New algorithms based on the Voronoi Diagram applied in a pilot study on normal mucosa and carcinomas

J. Sudbø [a,*], R. Marcelpoil [b] and A. Reith [a]
[a] *Section for Digital Pathology, The Norwegian Radium Hospital, Oslo, Norway*
[b] *Université Joseph Fourier, Grenoble, France*

Accepted 30 October 2000

An adequate reproducibility in the description of tissue architecture is still a challenge to diagnostic pathology, sometimes with unfortunate prognostic implications. To assess a possible diagnostic and prognostic value of quantitiative tissue architecture analysis, structural features based on the Voronoi Diagram (VD) and its subgraphs were developed and tested.

A series of 27 structural features were developed and tested in a pilot study of 30 cases of prostate cancer, 10 cases of cervical carcinomas, 8 cases of tongue cancer and 8 cases of normal oral mucosa. Grey level images were acquired from hematoxyline-eosine (HE) stained sections by a charge coupled device (CCD) camera mounted on a microscope connected to a personal computer (PC) with an image array processor. From the grey level images obtained, cell nuclei were automatically segmented and the geometrical centres of cell nuclei were computed. The resulting 2-dimensional (2D) swarm of pointlike seeds distributed in a flat plane was the basis for construction of the VD and its subgraphs. From the polygons, triangulations and arborizations thus obtained, 27 structural features were computed as numerical values. Comparison of groups (normal vs. cancerous oral mucosa, cervical and prostate carcinomas with good and poor prognosis) with regard to distribution in the values of the structural features was performed with Student's t-test.

We demonstrate that some of the structural features developed are able to distinguish structurally between normal and cancerous oral mucosa ($P = 0.001$), and between good and poor outcome groups in prostatic ($P = 0.001$) and cervical carcinomas ($P = 0.001$).

We present results confirming previous findings that graph theory based algorithms are useful tools for describing tissue architecture (e.g., normal versus malignant). The present study also indicates that these methods have a potential for prognostication in malignant epithelial lesions.

Keywords: Graph theory, Voronoi Diagram, diagnostic pathology, prognostication tissue architecture, reproducibility, oral mucosa, oral carcinoma, carcinoma of the cervix, carcinoma of the prostate

Abbreviations[1]: CCD, Charged Coupled Device; DT, Delaunay Triangulation; GG, Gabriel Graph; HE, Hematoxylin-Eosine; IOD, Integrated optical density; MST, Minimum Spanning Tree; sd, Standard deviation (used in the Appendix); UT, Ulam Tree; VD, Voronoi Diagram; 2D, Two-dimensional.

1. Introduction

Several authors have demonstrated that subjective grading of malignant lesions is associated with poor reproducibility and accordingly with reduced prognostic power [1–5], although some indicate the opposite [6]. Recently, semiquantitative studies of tissue architecture have shown promising diagnostic and prognostic results [2,4,7–9]. Nevertheless, the problem of intra- and inter-observer variability persists. Over the last decades, several studies have quantitatively studied the relation of structure and function in biological systems, including pathologically altered tissues [10–15]. *In vitro* transformation studies have demonstrated that the addition of carcinogens to contact inhibited ordered fibroblast monolayer cultures results in loss of contact inhibition, with cells displaying striking criss-crossing growth patterns, where the degree of criss-crossing pattern may reflect the extent of oncogenic transformation [16–19]. These *in vitro* findings are an indication that the biological status of cells also is expressed in the tis-

*Corresponding author: Jon Sudbø, MD, DDS, Department of Pathology, Section for Digital Pathology, The Norwegian Radium Hospital, Montebello, 0310 Oslo, Norway. Tel.: + 47 22934230; Fax: + 47 22935627; E-mail: jon.sudbo@rh.uio.no.

[1]For a complete list describing the structural features, see Appendix.

Analytical Cellular Pathology 21 (2000) 71–86
ISSN 0921-8912 / $8.00

sue architecture. Hence, it is biologically meaningful to extract structural features from tissues for diagnostic and prognostic purposes, and to do this in a quantitative manner might improve the prognostic value in some tissues [20–32]. Previous findings in transitional carcinomas of the bladder indicate that graph theory based methods are useful tools in grading of malignant lesions [29], but a prognostic value was not demonstrated.

We have undertaken the present study in order to develop tools for fast, strictly quantitiative and reproducible tissue architecture analysis in epithelial tissues (squamous cell carcinomas from the prostate, cervix and oral cavity and normal oral mucosa) and to evaluate the diagnostic and prognostic potential of these methods in such tissues. By employing graphs such as the Voronoi Diagram (VD [Figs 1–4]) [33] and its subgraphs, the Delaunay Triangulation (DT [Fig. 2]), Minimum Spanning Tree (MST [Fig. 5]) [34–38], Ulam Tree (UT [Fig. 6]) [39] and the Gabriel Graph (GG [Fig. 7]) [40], the structural manifestations of cellular interactions in tissues may be quantified [41–47]. A total of 27 structural features were developed, taking into consideration the shape of individual structural entities (polygons, triangulations, arborizations), particularly derived from the VD; clusterings, particularly from the GG, and studying the order or randomness in the distribution of pointlike seeds, particularly derived from the UT and MST.

2. Materials and methods

2.1. Material

The biological material investigated consisted of 8 cases of normal oral mucosa obtained from surplus tissue after plastic surgery on the gingiva in relation to implant surgery, 8 cases of carcinomas of the tongue, 10 cases of cervical carcinomas and 30 cases of prostate carcinomas. HE stained sections were made from paraffin embedded tissue blocks fixed in 4% formaldehyde.

2.2. Data acquisition

Grey level images from 5–7 μm thick HE stained sections were digitised using a charged coupled device (CCD) camera (Philips® LDH 0670/00 equipped with a Hamamatsu® AC Adaptor, type A3472) mounted on a Zeiss Axioplan 2 microscope using a Plan-Neofluar 40×/0.75 lens in addition to a Prior HI52V2 microscope stage. The final magnification was 400× at a resolution of 876 nm (0.9 micrometers) per pixel.

2.3. Segmentation

Local segmentation was used, and developed from an algorithm based on the size of the elements to be detected and their contrast to the background. Thresholding was based on the pixel darkness measured as integrated optical density (IOD). Any pixel with an IOD within a given range is turned ON, otherwise OFF. The algorithms for construction of a continuous area of interest (e.g., cell nucleus) were further based on mathematical morphology [48]. From the nuclear profiles, the geometrical center of gravity was computed. The resulting data were stored as files of coordinates, where the coordinates represented a center of gravity. Further analysis of these raw data with computation of structural features was done on a Pentium® based PC running Windows 98®. Among the software facilities developed was the possibility to define digitally a subset of pointlike seeds in order to run the analyses in a limited window of analysis (Fig. 4).

