Pattern Recognition Letters 11 (1990) 275–284
North-Holland

April 1990

Quantification of tissue sections
Graph theory and topology as modelling tools

Karsten RODENACKER and Paul BISCHOFF
Gesellschaft für Strahlen- und Umweltforschung mbH, Institut für Strahlenschutz, D-8042 Neuherberg, FRG

Received 15 March 1989
Revised 28 June 1989

Abstract: Tissue sections consist of distributed objects, the nuclear profiles. A methodological tool box is presented, based on graph theory and topology, to describe formally the arrangement of such objects, to define agglomerations of objects in a hierarchical manner and to compute quantitative features for every object. For a set of tissue sections (conisates) from the *cervix uteri*, a model description and an evaluation are given.

Key words: Morphology, morphometry, graph theory, topology, histometry, modelling, pathology, image processing.

Introduction

In pathology, the analysis of the structure of tissue sections plays an important role for any diagnostic process. The arrangement of the cells with their nuclei, the very constituents of the tissue, denotes a tissue structure or texture. The latter and the appearance of certain cell types in certain positions of the tissue build the basis of any diagnosis in visual pathology. To quantify these arrangements and variations, an adequate and interpretable model of tissue has to be designed. The term *model* means an analytical and/or algorithmic description, which might be motivated by the formation of the tissue, e.g. epithelial type, or by the changes resulting from a certain disease. This procedure can be called heuristical.

In this paper applications of graph theory and topology are shown on the basis of a hierarchical structure. Starting from the digitized image area (*pixel* representation), a hierarchy is built up which allows to define and to denominate the objects of interest and to analyse and to interpret them according to different points of view.

The implemented methods are illustrated on conisates of the *cervix uteri*. In these cases, it is of special importance to recognize and to discriminate a *severe dysplasia*, as a certain type of tissue disorder. The latter builds the cutting point for either surgical or less invasive interventions.

The outlined methods allow us to divide automatically whole pictures or subareas of pictures into *zones*. The latter can be measured and classified. E.g. this can be by counting of special objects, e.g. mitotic cells, inside certain zones and the weighting for a final diagnosis.

The paper describes more methodoligical approaches of Chaudhuri et al. (1988) and Rodenacker et al. (1988) and is based to a limited extent on work of Kayser & Höffgen (1988, 1984). Comparable exploits were presented in the medical field by Preston (1981) and Prewitt (1979), followed up by Bartels (1987) and Weber et al. (1988). More generally there are approaches referenced by Toussaint (1988) and Ahuja & Schachter (1983). Many ideas are directed by the book *On Growth and Form* (Thompson, 1961, 1917). The hierarchical structuring is based on the work of Granlund (1979),

Rosenfeld (1979) and Mesarovic (1970). The definition of graphs is similar to approaches of Serra (1982) and Klette & Voss (1985).

Model description

Starting from the (pattern) recognition paradigm, consisting of the working steps
- definition,
- measurement and
- classification of the objects of interest,

a certain hierarchy of objects has to be defined for which the recognition paradigm is valid on every level. Beginning with a rudimentary medical terminology, object hierarchy levels could be characterized as *cell* or *nucleus* level. Higher hierarchical levels could then for example be called *basal*, *intermediate* and *superficial area* or *plaster epithelium* (Oberholzer, 1983). An important premise here and in the following is, that *the parts* (pixel, tissue, cells, areas) *constitute the whole*. Other possible concepts of modelling are the hierarchy *chromatin*, *nucleoli*, *nucleus*, *cell*, *cell heap*, *part of organs* and eventually *the organs itself* (Kayser, 1988).

In general, a hierarchy is defined by a sequence of *hierarchy levels* HE_n, the *objects* $\boldsymbol{O}_n$ of each level and the *relation* $\boldsymbol{R}_n \subseteq \boldsymbol{O}_n \times \boldsymbol{O}_n$ between objects. Additionally for this type of hierarchy, the objects $\boldsymbol{O}_n$ have to be defined as subsets of objects of underlying hierarchy levels HE_k, $k < n$, by means of certain selection criteria and under application of the appropriate relationships. For further purposes every object is marked and flagged, respectively by a set of mostly quantitative *features*. The representation of relations is similar to the mathematical formulation of graphs (Harary, 1969). Objects may be represented by *nodes* and the relationship between objects by *edges*.

The definition of the relation $\boldsymbol{R}_n$ for hierarchy level HE_n reflects the model concept of the objects $\boldsymbol{O}_{n+1}$ of HE_{n+1}. In the next section, the concepts for the definition of objects, relations as well as the necessary graph operations are outlined. Graph operations are comparable with operations in set theory and mathematical morphology (Serra, 1982). The descriptors and operators are chosen accordingly. The mathematical terminology follows (Dieudonné, 1969).

Definition of objects and relations

Pixel hierarchy HE_1

Let a digital image $\boldsymbol{B}$ consist of rows × columns *pixels* (picture points) $P_{x,y}$ with a pixel value $p_{x,y}$ at position (x, y) with $x \in$ [1,columns] and $y \in$ [1,rows] (Figure 1).

