that of the 0-dimensional, 1-dimensional, and 2-dimensional block cells. To provide the possibility of analyzing and restoring the original image the information represented in the lists of cells must be augmented by the metric data: the coordinates of the 0-cells and of some intermediate points in the 1-cells which define the exact location of the latter. This is the content of the metric sublist.

The intermediate points are found by means of dissolving the line into digital straight line segments (DSS). This provides an efficient polygonal representation of the boundaries. We have succeeded in encoding the DSS in such a way that an exact reconstruction of the original boundaries is possible. For this purpose the coordinates of the DSS end points are accompanied by some additional parameters which exactly specify which of the possible DSS running through the given end points takes place. Hence our polygonal representation is an exact and economical encoding of the boundaries rather than an approximation. We are using an original fast method of recognizing the DSS [14].

The structure of the cell list may be explained by the example of the image of Fig. 12. The corresponding lists are shown in Table 1. The first column in the list of the branching points contains the identifiers of the 0-cells as used in Fig. 12. (The identifiers are not stored in the computer). The next two columns contain the coordinates of the points. The following 4 columns contain the identifiers of (or pointers to) all lines (1-cells) bounded by the current 0-cell. The lines are directed. A minus sign of a pointer denotes that the corresponding line starts from the current point, a plus sign corresponds to a line ending at it.

The first column in the list of the lines contains the identifiers of the lines. The next two columns contain pointers to the beginning and ending points of the line. If the line is closed and hence is neither starting nor ending at a branching point then

TABLE 1

The Cell List Structure for the Cellular Complex of Fig. 12

| <i>List of the branching points</i> | | | | | | |
|-------------------------------------|-------------|----|-------------|-------------|-------------|-------------|
| No. | Coordinates | | Lines | | | |
| | X | Y | East | South | West | North |
| <i>P1</i> | 10 | 24 | - <i>L3</i> | - <i>L1</i> | 0 | + <i>L2</i> |
| <i>P2</i> | 30 | 23 | 0 | + <i>L1</i> | + <i>L5</i> | - <i>L2</i> |
| <i>P3</i> | 17 | 17 | - <i>L4</i> | + <i>L3</i> | 0 | + <i>L6</i> |
| <i>P4</i> | 24 | 20 | - <i>L5</i> | 0 | + <i>L4</i> | - <i>L6</i> |

| <i>List of the lines</i> | | | | | | |
|--------------------------|-----------|-----------|-----------|-----------|--------|-----|
| No. | Points | | Regions | | Metric | |
| | Begin | End | Right | Left | Begin | End |
| <i>L1</i> | <i>P1</i> | <i>P2</i> | <i>R1</i> | <i>R2</i> | 1 | 4 |
| <i>L2</i> | <i>P2</i> | <i>P1</i> | <i>R1</i> | <i>R3</i> | 5 | 8 |
| <i>L3</i> | <i>P1</i> | <i>P3</i> | <i>R2</i> | <i>R3</i> | 9 | 11 |
| <i>L4</i> | <i>P3</i> | <i>P4</i> | <i>R2</i> | <i>R4</i> | 12 | 13 |
| <i>L5</i> | <i>P4</i> | <i>P2</i> | <i>R2</i> | <i>R3</i> | 14 | 15 |
| <i>L6</i> | <i>P4</i> | <i>P3</i> | <i>R3</i> | <i>R4</i> | 16 | 20 |

| <i>List of the regions</i> | | |
|----------------------------|-------|---------------------|
| No. | Label | Boundary start line |
| <i>R1</i> | 0 | + <i>L1</i> |
| <i>R2</i> | 5 | + <i>L3</i> |
| <i>R3</i> | 9 | - <i>L5</i> |
| <i>R4</i> | 0 | - <i>L6</i> |

| | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|
| Address | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Coord. | (10,24) | (11,26) | (30,25) | (30,23) | (30,23) | (29,10) | (10,11) |
| Address | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Coord. | (10,24) | (10,24) | (17,18) | (17,17) | (17,17) | (24,20) | (24,20) |
| Address | 15 | 16 | 17 | 18 | 19 | 20 | |
| Coord. | (30,23) | (24,20) | (23,17) | (12,11) | (16,16) | (17,17) | |

both pointers are zero. The next two columns contains pointers to the regions lying to the right and to the left sides of the line. The last two columns contain pointers to the beginning and the ending of the subsequence of the metric list as a sequence of coordinate pairs. Each pair corresponds to an end point of a digital straight segment of the line. Each coordinate pair may be accompanied by additional parameters which are used for the exact reconstruction of the DSS running through the given end points. These parameters are not shown in the example of Table 1, to make the presentation simpler.