2.4. Building a composite picture

A composite picture consisting of up to 50 fields of view was constructed by aligning each field of view according to a simple algorithm developed by the authors. For composite pictures generated by manual movement of the microscope stage, an algorithm using the binary mask of segmented nuclei was employed. One field of vision is composed of a matrix of 512×512 pixels, i.e., 512 separate columns and rows. When moving to the left in the visual field, the 128 left columns of the binary image are copied from the left to the right margin of the screen. The microscope stage is then moved until the binary mask is congruent to the grey level image. The offset of the movement thus is 384 columns (75% of the field of view), giving an overlap from one field to another of 25%. No further correction of the image alignment was performed. For automatic movement of the microscope stage a predetermined pattern of movement (spiralling) using the same offset as default was employed. Hereby, the stage was first moved one offset (384 columns) to the left, then one offset (384 rows) up, then two offsets (768 columns) to the right. Thereafter two offsets down, then three offsets to the left, then three offsets up, three followed by four offsets to the right and so on. The typical number of fields of view to be included was 25, 36 or 49, which makes up a square composite picture. For the automatic image acquisition, alignment of visual fields relied on the mechanical accuracy of the microscope stage.

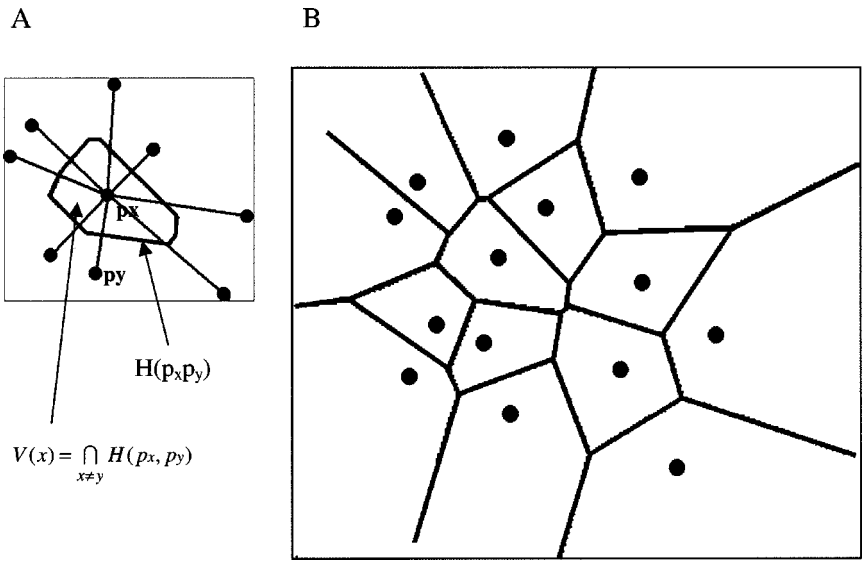

Fig. 1. A single Voronoi polygon, V_x (Panel A) is defined by the halfplanes $H(p_x p_y)$ that perpendicularly bisect the lines between a centre point and its neighbouring pointlike seeds. When we use this rule for every point considered, the area of interest is completely covered by adjacent polygons, constituting the Voronoi Diagram (VD). Panel B shows the VD for a set of randomly distributed pointlike seeds in a flat plane.

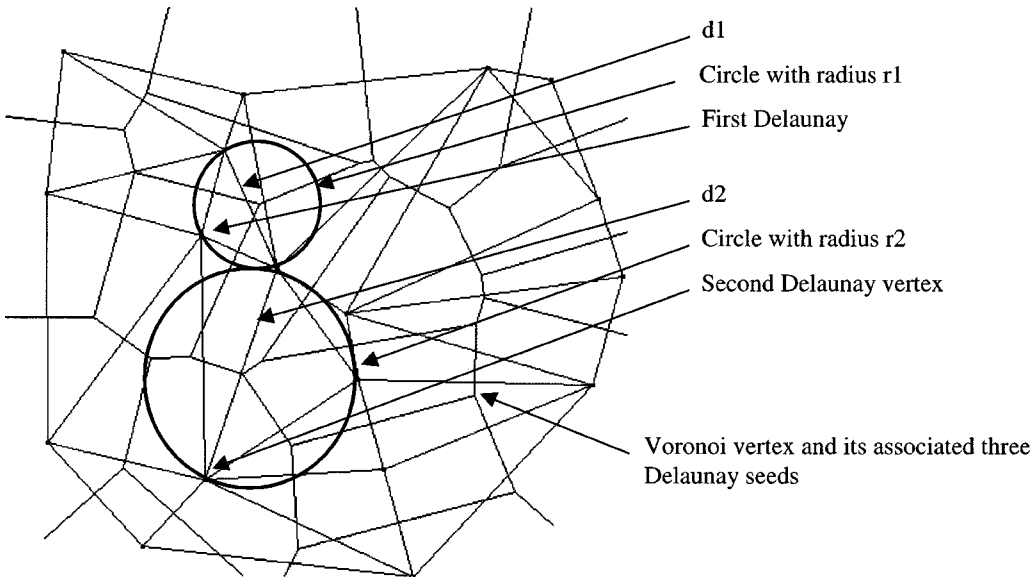

Fig. 2. The Delaunay Triangulation (DT) represents the dual of the VD and is constructed by drawing lines between pointlike seeds in adjacent Voronoi polygons. The completed construction is a triangular network that covers the whole area. A Delaunay network in two dimensions consists of non-overlapping triangles where no pointlike seeds in the network are enclosed by the circumscribing circles of any neighbouring triangle. The VD is shown in blue and the DT in red.

2.5. Space partitioning

Space partitioning in our context is based on the geometrical center of segmented cell nuclei. Computing the geometrical centers for each nucleus within a considered area creates a 2D swarm of pointlike seeds, from which the VD (Fig. 1) is constructed. All other graphs employed (DT, MST, GG and UT) are subgraphs of the main graph, the VD (Fig. 8). This graph was chosen as the principle tool for exploring the tissue structure, as it is considered to be the most informative [33]. The algorithms for generating these graphs