Definition. The *objects* $\boldsymbol{O}_1$ of HE_1 are the *pixels*:

$$O_1^{x,y} := P_{x,y}. \tag{1}$$

The *relation* $\boldsymbol{R}_1$ is defined for every pair of objects $O_1^{x,y}, O_1^{x',y'} \in \boldsymbol{O}_1$ by a certain neighborhood criterion N

$$O_1^{x,y}\, \boldsymbol{R}_1\, O_1^{x',y'} := N(P_{x,y}, P_{x',y'}) \tag{2}$$

with

$$N(P_{x,y}, P_{x',y'}) :== (|x-x'|+|y-y'| \leqslant 1) \quad (2.1)$$

for 4-neighborhood or

$$N(P_{x,y}, P_{x',y'}) :== (\max(|x-x'|, |y-y'|) \leqslant 1) \tag{2.2}$$

for 8-neighborhood. *Features* of the objects $\boldsymbol{O}_1$ are their pixel value $p_{x,y}$ and their coordinates (x, y) only.

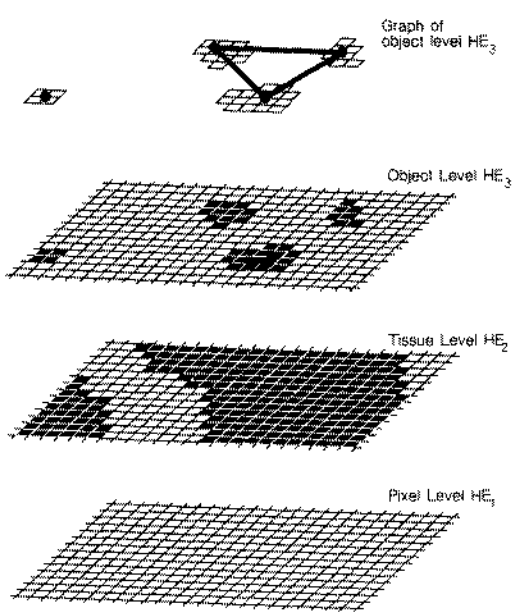

Figure 1. Illustration of the hierarchical level concept.

Tissue hierarchy HE_2

The objects O_2 are the *background* and the *tissue*. A selection criterion is applied on every object $O \in O_1$ to decide whether it belongs to the background or to the tissue. A relation between these objects is not necessary. In practice, the tissue area is interactively chosen as the area of interest for the succeeding steps of analysis. Features may be the area (the number of objects from O_1 per object from O_2) or the number of R_1-connected components (see later).

Object hierarchy HE_3

Here, the objects O_3 are the set of nuclear profiles inside the tissue. Pixels of single nuclear profiles are R_1-connected. Let for simplicity the pixels of nuclear profiles be segmentable by a certain selection criterion A. Nuclear profiles have no holes. A formal definition is given below:

Definition. The relation R_1' is defined by

$$O_1^{x,y}, O_1^{x',y'} \in O_1,\ O_1^{x,y} R_1' O_1^{x',y'} :\Leftrightarrow$$
$$\exists \text{ a sequence of objects } O_1^{x_i,y_i},\ i=1,\dots,n,$$
$$\text{with } O_1^{x_1,y_1} = O_1^{x,y} \text{ and } O_1^{x_n,y_n} = O_1^{x',y'} \text{ with}$$
$$O_1^{x_i,y_i} R_1 O_1^{x_{i+1},y_{i+1}} \text{ for } i=1,\dots,n-1. \quad (3)$$

Objects in relation R_1' are called R_1-*connected*.

Lemma. *R_1' is an equivalence relation.*

Proof. Reflexivity ($x R_1' x$) is clear by definition of relation R_1, formulae (2.1) and (2.2), symmetry ($x R_1' x' \Rightarrow x' R_1' x$) follows directly from (3) renumbering the sequence of objects from n down to 1 and transitivity ($x R_1' x', x' R_1' x'' \Rightarrow x R_1' x''$) follows from merging the sequences from x to x' and from x' to x''. □

Hence any subset $N \subseteq O_1$ is partitioned into disjunct classes N/R_1' with respect to relation R_1'.

Definition. Let N be the set of pixels (O_1) selected by the selection criterion A inside tissue (O_2). Then the equivalence classes of N with respect to R_1' are the *objects* O_3, the nuclear profiles:

$$O_3 := N/R_1'. \quad (4)$$

To fill up holes, each object $O_3 \in O_3$ is merged with all pixels of the background R_1-connected to any pixel of O_3, but not to any other object in O_3. These pixels are called *inside* the object O_3.

The *relation* R_3 is inherited from HE_1 in the following way: Let $S \subseteq O_1$ be the set of *zones of influence* (ZOI) of the nuclear profiles O_3 with respect to the tissue O_2. S is a skeleton (Serra, 1982, pp. 375–423) and represents an O_3-object-connectivity preserving partition of the tissue. Each zone of influence is related to one and only one nuclear profile with

$$\forall O \in O_2\ \exists_1 \text{ZOI}_O \in S \text{ with } O \subset \text{ZOI}_O. \quad (5)$$

The *relation* R_3 is defined as

$$\forall O, O' \in O_3:\ O R_3 O' :\Leftrightarrow \text{ZOI}_O R_1 \text{ZOI}_{O'}. \quad (6)$$

In other words, relation R_3 is true for two objects (4) if and only if there exists at least between two pixels of the corresponding zones of influence the relation R_1. With this method the pixel oriented relation R_1 is inherited by the object oriented relation R_3. The *features* are listed in Table 1.