It may be seen from Table 1 that the coordinates of a branching point which is incident with N lines are repeated in the metric list N times. This is necessary since the coordinates are accompanied by the additional parameters which are different for the N different DSS meeting at the point.

The first column in the list of regions contains the identifiers of the regions. The next column contains the labels (gray values) of the regions. The last column contains a pointer to a line belonging to the boundary of the region. Starting from this line in the proper direction one may reconstruct the complete sequence of the lines composing the boundary. The boundary is directed in such a way that the region is always lying to the right-hand side of the boundary. The minus sign indicates that the starting line should be traversed from the end to the beginning to obtain the correct direction of the boundary.

6. APPLICATIONS

The described technique has been successfully used since 1985 for solutions of various problems in image analysis and computer graphics. The basis of this work is a computer program for automatically converting a raster image into a cell list. The image must be previously segmented into not more than a few hundred "uniform" regions. Each pixel of a region must have the same label, different from the labels of the adjacent regions. The number of labels may be up to 256.

The program scans the image line by line, looks for new boundaries, and starts the boundary tracing procedure which is combined with the recognition of digital straight segments. Both procedures write records into the cell list whose structure was described in the previous section.

To accelerate the tracing, the program has been implemented on the image processing system ROBOTRON A6472 [15] containing a fast pipeline processor which is able to label all 0-elements of the given image of $512 \times 512 \times 8$ bits by the "direction change labels" in a time of less than 1 s. The labels indicate the direction of the next step when tracing the boundary. They also indicate the branching points and the "north-west corners" which are the only possible locations of new boundaries when scanning the lines from the upper left to the lower right corner of the image. Due to the direction change labels, the boundary tracing which cannot be realized by the "pseudo-parallel" processor of the A6472 becomes rather fast: the cell list for an average image of $512 \times 512 \times 8$ bits containing about 150 regions is produced by a rather slow version of a mini-computer similar to the PDP-11 in about 80 s. The cell list has a rather high data compression rate; e.g., when using it for encoding technical drawings the compression ratio is about 20 times.

The cell lists produced in this way make any topological and geometrical analysis of the images especially efficient and simple. Adjacent regions, regions contained in each other, lines and points incident with each other or with some regions may be directly found in the lists. The distance between two points, the area of a region, the perimeter of a boundary, the angle between two line segments, etc. may be easily

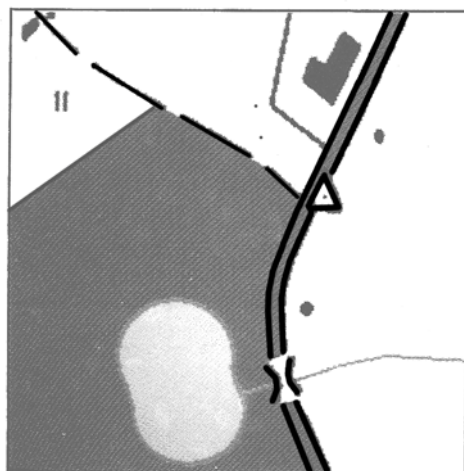


FIG. 13. Example of an image reconstructed from a cell list

calculated by the usual well-known formulae dealing with coordinates. All geometrical transformations, e.g., translation, scaling, rotation, etc. may be performed by recalculating all the coordinates of the 0-cells and the intermediate points according to the well-known formulae.

To obtain the transformed image in the raster representation the cell list may be converted back into the raster format. For this purpose the DSS composing the boundaries are being drawn in a refresh memory of the A6472 system and then the regions are being filled with the original labels. The whole procedure takes about 15 s for a 512×512 image of 150 regions when using the above-mentioned slow mini-computer (14 s are spent for drawing the boundaries by the slow host computer and about 1 s for filling the regions by the pipeline processor). An example of an image exactly reconstructed from a cell list having the total length of 3922 bytes is shown in Fig. 13.

In the cases when the boundaries in the original image are eroded by noise they must be smoothed. One of the most efficient ways of doing this is the polygonal approximation of the boundaries with an arbitrary predefined tolerance. The use of the cell list enables one to realize a fast approximation procedure: there is no need to perform the calculations for all boundary elements; it suffices to do this for the end points of the DSS which are stored in the metric list. In a similar way the determination of the middle axis of a stripe may be performed much faster when analyzing the coordinates of the DSS end points rather than all the boundary elements. This procedure has proved to be also much more precise and flexible than that of erosion.