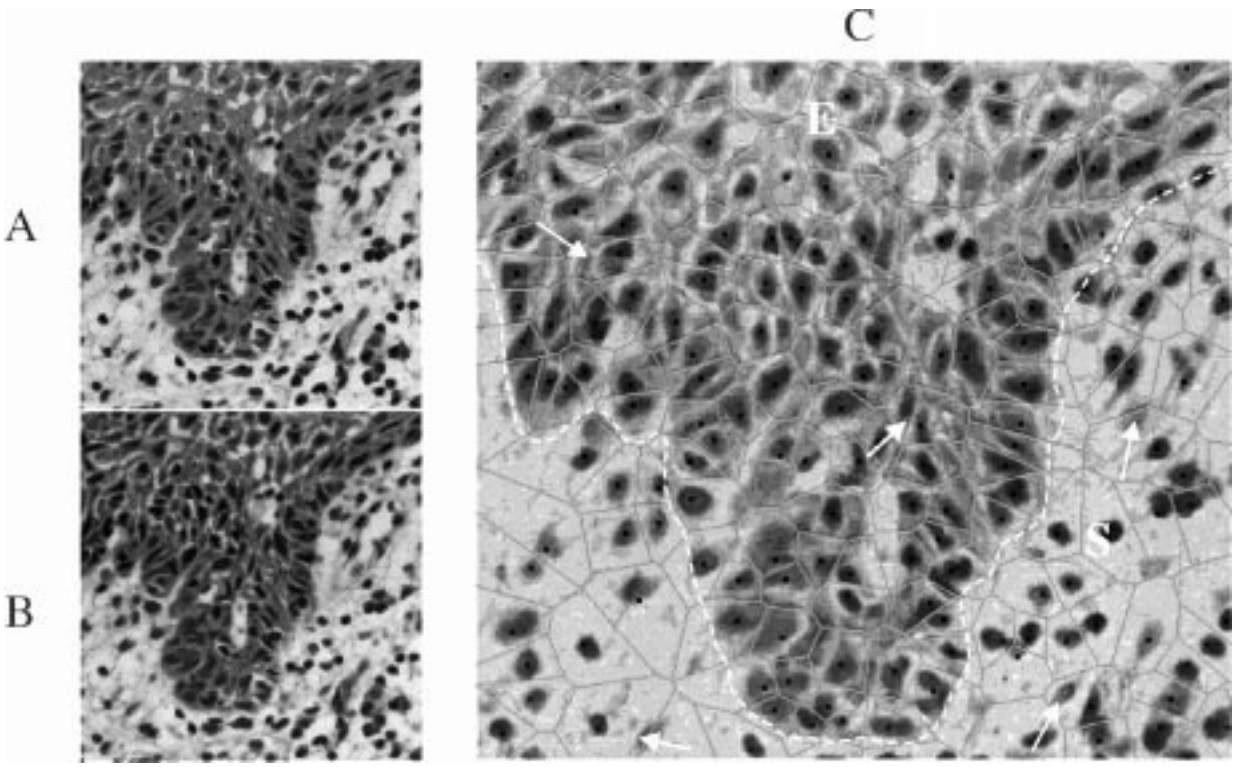

Fig. 3. An epithelial island from the part of an oral squamous cell carcinoma (E, Panel C) bordering onto the underlying connective tissue (S, Panel C). Panel A is a detail from a HE stained section from the invasive front. The corresponding grey level image is shown in Panel B. Panel C shows Voronoi polygons superimposed on the enlarged grey level image shown in Panel B. The pointlike seed are given as black dots superimposed on the cell nuclei, and represent the geometrical centre of gravity of each nucleus. Each Voronoi polygon represents the area of influence of one pointlike seed, which in epithelial tissues roughly correspond to the extension of a cell. Some cell nuclei are missed (white arrows), but the overall precision in the segmentation is acceptable. The dotted line denotes the basal membrane, which in this case also comprises a part of the border of the window of analysis. The scale bar is 0.10 mm (400× magnification).

have been presented elsewhere [9,33,39,49] and are only briefly commented here. The schematic relationships between the VD versus DT and VD versus GG are shown in Figs 2 and 5, respectively. Figure 3 shows relationship of the VD to tissue structures. In Panel A (HE stained section) an epithelial island (E) can be seen bordering onto the underlying stroma (S). Panel B is the corresponding grey level images. To the right the detail is shown in larger magnification, where the pointlike seeds are superimposed on the cell nuclei, and Voronoi polygons are constructed to make up the VD for the considered area. A program for eliminating border effects in marginal polygons (Fig. 8) was also developed. The window of analysis was defined digitally, by defining a closed contour with a digitizing pad and storing the coordinates of the contour. Only pointlike seeds within the contour were included in the analysis. The coordinates of the part of the contour crossing a marginal polygon were defined as the new edge in the polygon.

VD: The VD for a set of random distributed pointlike seeds is shown in Fig. 1. When two points, p_x and p_y are in a plane π, a half-plane, denoted $H(p_x p_y)$, is defined by the perpendicular bisector of $p_x p_y$. The locus of points closer to p_x than to any other point is the intersector of $N-1$ perpendicularly oriented halfplanes, where N is the number of points in the considered space. Hence,

$$V(x) = \bigcap_{x \neq y} H(px, py) \quad \text{(Panel A, Fig. 1).}$$

A single Voronoi polygon is defined by the intersection ($\cap$) of $N-1$ halfplanes in the considered space, i.e., the center point and its surrounding pointlike seeds (Panel A, Fig. 1). When applied to every point in the considered area, this rules gives the VD (Panel B, Fig. 1).

DT: The DT represents the dual of the VD and is constructed by drawing the lines between the pointlike

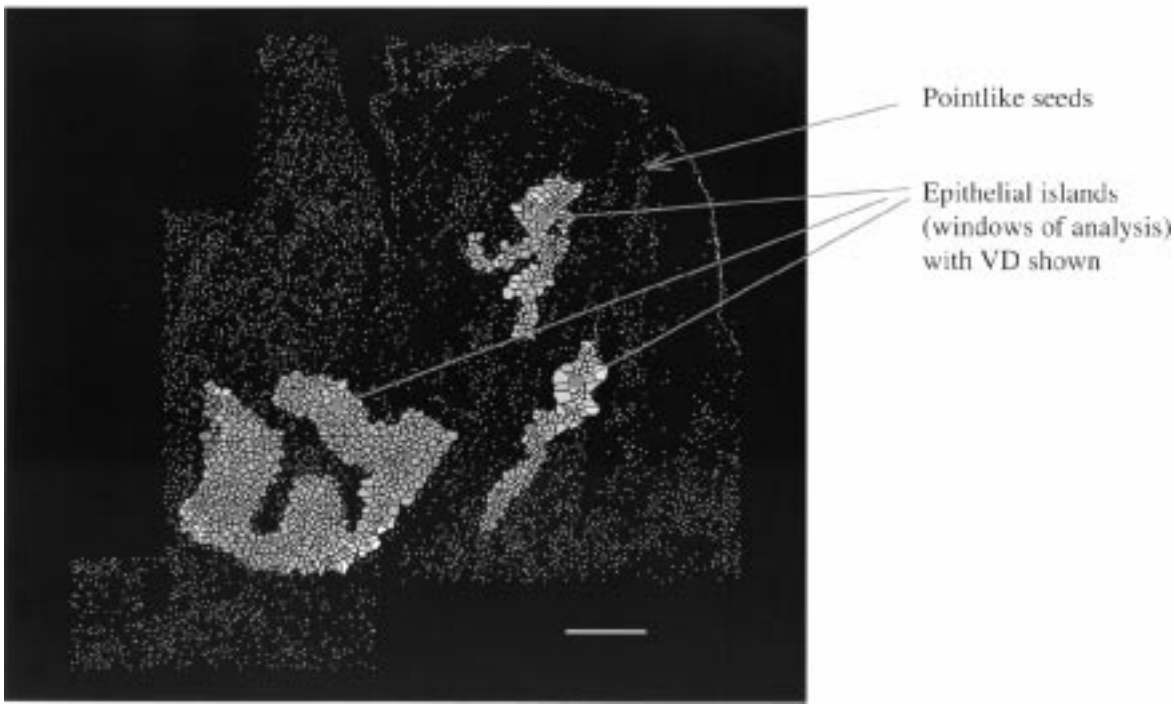

Fig. 4. VD with Voronoi polygons shown in pseudocolours, where the areas selected for analysis represent an epithelial island in the underlying stroma. Note in particular that border effects of marginal polygons (the polygons in the periphery of the analysis windows) have been eliminated. The sample is from the invasive front of an oral squamous cell carcinoma of the tongue. The scale bar is 0.5 mm.