This rather broad description also shows how, from a topological point of view, the amorph set of picture points $P_{x,y}$ is structured into objects *and* topologized by the definition of a neighborhood relation (2). Subsequently this structure and topology is inherited to the next hierarchy level. Beside the mere locally oriented neighborhood relation, other relations, derived from certain object

Table 1
List of computed features HE_3

	Cell nucleus section	
2.	Area	A
3.	Perimeter	P
4.	Point of gravity x	KX
5.	Point of gravity y	KY
6.	Orientation	THETA
7.	Mean of grey values	M1
	Zone of influence	
8.	Area	AZ
9.	Perimeter	PZ
10.	Point of gravity x	KXZ
11.	Point of gravity y	KYZ
12.	Orientation	THETAZ

feature constellations, can be defined arbitrarily.

The graph representation, outlined in the following, is quite feasible for the definition and combination of objects and hierarchy levels after leaving the more image oriented levels.

Graph representation

Definition. A *graph* is a 2-tuple $G := (V, E)$ with

$$\begin{aligned} V &= \text{set of nodes} \quad \text{and} \\ V \times V \supseteq E &= \text{set of edges.} \end{aligned} \tag{7}$$

The relation neighborhood for cell or nucleus profiles in histology is considered symmetrical. Hence there exists for each edge (v_i, v_k) another edge (v_k, v_i) between nodes v_i and v_k. Represented in an adjacency matrix the upper and lower triangle matrix are transposed. However the definitions given in the folowing are valid for the more general case of directed graphs.

In all definitions, the operators $\cap$, $\cup$ and $/$ are used as symbols for set union, set intersection and set difference. Additionally pr_i denotes the projection

$$\mathrm{pr}_i(X_1 \times \cdots \times X_n) \to X_i, \quad 1 \leqslant i \leqslant n, \; n \in \mathbb{N}, \tag{8}$$

which results in the i-th component of the cross product.

Definition. A graph $G_1 = (V_1, E_1)$ is *included* in another graph $G = (V, E)$:

$$G_1 \subseteq G \quad :\Leftrightarrow \quad V_1 \subseteq V, \; E_1 \subseteq E. \tag{9}$$

Definition. Let $G_i = (V_i, E_i)$, $i = 1, \ldots, m \in \mathbb{N}$ a set of graphs. A graph $G = (V, E)$ is called *base graph*, if

$$\forall i = 1, \ldots, m \in \mathbb{N} \; G_i \subseteq G. \tag{10}$$

An example of a base graph is the graph $G_A = (V_A, E_A)$, constructed from all subgraphs $G_i = (V_i, E_i)$, where every node is linked to each other node; hence

$$V_A = \bigcup V_i, \qquad E_A = V_A \times V_A. \tag{11}$$

In the following let $G = (V, E)$ be a base graph for graphs $G_i = (V_i, E_i)$, $i = 1, 2$, if not otherwise mentioned.

Boolean graph operations

Definition. The *union* $\cup$ of two graphs $G_i = (V_i, E_i)$, $i = 1, 2$ (Figure 2c) is defined as

$$G_1 \cup G_2 := (V_1 \cup V_2, E_1 \cup E_2). \tag{12}$$

Definition. The *intersection* $\cap$ of two graphs $G_i = (V_i, E_i)$, $i = 1, 2$ (Figure 2f) is defined as

$$G_1 \cap G_2 := (V_1 \cap V_2, E_1 \cap E_2). \tag{13}$$

Definition. The operation *join* $+_G$ of two graphs $G_i = (V_i, E_i)$, $i = 1, 2$ with respect to graph G (Figure 2d) is defined as

$$\begin{aligned} G_1 +_G G_2 := (V_1 \cup V_2, E \cap [V_1 \times V_2) \cup (V_2 \times V_1)] \\ \cup E_1 \cup E_2). \end{aligned} \tag{14}$$

This operation can be considered as an extension of the union of graphs, where additionally the edges are added from the base graph G, whose cor-

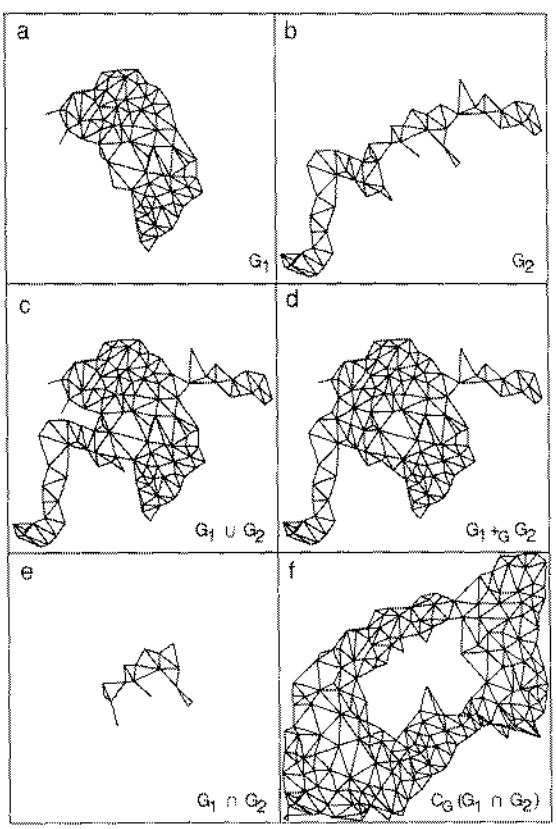

Figure 2. Base graph operations with base graph G in Figure 5b. (a) Graph G_1. (b) Graph G_2. (c) $G_1 \cup G_2$. (d) $G_1 +_G G_2$. (e) $G_1 \cap G_2$. (f) $C_G(G_1 \cap G_2)$.

responding nodes are both in $V_1 \cup V_2$ (Harary, 1969).