The most efficient applications of the cell lists are in the object recognition and structural image analysis [16]. The possibility of analyzing the coordinates of polygon vertices instead of carrying out procedures similar to the mask matching makes the recognition extremely tolerant to geometrical transformations of the objects to be recognized and to small distortions of the boundaries. The recognition is very reliable and relatively fast: the recognition time by a mini-computer similar to PDP-11 is about 50 ms per object.

The realized recognition technique [16] is a generalized version of the isomorphism of labeled subgraphs. The most important modifications are the following:

(a) The structure of an image is described by means of a cellular complex (more precisely, block complex) rather than by a graph.

(b) The "label" of a region in the image to be analyzed is represented by the full set of its topological and metric features contained in the cell list. Such a description is geometrically exact which means that the region and its neighborhood may be exactly reconstructed starting from the features.

(c) The concordance between the image to be analyzed and the prototypes of the object classes to be recognized is being tested by means of certain predefined one- and two-place predicates of the regions and other parts of the image. The predicates are rather various and different for different object classes. They are specified in the definition of a class.

(d) The definitions of the classes to be recognized are organized hierarchically: the classes of more complicated objects are defined by means of a composition of certain elementary objects which have been defined previously.

(e) The definitions of the classes are determined interactively in the form of special tables ("knowledge bases") specifying the necessary one- and two-place predicates which the features of the elementary objects must satisfy to compose an object of the class to be defined.

This technique has been successfully used for the recognition of handwritten characters and of some model images having the nature of children's drawings [16].

6. CONCLUSION

The application of the well-known notions of the topology of complexes brings many advantages for precisely and economically describing the images, for fast and topologically precise geometrical transformations of images, and for implementing flexible and reliable methods for structural image analysis.

ACKNOWLEDGMENTS

The author wishes to express his deepest appreciation to Dr. Wolfgang Jansen from the Central Institute for Cybernetics (Berlin) and Dr. Ludwing Stammeler from the University of Halle for many important recommendations and valuable discussions.

REFERENCES

1. J. R. Munkers, *Elements of Algebraic Topology*, Addison-Wesley, Menlo Park, CA, 1984.
2. A. Rosenfeld, Connectivity in digital pictures, *J. ACM* **17**, 1970, 146-160.
3. T. Pavlidis, *Structural Pattern Recognition*, Springer-Verlag, New York, 1977.
4. V. A. Kovalevsky, Discrete topology and contour definition, *Pattern Recognit. Lett.* **2**, No.5, 1984, 281-288.
5. A. Rosenfeld, Adjacency in digital pictures, *Inform. and Control* **26**, 1974, 24-33.
6. H. J. Feng and T. Pavlidis, *IEEE Trans. Circuits Systems* **22**, 1975, 427-439.
7. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1982.
8. H. Elliott and L. Srinivasan, An application of dynamic programming to sequential boundary estimation, *Comput. Graphics Image Process.* **17**, No.4, 1981, 291-314.
9. R. Klette and K. Voss, Theoretische Grundlagen der digitalen Bildverarbeitung. IV. Orientierte Nachbarschaftsstrukturen, *Bild und Ton* **39**, No.7, 1986, 213-219.
10. V. A. Kovalevsky, Discrete geometry and image processing. In: *Proceedings, Int. Conference on Computers, Systems and Signal Processing*, Bangalore, India, Dec. 9-12, 1984, 934-938.
11. V. A. Kovalevsky, On the topology of digital spaces. In: *Proceedings, Digital Image Processing*, Fleeth/Mirrow (GDR), April 1986, Technical University of Dresden 1987, 56-77.
12. V. A. Kovalevsky, The topology of cellular complexes as applied to image processing. In: *Proceedings, Second Int. Conf. CAIP'87*, Wismar (GDR), 1987. Akademie-Verlag, Berlin, 162-173.
13. W. Jansen, personal communication.
14. V. A. Kovalevsky and A. Wiczorek, *Fast Recognition of Digital Straight Segments*, Technical Report 86-73-03, Institute of Cybernetics and Information Processing, Berlin, GDR, 1986.
15. V. Kempe, B. Rebel, and W. Wilhelmi, The interactive image processing console A6471. In: *Proceedings, 6th International Conference on Pattern Recognition*, Munich, 1982, 607-610.
16. V. Kovalevsky, Structural image analysis. In: *Proceedings, 8th International Conference on Pattern Recognition*, Paris, 1986, 358-368.