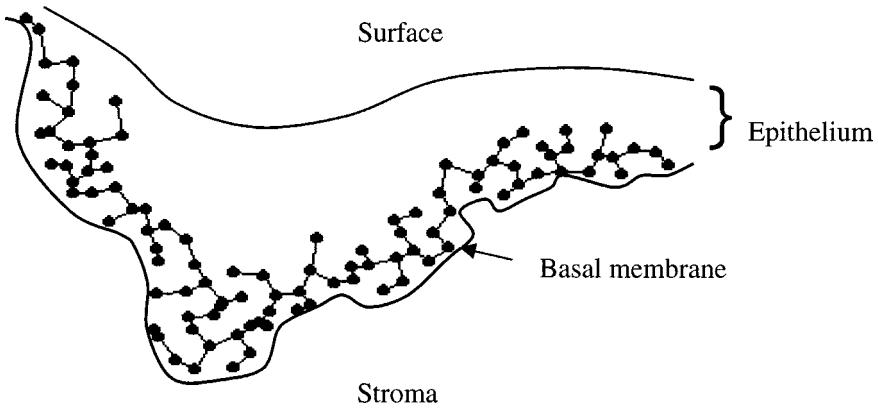

Fig. 5. The minimum spanning tree (MST) schematically drawn for the basal cell layers in normal oral mucosa. The MST represents the length of the shortest paths (or tours) through each point exactly once [37,38,50].

seeds in adjacent Voronoi polygons (Fig. 2). The completed construction is a triangular network that covers the considered area. A Delaunay network in two dimensions consists of non-overlapping triangles where no pointlike seeds in the network are enclosed by the circumscribing circles of any neighbouring triangle.

MST: Considering n distinct points in a d-dimensional space allows for $(n-1)!/2$ closed paths (or tours) through the space. Determining $L(n,d)$, the minimum tour length is possible by defining the smallest constant $\alpha(d)$ such that

$$\limsup_{n\to\infty} \frac{L(n,d)}{\frac{d-1}{n^d\cdot\sqrt{d}}} \leqslant \alpha(d).$$

Additionally, $\beta(d)$ given by

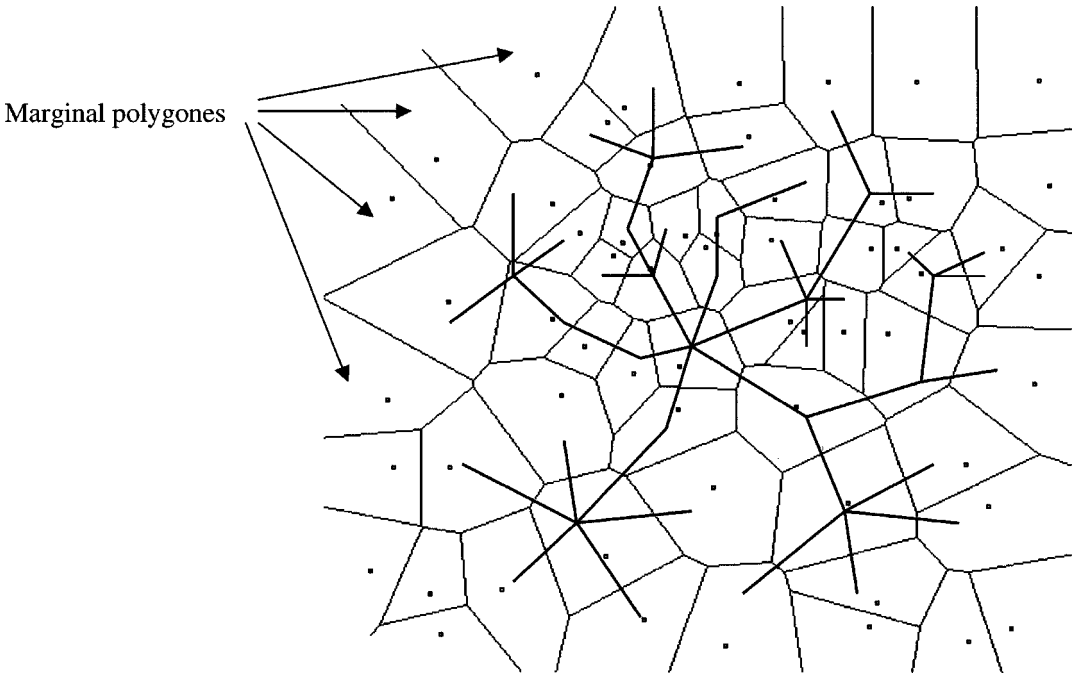

Fig. 6. The Ulam Tree represents a mathematical object growing in space and time according to specified rules [39]. The UT is generated from the VD, in such a manner that the "branches" of the tree only traverses polygons that are not traversed by any other branch of the tree. The structural feature ELH_av (average Edge Length Heterogeneity) is derived from the UT.

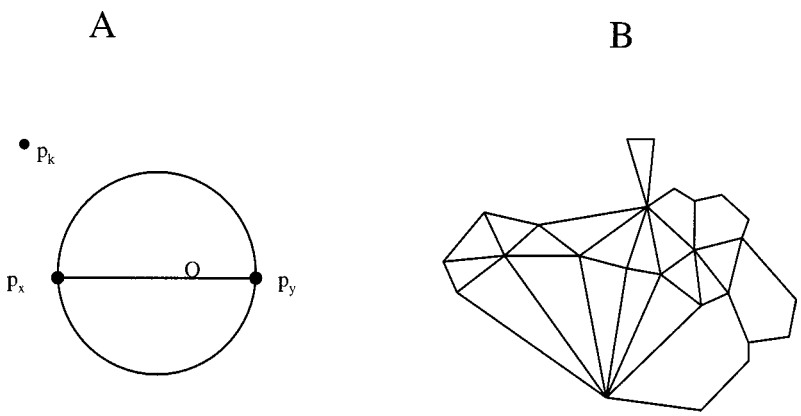

Fig. 7. The neighbourhood relation of two points (p_x and p_y) according to the Gabriel graph. Two points, p_x and p_y, are considered as neighbours if the circle on which they are placed, is empty. p_k on the other hand, is not a neighbour to either p_x or p_y (Panel A). The Gabriel graph (Panel B) is similar to the Delaunay Triangulation but contains polygons in addition to triangles.

$$\limsup_{n\to\infty} \frac{L(n,d)}{\frac{d-1}{n^d\cdot\sqrt{d}}} = \beta(d),$$

applies to almost all optimal tours in the considered d-dimensional space. The above limit fails only for a negligible subset of tours [37]. Furthermore, the above approximation applies to any dimension [50]. The solution to the problem can be reached by several different computations. For our purpose, the MST was derived by a decimation of the DT [51].