Definition. The *complement* C_G of a graph $G_1 = (V_1, E_1)$ with respect to graph G (Figure 2e) is defined as

$$C_G(G_1) := (V \setminus V_1, E \cap [(V \setminus V_1) \times (V \setminus V_1)]). \quad (15)$$

This definition avoids the existence of graphs with edges, whose corresponding nodes are not in the set of nodes. From the definition (15) it follows

$$G_1 \cup C_G(G_1) \neq G, \quad (15.1)$$

$$G_1 +_G C_G(G_1) \neq G, \quad (15.2)$$

and additionally in contrast to the classical definition of a complement

$$G_1 \neq C_G(C_G(G_1)). \quad (15.3)$$

There is in (15.2) and (15.3) equality for a graph $G_1 = (V_1, E_1)$ and base graph $G = (V, E)$, with

$$E_1 = E \cap (V_1 \times V_1),$$

hence is *closed* (see later).

Definition. The set of *G-reachable* nodes $N_G(G_1)$ of $G_1 = (V_1, E_1)$ (Figure 3a) is defined as

$$N_G(G_1) := \mathrm{pr}_2([V_1 \times (V \setminus V_1)] \cap E) \cup \mathrm{pr}_2([(V \setminus V_1) \times V_1] \cap E) \quad (16)$$

and considered as a graph with an *empty* set of edges.

Definition. The *closure* of a graph $G_1 = (V_1, E_1)$ with respect to G (Figure 3b) is defined as

$$A_G(G_1) := (V_1, E \cap (V_1 \times V_1)), \quad (17)$$

or, using the join operation (14),

$$A_G(G_1) := G_1 +_G G_1. \quad (17.1)$$

Derived graph operations

Based on the above defined boolean graph operations the following operations, oriented at mathematical morphology (Serra, 1982), can be defined.

Definition. The *dilation* of a graph $G_1 = (V_1, E_1)$ with respect to G (Figure 3c) is:

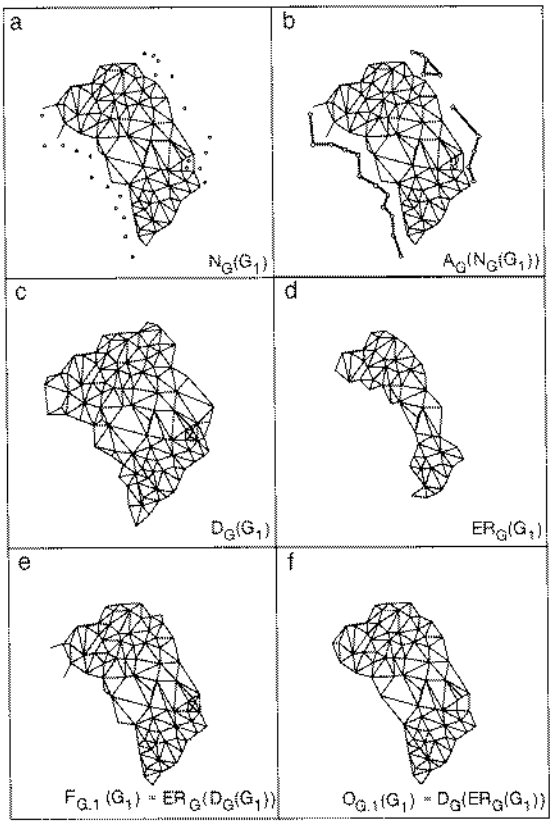

Figure 3. Derived graph operations with base graph G in Figure 5b. (a) The set of reachable nodes $N_G(G_1)$. (b) The closure $A_G(N_G(G_1))$. (c) The dilation $D_G(G_1)$. (d) The erosion $ER_G(G_1)$. (e) The closing $F_{G,1}(G_1)$. (f) The opening $O_{G,1}(G_1)$.

$$D_G(G_1) := A_G(G_1 +_G N_G(G_1)). \quad (18)$$

The graph G_1 is thus dilated by the *G-reachable* nodes with their corresponding edges (Klette et al., 1985).

Definition. The *erosion* of a graph $G_1 = (V_1, E_1)$ with respect to G (Figure 3d) is:

$$ER_G(G_1) := C_G(D_G(D_G(G_1))). \quad (19)$$

Definition. The *opening* of a graph $G_1 = (V_1, E_1)$ with respect to G with radius r, the degree of reachability, (Figure 3e) is:

$$O_{G,r}(G_1) := D_G(\cdots D_G(ER_G(\cdots ER_G(G_1)\cdots))\cdots), \quad (20)$$

where ER_G and D_G are r-times applied.