GG: In a *Gabriel graph* (Fig. 7, Panel B), pointlike seeds are connected by an edge only if the circle defined by the diameter connecting the nodes contains no other pointlike seeds (Panel A, Fig. 7). The two points (p_x and p_y) are neighbours if and only if the circle (with a center O) on which they are placed is empty.

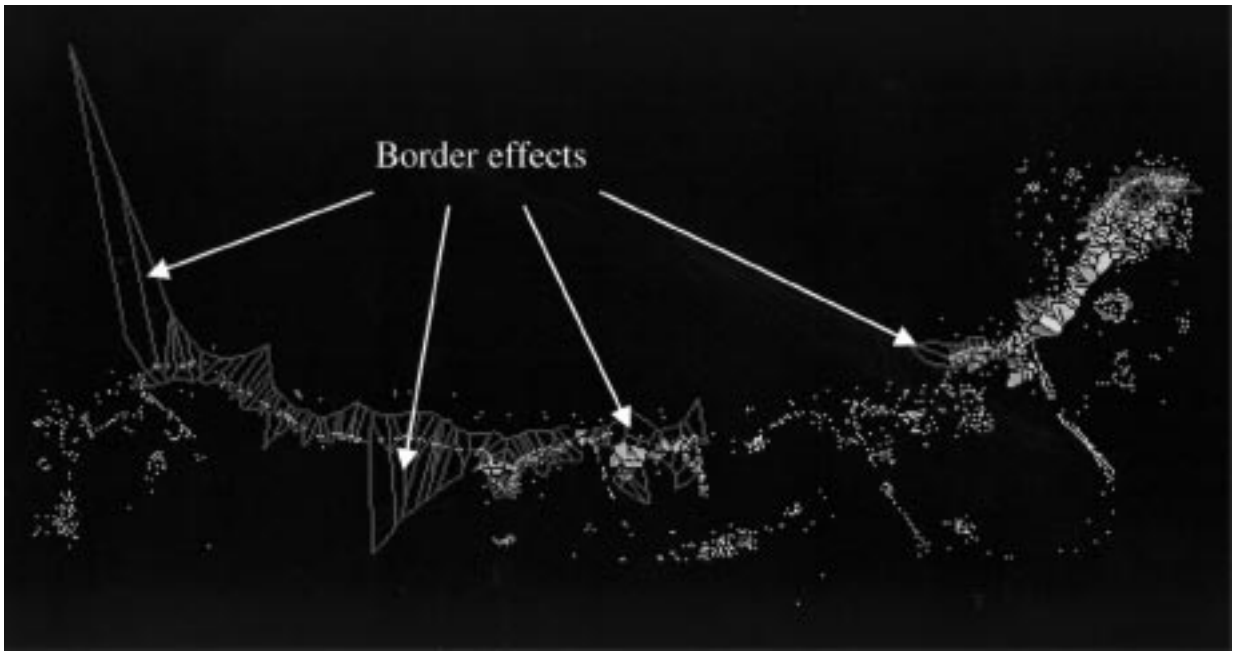

Fig. 8. A 2D swarm of pointlike seeds, generated from oral squamous cell epithelium. Several areas for analysis have been cut out digitally. Any polygon that could be modified by making the area of interest slightly larger (i.e., include additional number of seed peripheral to the considered swarm of points) are considered as border polygons. Border effects in the peripheral Voronoi polygons (arrows) are obvious. Scale bar 0.25 mm.

Accordingly, p_x and p_y (and not p_k) are neighbours (Panel A, Fig. 7). A graph similar to the DT can be constructed. However, it differs from the DT in that it contains polygons in addition to triangles (Panel B, Fig. 7). Cases are matched by comparing the lengths and orientations of the edges associated with each pointlike seed in one graph with those of every pointlike seed in a second graph. The Gabriel graph is particularly sensitive to subtle differences in the number or relative positions of pointlike seeds, making it a suitable tool for detecting changes in cellular organization within tissues.

UT: The *Ulam Tree* represents a mathematical object growing in space and time according to specified rules [39]. The UT is generated from the VD, in such a manner that the "branches" of the tree only traverses polygons that are not traversed by any other branch of the tree (Fig. 6).

2.6. Topographical analysis

Applying the 27 algorithms we have developed on the polygons obtained when constructing the VD and its subgraphs (space partitioning) we have performed topographical analysis on tissue specimens from normal and cancerous epithelium. The number of epithelial cells included in the analysis varied from 1500–5000. For a more detailed description of the VD-based algorithms, see Appendix.

2.7. Border effects

Marginal seeds represent a source of error as they yield polygons with a morphology that deviates from the population as a whole (Fig. 8). Structural features derived from marginal polygons are irrelevant, as they represent a non-representative population, and thus are a source of errors. Accordingly, we developed software to eliminate these aberrations (Fig. 4).

2.8. Temporal aspects

Scanning 30–50 fields of view is done in 8–12 minutes, depending on whether it is done manually or automatically. Segmentation requires another 5–10 minutes, depending on the number of cells in the specimen. For computation of structural features, 30–90 seconds are required, giving a total time expenditure of approximately 14–25 minutes.

2.9. Statistical analysis

Student's t-test was used for comparison of groups. All P-values were two-tailed, and values less than 0.05 were considered to indicate statistical significance.

Table 1
Diagnostic value of 10 structural features in *normal and malignantly changed oral mucosa*

Structural feature¶	Mean values with ranges given in parentheses		
	Normal oral mucosa ($n = 8$)	Carcinoma of the tongue ($n = 8$)	P-value* Student's t-test
RF_av‡	0.81 (0.73–0.87)	0.67 (0.57–0.76)	0.01
RF_dis‡	0.74 (0.66–0.78)	0.67 (0.56–0.79)	0.05
A_dis	0.45 (0.37–0.54)	0.39 (0.33–0.45)	0.09
MSTEL_av	5.9† (4.3–6.8)	4.2 (3.3–4.8)	0.05
PTS	0.74 (0.66–0.87)	0.69 (0.54–0.74)	0.10
DEL_av‡	26.8† (25.9–27.3)	25.8 (4.9–26.7)	0.001
ELH_av‡	0.41 (0.33–0.58)	0.32 (0.23–0.48)	0.02
DKNN_av‡	728† (703–738)	693 (641–723)	0.03
NNRR	32.8 (29.1–33.6)	31.9 (28.4–33.2)	0.16
RMPB‡	15.7† (13.6–17.1)	12.9 (11.6–14.6)	0.03

* Two-tailed.

† Measured in pixels.

‡ Structural features for which the differences between groups reaches statistical significance.

¶ See Appendix for an explanation of the separate structural features.

Results from running 10 different form parameters on altogether 16 samples of oral mucosa, 8 cases from normal mucosa and 8 cases with carcinomas of the tongue. Numbers in parentheses denote the range. 5000 cells were included in the analysis. For a further description of the form features used, see Appendix. P-value for the best structural feature was 0.001 (DEL_av).

3. Results

A total of 10 of the 27 structural features we investigated were able to distinguish between normal and malignantly altered tissue and/or were shown to have a possible prognostic value (Tables 1–3).

3.1. Normal oral mucosa versus carcinoma of the tongue (Table 1)

Eight biopsies from assumptively normal oral mucosa (acquired during gingivoplastic procedures in relation to serial extraction of teeth) and 8 cases of oral carcinomas of the tongue were compared with regard to values of 10 structural features. Of these, 6 features made it possible to distinguish between normal oral mucosa and carcinoma of the tongue, usually situated at the lateral border of tongue, bordering onto the floor of the mouth. This pertains to the features RF (**r**oundness **f**actor, from the VD, $P = 0.01$), RF_dis (disorder of the roundness factor disorder derived from the VD, $P = 0.05$), DKNN_av (average **d**istance to the **K**-**n**earest **n**eighbours, $P = 0.03$), DEL_av (average Delaunay Edge Length, $P = 0.001$), ELH_av (average **e**dge **l**ength **h**eterogeneity of the Ulam Tree, $P = 0.02$) and RMPB (**r**adius of the **m**aximum **p**ercolating **b**all, percolating the Delaunay network, $P = 0.03$). The ability of these methods to discern normal and cancerous oral epithelium points to a diagnostic potential as they obviously detect structural differences between normal and malignantly changed mucosa.

Table 2
Prognostic values of 10 structural features in *carcinomas of the cervix*

Structural feature¶	Mean values with ranges given in parentheses		
	Cervical carcinomas with good prognosis ($n = 4$)	Cervical carcinomas with poor prognosis ($n = 6$)	P-value* Student's t-test
RF_av	0.74 (0.55–0.86)	0.69 (0.54–0.81)	0.31
RF_dis	0.69 (0.56–0.78)	0.69 (0.56–0.78)	0.24
A_dis‡	0.45 (0.37–0.54)	0.35 (0.27–0.41)	0.02
MSTEL_av	4.9† (4.3–5.8)	5.2 (3.3–6.8)	0.33
PTS	0.78 (0.66–0.87)	0.65 (0.54–0.74)	0.14
DEL_av	26.8† (25.9–27.3)	26.3 (24.9–27.7)	0.31
ELH_av‡	0.57 (0.39–0.68)	0.38 (0.29–0.43)	0.02
DKNN_av	725† (650–795)	695 (630–740)	0.09
NNRR‡	36.0 (30.6–41.2)	30.1 (28.4–31.9)	0.03
RMPB‡	16.1† (14.6–17.1)	13.9 (11.6–15.1)	0.04

* Two-tailed.

† Measured in pixels.

‡ Structural features for which the differences between groups reaches statistical significance.

¶ See Appendix for an explanation of the separate structural features.

Results from running 10 different form parameters on altogether 10 cases of carcinomas of the cervix, 4 with good (relapse-free survival more than 12 years) and 6 with poor (relapse-free survival less than 5 years) prognosis when 5000 cells are include in the analysis. The featured DEL_av and ELH_av display significant differences in the two prognosis groups in this test set. P-values for the best structural feature are 0.001 (DEL_av and ELH_av).

3.2. Carcinomas of the cervix (Table 2)

Altogether 10 biopsies; 4 with a good (relapse-free survival more than 12 years) and 6 cases with a poor (relapse-free survival less than 5 years) prognosis were examined. A total of 4 structural features made it possible to distinguish between the two outcome groups. These were A_dis (**a**rea **dis**order, from the VD, $P = 0.02$), ELH_av ($P = 0.02$), NNRR (**n**umber of **n**earest neighbours within a **r**estricted **r**adius of 75 pixels, $P = 0.03$) and RMPB ($P = 0.04$).

3.3. Carcinomas of the prostate (Table 3)

The values of the same 10 form features as above when applied to 30 cases of carcinomas of the prostate (15 cases with good and 15 cases with poor prognosis) are shown. Altogether 5 structural features made it possible to distinguish between carcinomas of the prostate with a good and poor prognosis. These structural features were RF_dis ($P = 0.01$), A_dis ($P = 0.02$), DEL_av ($P = 0.001$), ELH_av ($P = 0.02$), DKNN (average distance to the K nearest neighbour, from the DT, $P = 0.01$) and RMPB ($P = 0.03$).

Note in particular that DEL_av and/or ELH_av as significant descriptors are common to all three sets of tissues.

4. Discussion

We present data from several tissues that demonstrate the possible diagnostic and prognostic value of

Table 3
Prognostic value of 10 structural features in *carcinomas of the prostate*

Structural feature¶	Mean values with ranges given in parentheses		
	Prostate carcinomas good prognosis ($n = 15$)	Prostate carcinomas poor prognosis ($n = 15$)	P-value* Student's t-test
RF_av	0.65 (0.52–0.78)	0.71 (0.54–0.81)	0.31
RF_dis‡	0.61 (0.53–0.68)	0.74 (0.63–0.81)	0.01
A_dis‡	0.55 (0.37–0.64)	0.35 (0.27–0.41)	0.02
MSTEL_av	5.5† (4.3–6.8)	5.4 (4.1–6.2)	0.33
PTS	0.73 (0.63–0.87)	0.65 (0.54–0.74)	0.24
DEL_av‡	27.1† (25.9–28.3)	25.8 (24.9–26.7)	0.001
ELH_av‡	0.41 (0.33–0.58)	0.32 (0.23–0.48)	0.001
DKNN_av	705† (650–740)	680 (640–715)	0.11
NNRR‡	29.8 (29.1–30.6)	33.8 (29.4–35.2)	0.01
RMPB‡	13.4† (11.7–14.4)	16.2 (13.4–17.4)	0.03

* Two-tailed.

† Measured in pixels.

‡ Structural features for which the differences between groups reaches statistical significance.

¶ See Appendix for an explanation of the separate structural features.

Results from running 10 different form parameters on altogether 30 cases of prostate carcinomas, 15 with good (relapse-free survival more than 10 years) and 15 with poor (relapse-free survival less than 3 years) prognosis.5000 cells were included in the analysis. P-value for the best structural feature is 0.02 (A_dis and ELH_av).

tissue architecture analysis by VD-based algorithms. However, it should be kept in mind that we have tested out a large number of putatively informative structural features (27) on a limited number of objects (56). Testing out such a large number of structural features on a limited number of objects statistically is not without pitfalls [52]. The interpretations from this preliminary study must therefore be cautious, and a limited number (2–4) of structural features (e.g., DEL_av and ELH_av) should be applied to an independent test set with a number of objects considerably larger than the number of features explored [52]. Furthermore, these methods have to be tested out in a series of different tissues, to avoid the serious problem of over-fitting, where selected features characterise well the samples in specific training data, but not the general classes (e.g., prostatic carcinomas, but not all carcinomas). Over-fitting typically occurs when the number of features analyzed is high in relation to the number of samples considered [52].