Definition. The *closing* (not *closure* (17)) of a graph $G_1 = (V_1, E_1)$ with respect to G with radius r (Figure 3f) is:

$$F_{G,r}(G_1) := ER_G(\cdots ER_G(D_G(\cdots D_G(G_1)\cdots))\cdots). \quad (21)$$

The dilation as defined above (18) is comparable with the same in mathematical morphology if the neighborhood relations N used (2.1, 2.2) are considered as graphs and applied in the classical definition of the dilation as a so called *structuring element*

$B =$  or $B =$ 

for 4- or 8-neighborhood, respectively.

Graph hierarchy HE_4

In the following two different object classes $\boldsymbol{O}_{41}$ and $\boldsymbol{O}_{42}$ are defined. $\boldsymbol{O}_{41}$ bases on the model of a *parent-child relation*, which is called *layer partition* (L). $\boldsymbol{O}_{42}$ can be considered as a coarse layer partition. It is called region partition (R). The hierarchical interdependencies are shown in Figure 4.

Hierarchy level HE_{41}

As with every human parent-child relationship at least one 'parent' or one child has to be designated as the starting root. For epithelial tissue sections the basal cell layer represented by their nuclear profiles (objects of HE_3) is usable as starting object. The choice of the latter represents the selection criterion.

Definition. Let $V_1 := \{\text{nuclear profiles of the basal layer}\} \subset \boldsymbol{O}_3$ be the starting set of objects, which becomes by application of the closure A_G the starting graph $L_1 = A_G((V_1, \{\,\}))$.

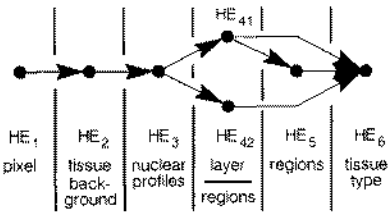

Figure 4. Hierarchical interdependencies.

The following iteration with respect to the base graph G:

$$\begin{aligned} i = 1: \quad & L_1 = \text{interactively selected basal layer}, \\ & S_1 = L_1, \\ i = i+1: \quad & L_i = C_{D_G(L_i)}(S_i), \\ & S_{i+1} = S_i +_G L_i \end{aligned} \quad (22.1)$$

terminates by exhausting the set of existing nodes V or by reaching stability $S_i = S_{i+1}$ respectively. The last value of i represents the number of layers generated concerning the start graph L_1 and the base graph G (Figure 5b).

The set of generated layers construct the set of *objects* $\boldsymbol{O}_{41}$ of hierarchy level HE_{41}. *Relation* $\boldsymbol{R}_4$ is defined by the sequence of layers

$$\boldsymbol{O}_{41} :== \{L_1, \ldots, L_n\} \quad (22.2)$$

with

$$L_i R_4 L_j \;:\Leftrightarrow\; |i-j| = 1 \quad \forall L_i, L_j \in \boldsymbol{O}_{41}. \quad (23.3)$$

The latter represents an order relation in terms of the generating sequence. Figure 5a shows the objects $O \in \boldsymbol{O}_3$ in black with the borders of the corresponding zones of influence ZOI_O. Figure 5b

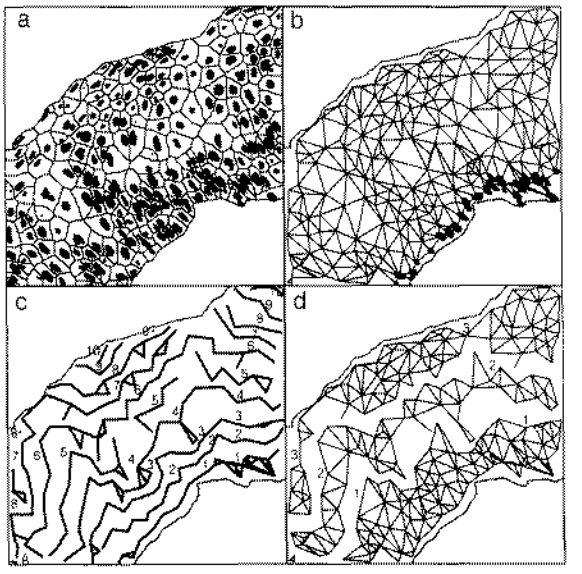

Figure 5. A tissue section processed. (a) Binary image with objects $\boldsymbol{O}_3$ and borders of zone of influences. (b) Relation $\boldsymbol{R}_3$ of objects $\boldsymbol{O}_3$, the base graph G, as well as the chosen basal layer. (c) Objects $\boldsymbol{O}_{41}$ with relation $\boldsymbol{R}_3$ (layers). (d) Objects $\boldsymbol{O}_{42}$ with relation $\boldsymbol{R}_3$ (*regions*).

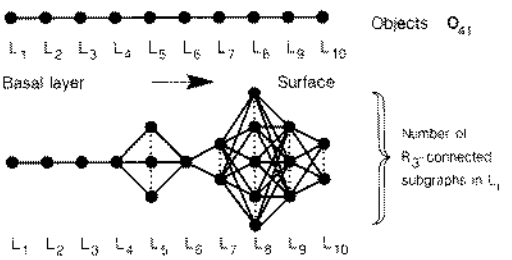

Figure 6. The graph of hierarchy level HE_{41}.

shows the basegraph G derived from HE_3 and the interactively chosen objects of the basal layer, the set V_1. The layers, objects O_{41}, represented in Figure 5c, can easily be shown with their relation R_{41} as a graph (Figure 6).