The main graph in our context (VD) encompasses a number of subgraphs, such as the DT, MST, UT and GG, and is generally regarded as the most informative graph [33]. The VD is generated from the point-like seeds representing the centres of gravity within cell nuclei of the considered tissues. The Voronoi polygon represents the region of influence of each seed. A priori, this does not have any direct biological correlate. However, in epithelial tissues, with only sparse intercellular substance, it roughly corresponds to the somata of the epithelial cells. The shape and size of cells is a structural feature of considerable interest in tradi-

tional histological assessment of pathologically altered tissue, e.g., carcinomas. This relationship breaks down when we consider the stromal tissue, with abundance of intercellular substance (Panel C, Fig. 3). Thus, the algorithms we have developed directly derived from the VD we believe are best suited for tissues with a minimum of intercellular substance, as in epithelial tissues, although graph theory based methods could be applied to any tissue. Also, there is no reason that other features based on, e.g., the MST or UT should have such limitations.

The segmentation algorithms we developed performed with an acceptable level of precision (Panel C, Fig. 3). The algorithms represent a compromise of speed and precision. For the calculations, a minimum number of 1000–1500 cells were included. We have chosen such a fairly large number of objects to be included in the analysis, as preliminary runnings of computations indicated that the values of the structural features did not stabilise until at least 1000–1500 objects were included, the exact number depending on which tissue was analysed. The imaginary flat plane we consider in fact represents a 3-dimensional entity. Depending on the thickness of the sections considered some cell nuclei might be below or above the focal plane and therefore missed in the segmentation. However, for the sections thickness we have employed (5–7 μm), this has not been a major problem (Fig. 3). However, several observer note that in the most aggressive carcinoma of the oral cavity, the epithelial cells bordering onto the underlying stroma show a distinct blurring of their structural features, with a typically glossy appearance of the somata and nuclei (M. Bryne, personal communication). It is conceivable that because of this, a considerable number of nuclei in the area of interest could escape segmentation, with a resultant error in the estimation of the structural features. If this is abundant, a distinct prognostic group of lesions might be missed, at least with the HE staining procedures. In our study, however, this was not a prominent feature, and the precision in the segmentation is acceptable (Fig. 3).

A particular point of interest when investigating the invasive front of carcinomas, are the border effects, that tend to have a dominant effect when the epithelial islands become small and numerous (Fig. 8). In particularly aggressive lesions, the gross structures of the invasive front of carcinomas tend to disintegrate, with multiple small fingerlike projections into the underlying stroma. In sections, these projections will present as small epithelial islands, consisting of a very limited number of cells. This poses a possible problem with regard to border effects, as these will become dominant in small cluster of cells, perhaps eliminating them entirely. Again, these cases may be of particular prognostic interest.

Only structural features of the epithelial tissue have been investigated in this study. Most likely, the tumor-host response will result in structural alterations of diagnostic and prognostic value also in the underlying stroma. Such features could be the amount of inflammatory response, which can easily be assessed by our methods. However, current algorithms do not detect the specific nature of subepithelial lymphocyte infiltration. New methods for segmentation of immunohistochemically stained cell nuclei [53,54] might contribute to shedding more light on biological information contained in the pattern of inflammatory response.

Twenty-nine structural features based on algorithms derived from the Voronoi Diagram or its subgraphs on different sets of epithelial tissues, 10 of which were demonstrated to have a diagnostic or prognsotic potential. The ultimate test for these methods will be to employ a limited number of structural features (e.g., DEL_av and ELH_av) on an independent test set [52].

Appendix

1) A_dis Area disorder. The Area denotes the area of a singleVoronoi polygon, measured in pixels. Area disorder reflects the variation in the polygons associated with considered population of pointlike seeds:

$$\text{A\_dis} = 1 - \frac{1}{1 + \frac{\text{Area\_sd}}{\text{Area\_av}}}.$$

The feature acquires the value of 0 if all polygons have the same area and tends towards 1 otherwise. The entire population except the marginal polygons is considered.

2) DEL_av Average Delaunay Edge Length. This feature sums up the edge length of the edges of the DT's and divides it by the number of non-marginal seeds.

3) DEL_dis Delaunay Edge Length disorder. This feature considers the standard deviation (sd) in the lengths of the edges of the Delaunay triangles linking non-marginal seeds:

$$\text{DEL\_dis} = 1 - \frac{1}{1 + \frac{\text{DEL\_sd}}{\text{DEL\_av}}}.$$

4) DENS This feature represents the density of the entire population, except for the marginal polygons, which are eliminated:

$$\text{DENS} = \frac{\text{Polygon\_nb}}{\text{Area\_sum}}.$$

Polygon_nb denotes the number of polygons.

5) DEP_av Average Delaunay Edge Probability. One edge of the DT belongs to two triangles, each being associated with an overlapping circle.

$$\text{DEP\_av} = 1 - \frac{d1 + d2}{r1 + \frac{r2}{\sqrt{3}}}.$$

Here, $d1$ denotes the distance of the first vertex to its nearest neighbour. $d2$ denotes the distance of the second vertex to its nearest neighbour. $r1$ is the radius of the first circle associated with the two triangles sharing the considered edge $d1$. $r2$ is the radius of the first circle associated with the two triangles sharing the considered edge $d2$ (Panel B, Fig. 2). $\sqrt{3}$ is a normalization factor with respect to the triangular lattice.

6) DEP_dis Delaunay Edge Probability disorder. This feature denotes the disorder of the abovementioned Delaunay edge probability, and is given by the following equation:

$$\text{DEP\_dis} = 1 - \frac{1}{1 + \frac{\text{DEP\_sd}}{\text{DEP\_av}}}.$$

7) DFRAC_av This feature denotes theaverage fractal dimension of the Ulam Trees, more precisely an application of the Hausdorff fractal dimension which express the properties of topological defects in the tree structure, e.g., as related to a highly regular tree. Consider a closed contour (e.g., Ulam Tree) within a 3-dimensional space, on which two points, x and y, are placed. The Ulam Tree may be viewed as projected onto a 2-dimensional flat plane with a unit of length correponding to the size of a pixel. The mean value of the area A covered by traversing in N steps from x to y along branches of the tree is given by $A \sim r^2$ where r is given by the equation

$$r = \left] \sqrt{|x - y|^2} \right[.$$

The quantity r effectively expresses the radius of the circle or square that has the same area as the projection of the traversed part of the tree projected onto the 2D plane where][denotes the mean values for all the possible contours within the considered space. x and y are expressed as vectors.

The Hausdorff dimension of the Ulam Tree is given by the general equation

$$D_{\text{H}} = \frac{1}{\Delta} \quad [56].$$

A value of $\Delta \neq 1$ indicates that the contour is fractal, i.e., how fuzzy the contour is.