Of this set of objects O_{41}, *features* are calculated as mean, standard deviation (SD), minimum and maximum of the corresponding object features O_3 (see Table 1) extended by graph features like mean and SD of distances of nodes, number of edges, nodes, R_3-connected objects (similar definition (3)) and articulation points as well as the maximal extension of the graph (Rodenacker et al., 1988).

Such a partition may reflect the behavior of differentiation of the tissue starting from the basal layer.

Hierarchy level HE_{42}

A second way of object definition for another hierarchy level is based on the partition of the base graph into a number of preferably connected regions. A certain marker variable of each object $O \in O_3$ is e.g. accumulated and divided by the number of regions. Then the partition algorithm tries to fill up the different regions with objects balanced in terms of the marker variable. The sequence of included objects depends on a predefined order over the objects O_3. According to the methods of pathologists the number of regions here is chosen to be three. The marker variable is the area of the zone of influence ZOIA (Table 1) of each object. The order over the objects O_3 is defined as following:

Definition. Let the objects O_3 be ordered using hierarchy HE_{41} in the following way:

$$\forall O_{ij} \in L_i: \quad O_{i1}, \ldots, O_{im_i}, \tag{23.1}$$

the defined layers (22.1), with any existing order.

Now all objects $O \in \bigcup L_i$, $i = 1, \ldots, n$ can be ordered by

$$O_{11}, \ldots, O_{1m_1}, O_{21}, \ldots, O_{2m_2}, O_{31}, \ldots, O_{nm_n}. \tag{23.2}$$

This sequence of objects can then be cut off into a number of intervals using a certain condition. The aim of the condition applied here was to generate regions with at least similar areas in terms of zones of influence (ZOIA, Table 1). Objects in the given order are *joined* as far as $\sum \text{ZOIA}/3$ (three regions) for each region is reached.

The resulting region partition, the set of objects O_{42}, shown for one section in Figure 5d, reflects the way in which pathologists differentiate epithelial tissue into *basal*, *intermediate* and *superficial areas*. The relation chosen is similar to hierarchy level HE_{41}.

Hierarchy level HE_5

Another way to define regions can be performed joing (14) an appropriate number of neighbored layers (O_{41}). This would also correspond with a view of tissue, in which the cells, in our model approach represented by nuclear profiles (O_3), are arranged in layers and a certain number of subsequent layers constitute the basal, intermediate and superficial areas. A procedural description could be:

(i) Partition of the set of objects O_{41}, ordered by (24.2), into intervals $[O_3^1, O_3^i], [O_3^{i+1}, O_3^j], [O_3^{j+1}, O_3^n]$ as described for HE_{42}, see above,

(ii) The *objects* O_5 are the closures of the object intervals with respect to base graph G.

The *relation* R_5 is similar to R_{41}.

Hierarchy level HE_6

The hierarchy levels defined up to now are shown in Figure 6 with references. The next level HE_6 is either based on HE_{41} or HE_{42} or HE_5. Its objects again reflect an increased degree of abstraction. They can now be called *tissue* or *layer types*.

Definition. Let O_6 be the set of linear regression functions f_{xxx} derived per tissue from the set of objects O_{41}, O_{42} or O_5 for each feature *xxx* of the

used hierarchy level, defined by

$$O_6 := \{f_{xxx}(j) := A_{xxx} + B_{xxx} \cdot j\}. \tag{24}$$

Each tissue consisting of the chosen object set results in one function f_{xxx} per feature *xxx*. Obviously, the relations are chosen as an *order relation* accordingly to define the sequence of objects for regression computation. The features of this object are the coefficients of the regression function A_{xxx} and B_{xxx}.

Definition. The regression function parameters are calculated by

$$B_{xxx} = \frac{n \sum x_j \cdot j - \sum x_j \cdot \sum j}{n \sum x_j^2 - (\sum x_j)^2},$$
$$A_{xxx} = \frac{\sum j - B_{xxx} \sum x_j}{n}, \tag{25}$$

with x_j the value of feature *xxx* of object *j*.

Application example

Preprocessing and segmentation HE_1, HE_2 and HE_3

In the reported approach, 32 fields from 7 routine specimens (conisates of *cervix uteri*) were interactively chosen. For different grades of dysplasia, typical areas were individually diagnosed, scanned with an AXIOMAT microscope (50× objective) and digitized via a TV-camera using a 546 nm narrow bandpass filter. A picture covers a square area of 1/4 µm edge length with 512×512 pixels. This, for histological purposes, unusually high magnification is mainly necessary for the segmentation as well as for the computation of the nuclear profile orientation (Table 1, feature THETA). On the image processing system IPS (Kontron, Eching, FRG), a segmentation routine has been developed for the almost automated discrimination of the nuclear profiles. Subsequently the section area is interactively marked, the zones of influence of the nuclear profiles (Figure 5a) are computed (Arcelli, 1979) and the profiles belonging to the basal layer are marked (Figure 5b). Acceptable segmented images require only these interactions. Sections with dense distributed cells often demand heavy interactive improvement of the automatic segmentation of the nuclear profiles.

Feature extraction HE_{41}, HE_{42}, HE_5 and HE_6

The feature values listed in Table 1 are determined per sections from the nuclear profiles and the corresponding zones of influence (ZOI), the objects of HE_3. They represent the basis for all further evaluations. Graph feature values or feature values from objects of HE_4 or higher, respectively are computed as the mean value (M1), the standard deviation (M2), the maximum (MAX) and the minimum (MIN) from the underlying objects of HE_3. In addition, graph features are evaluated, e.g. the number of R_3-connected graphs of an object or the number of R_3-edges. Based on the euclidean distances of the centre of gravity, edge lengths can be derived as features.

Results

The results of the measured 32 section images will be illustrated only briefly. Figures 7 and 8 show two scatter plots of features of tissue types (HE_6) derived from layers (HE_{41}) and regions (HE_{42}). The analysis of single objects (HE_3) has been shown in (Burger et al., 1987). In the scatter

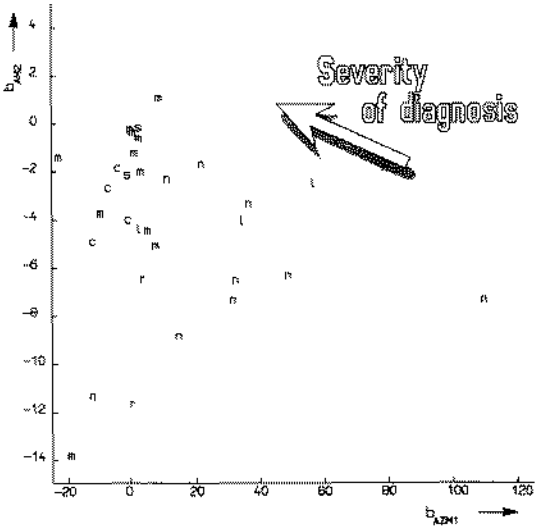

Figure 7. Plot of slope of feature AZM1 (mean of area of ZOI per layer) B_{AZM1} against slope of feature AM2 (standard deviation of aera of nuclear profile per layer) B_{AM2}.

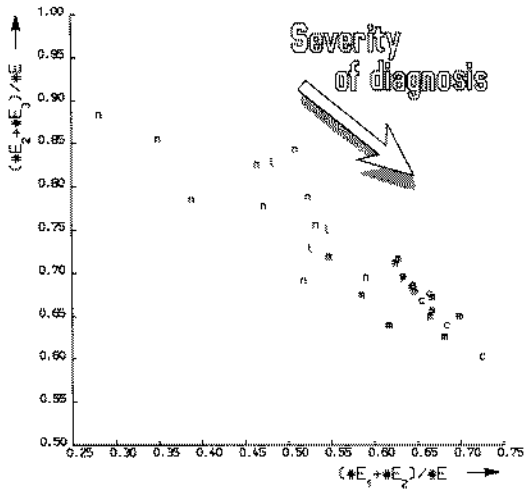

Figure 8. Plot of relative object frequency of first and second region $(\#E_1+\#E_2)/\#E$ against relative object frequency of second and third region $(\#E_2+\#E_3)/\#E$. $\#E_i$ is the total number of objects O_3 in region i.

plots, every letter **n, r, l, m, s, c** represents one section image. The abbreviations stand for the visual diagnoses: **n**ormal, **r**eserve cell hyperplasia, **l**ight, **m**oderate and severe dysplasia and **c**arcinoma in situ. Sections of type **r** were considered as **n**ormal. Additional diagnoses like **n/a** and **m/a** for **a**trophic as well as transitions like **l/m** or **s/c** were taken as **n**, **m**, **l** or **s**, respectively.

Figure 7 shows the slope B_{AM2} of the linear regression line of feature AM2, the standard deviation of the nuclear profiles A in one layer (HE_{41}), plotted against the slope B_{AZM1} of the linear regression line of mean of the corresponding area of zones of influence (AZM1). It can be recognized how the severeness of diagnosis decreases from top left to bottom right. The insufficient differentiation of cases of type **m**, **s** and **c** can partly be explained by the fact that the diagnostical scheme is based on arrangement characteristics *and* on object type appearances, like e.g. a mitotic cell in the intermediate region. However, the computed features depend *only* on the arrangement of nuclear profiles (AZM1) and the deviation of nuclear profile areas (AM2). Particularly, the comparison of visual appearance of sections with the resulting position in the scatter plot suggests that the medical diagnosis should be differentiated according to arrangement *and* object characteristics.

Figure 8 illustrates the features frequency of objects in region 1 and 2 ($\#E1+\#E2$) plotted against the same of region 2 and 3 ($\#E2+\#E3$) (Figure 5d) divided by the total number of objects ($\#E$) using the same set of section images. Here, the severity of diagnosis decreases from bottom right to top left. The two normal cases **r** would be misclassified at least as type **m**. They consist of extremely thin epithelial areas which cannot be partitioned reasonably.

Summary

The main goal of this paper is the description of methods based on graph theory and topology for the *gathering* and the *definition* of distributed objects. The methods are implemented in a PASCAL program as an external procedure of ILIAD (Erikson, 1982) on a VAXstation II/GPX (Digital, Maynards, MA, USA) using extensively the emulated Tektronix 4014 screen. It is possible to develop adequate routines interactively as well as to apply command files automatically. An application in histology is demonstrated. Starting with a base graph representing a neighborhood relationship between nuclear profiles, the definition of an hierarchy reflecting the architecture of sections of epithelial tissues is modelled. Another model reflects the diagnostic scheme of pathologists. A section area is divided into regions and the appearance of certain events, e.g. the occurrence of mitotic cells in different regions is registered. In this paper, the methods of partitioning section images accordingly are shown.

The described procedure is applicable for any spatial structure, not limited to medical problems or to 2-dimensional sections.

Acknowledgements

The authors would like to express their thanks to their colleagues Peter Gais, Uta Jütting and Michaela Aubele for help. For important stimulations, the preparation of the material and many discussions about the proceedings of pathologists, Prof. Dr. M. Oberholzer (Institute of Pathology, University Basel) and Prof. Dr. W. Gössner (Insti-

tute of Pathology, Gsf, München) are gratefully acknowledged. Last but not least we have to thank the head of the group Dr. Georg Burger for many valuable suggestions.

References

Ahuja, N. and B.J. Schachter (1983). *Pattern Models*. Wiley, New York.

Arcelli, C. (1979). A condition for digital points removal. *Signal Processing* 1, 283–285.

Bartels, P. (1987). Expert systems in histopathology and machine vision. Lecture held at 1st Int. Conf. on Artif. Intell. Sys. as Diagnostic Consult. for the Cytologic and Histologic Diagnosis of Cancer, Feb. 1–3, 1987, North Hollywood.

Burger, G., M. Oberholzer, K. Rodenacker and U. Jütting (1987). Tissue section analysis of cervical intraepithelial neoplasias. In: G. Burger, J.S. Ploem and K. Goerttler, Eds., *Clinical Cytometry and Histometry*, Academic Press, London, 531–533.

Chaudhuri, B.B., K. Rodenacker and G. Burger (1988). Characterization and featuring of histological section images. *Pattern Recognition Letters* 7, 245–252.

Dieudonné, J. (1969). *Eléments d'Analyse Vol. I, Fondements de l'analyse moderne*. Gauthier-Villar, Paris.

Eriksson, O., E. Bengtsson, T. Jarkans, B. Nordin and B. Stenkvist (1982). Tuning of an interactive software system for image analysis to quantitative microscopy, *Proc. Int. Symp. Medical Imaging and Image Interpretation* (ISMII 82), Berlin, 549–553.

Gössner, W. and M. Oberholzer (1988). Einführung: Gestaltwahrnehmung in der Histopathologie. In: G. Burger, M. Oberholzer and W. Gössner, Eds., *Morphometrie in der Zyto- und Histopathologie*, Springer, Berlin, 1–4.

Granlund, G.H. (1978). In search of a general picture processing operator. *Computer Graphics and Image Processing* 8, 155–173.

Harary, F. (1969). *Graph Theory*. Addison-Wesley, Reading, MA.

Kayser, K. and H. Höffgen (1984). Pattern recognition in histopathology by orders of textures. *Medical Inform.* 9, 55–59.

Kayser, K. (1988). Syntaktische Strukturanalyse in der Histopathologie. In: G. Burger, M. Oberholzer and W. Gössner, Eds., *Morphometrie in der Zyto- und Histopathologie*. Springer, Berlin, 164–178.

Klette, R. and K. Voss (1985). Theoretische Grundlagen der digitalen Bildverarbeitung II. Nachbarschaftsstrukturen. *Bild und Ton* 11, 325–331.

Mesarovic, M.D. (1970). *The Theory of Multilevel Hierarchical Systems*. Academic Press, London.

Oberholzer, M. (1983). *Morphometrie in der klinischen Pathologie*. Springer, Berlin.

Oberholzer, M., P. Dahlquen, W. Gössner and P.U. Heitz (1988). Aufgaben, Möglichkeiten und Grenzen einer quantitativen Pathologie. In: G. Burger, M. Oberholzer and W. Gössner, Eds., *Morphometrie in der Zyto- und Histopathologie*. Springer, Berlin, 118–131.

Preston, K. (1981). Tissue section analysis: Feature selection and image processing. *Pattern Recognition* 13, 17–36.

Prewitt, J.M.S. (1979). Graphs and grammars for histology: An introduction. *Proc. Third Annual Symp. Comp. Appl. and Medical Care*, Washington.

Rodenacker, K., B.B. Chaudhuri, P. Bischoff, P. Gais, U. Jütting, M. Oberholzer, W. Gössner and G. Burger (1988). Strukturbeschreibung und Merkmalsgewinnung in der Histometrie am Beispiel von Plattenepithelien. In: G. Burger, M. Oberholzer and W. Gössner, Eds., *Morphometrie in der Zyto- und Histopathologie*. Springer, Berlin, 179–199.

Rodenacker, K., P. Bischoff and B.B. Chaudhuri (1987). Featuring of topological characteristics in digital images. *Acta Stereologica* 6(3), 945–950.

Rosenfeld, A. (1979). *Picture Languages*. Academic Press, London.

Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.

Thompson, D. (1961, 1917). *On Growth and Form*, abridged edition by J.T. Bonner, University Press, Cambridge.

Toussaint, G.T. (ed.) (1988). *Computational Morphology*, North-Holland, Amsterdam.

Weber, J.E., P.H. Bartels, W. Griswold, W. Kuhn, S.H. Paplanus and A.R. Graham (1988). Colonic lesion expert system: Performance evaluation. *Analyt. Quant. Cytol. Histol.* 10(2), 150–159.