The algorithm for computing Δ is fairly simple. From the projection of the tree on a 2D plane (with pixels as the unit of length), one defines the upper, lower, left and right margins of the projected tree, which yields a rectangle with sides lx, ly. The r accordingly is defined as

$$r = \sqrt{lx \cdot ly},$$
$$\log N = \log k + \frac{1}{\Delta \cdot \log r}.$$

N can be derived from the number of nearest neighbouring distances required to build the Ulam Tree within a square where p_x and p_y lies on two opposing margins of the square.

8) DKNN_av Average distance to the K nearest neighbours. K was set to 16 after preliminary simulations showed that normal oral mucosa was statistically different from carcinomas of the tongue with regard to this feature (Table 1). Concidering the K nearest neighbours of a seed (according to the Delaunay neighbourhood), if all of them are non-marginal seeds, DKNN represents the sum of the distances to those neighbours. DKNN_av represents the average of DKNN over all the population where K neighbourhood does not contain any marginal seeds.

9) DKNN_dis Disorder of the DKNN.

$$\text{DKNN\_dis} = 1 - \frac{1}{1 + \frac{\text{DKNN\_sd}}{\text{DKNN\_av}}}.$$

10) DRT Divergence from the Regular Tree. Transposing the tree into a matrix, M, quantitates the neighbourhood from the Ulam tree, composed of two orthogonal properties which are the integration and the topological properties of the tree. The integration property is the number of points

added to the tree at a given level of expansion, while the topological property is the number of simple, double, triple etc junctions in the tree.

$$\mathrm{DRT} = \sum_x \frac{\sum_y |M_p xy - M_{\mathrm{ref}} xy|}{Wx}.$$

Hence, the DRT is the sum of the weighted absolute differences between all the elements of the actual tree matrix (M_p) and a theoretical tree reference (M_{ref}) [51].

11) ELH_av This is the average edge length heterogeneity of the UT's. ELH is representative of the intrinsic node to node distance variations in the current tree [39].

12) HA_av Average area of holes within the considered area. Concidering the empty circles associated with each DT, a triangle is set to be a hole if a ball of the current radius is able to go through at least one edge of the triangle (i.e., the radius of the percolating ball is equal to or less than at least one of the edges of the triangle). The number of DTs that can be percolated by the ball is the average number of DT's belonging to the defined holes within the architecture. Considering a Voronoi Edge (VE, denoted ls) and the Delaunay edge that it bisects (lg), let $r1$ and $r2$ be the radius of the circles associated with the vertices of VE. If the length of ls is smaller than or equal to both $r1$ and $r2$, then the two DT's associated with the vertices of VE and accordingly belong to the same hole.

13) HA_dis Disorder of the whole area of the holes:

$$\mathrm{HA\_dis} = 1 - \frac{1}{1 + \frac{\mathrm{HA\_sd}}{\mathrm{HA\_av}}}.$$

14) MSPDG Minimal step percolating the Delaunay network (PPDG). MSPDG is the length of the minimum step required to be able to move without interruption from neighbour to neighbour in a Delaunay network that comprises at least 50% of the entire network.

15) MSTEL_av Average Minimum Spanning Tree Edge Length. The MST is a tree that spans the entire population in such a way that the sum of the Euclidian edge length is minimal [49,55].

16) MSTEL_dis Edge length disorder of the minimum spanning tree:

$$\mathrm{MSTEL\_dis} = 1 - \frac{1}{1 + \frac{\mathrm{MSTEL\_sd}}{\mathrm{MSTEL\_av}}}.$$

17) NNRR_av Average number of neighbours within a restricted radius. Considering a circle of which radius is set to 75 pixels around one seed, NNRR is the number of other seeds (neighbours) lying within the circle. NNRR_av is the average number of seeds over all the population located at least at 75 pixels from the border of the analysis window.

18) NNRR_dis Disorder in the Number of Neighbours within a Restricted Radius.

$$\mathrm{NNRR\_dis} = 1 - \frac{1}{1 + \frac{\mathrm{NNRR\_sd}}{\mathrm{NNRR\_av}}}.$$

Considering a circle of which radius is set to 75 pixels around one seed, NNRR is the number of other seeds (neighbours) lying within the circle. NNRR_dis is the disorder in the number of seeds over all the population located at least at 75 pixels from the border of the analysis window.

19) PTS probability of topological stability. The center of gravity of the nuclei are stored as coordinates. For a small error tolerance (one pixel in diameter), we move at random all the pointlike seeds and rebuild the graph. As long as one seed keeps the same neighbourhood as without any disorder, the local PTS is computed to be 1, otherwise it is computed to be zero. At each step in the simulation (with one pointlike seed moved one pixel in any direction) the PTS represents the number of seeds that kept (after disorder is introduced) the same neighbourhood as initially. The more the PTS is close to 1, the more stable the graph is to a random alteration. The closer the PTS is to zero the less valid local topology is, only statistics can then be computed about local topology. The marginal polygons are not considered.

20) PTS_av Average of 10 runs of the PST computation.

21) PTS_dis is the PTS disorder, given by the equation

$$\mathrm{PTS\_dis} = 1 - \frac{1}{1 + \frac{\mathrm{PTS\_sd}}{\mathrm{PTS\_av}}}.$$

22) RMPB Radius of the Maximum Percolating Ball. Considering the empty circles associated with each DT, a triangle is set to be a hole if the radius of its associated empty circle is below the current threshold. If a ball of the current radius is

able to go through one edge of the DT, then the neighbouring triangle can be associated to the same hole as the considered triangle. The RMBP is the value of the radius of the biggest ball that is able to percolate the Delaunay network. In other terms, it is the maximum allowed radius to obtain a hole that is larger in area than 50% of the area of all the holes.

23) RF_av Average Roundness Factor, given as

$$\text{RF\_av} = \frac{4\pi}{\text{perimeter}^2}.$$

This equals the roundness factor of one polygon average RF is the average over the entire population except for the marginal polygons.

24) RF_dis Roundness Factor disorder.

$$\text{RF\_dis} = 1 - \frac{1}{1 + \frac{\text{RF\_sd}}{\text{RF\_av}}}.$$

This features acquires the value of 1 if all the RF's are the same and tends towards zero otherwise. The entire population except the marginal polygons were considered.

25) WGC Weighted global compacity. Consider one Voronoi vertex and its associated three Delaunay seeds (Fig. 2). To each of these seeds corresponds a set of nearest neighbours located at distance $d1, d2, d3$ etc. If r is defined as the radius of the considered Delaunay circle, the compacity is defined as follows:

$$\text{WGC} = \frac{d1 + d2 + d3 \cdots dn}{9r^2}.$$

The weighted compacity (WC) is equal to C multiplied by the area of the DT. By doing this, one takes into account the possibility that two different DT's can have exactly the same proportions, but different size and accordingly represent two differents structural entities. Only considering the unweighted compacity does not take this possibility into consideration. Only considering the WGC is the average WC over all non-marginal polygons.

26) WGC_av Average Compacity.

$$\text{WGC\_av} = \sum_{i=1}^{n} \frac{d1 + d2 + d3}{9r^2 \cdot n}.$$

27) WGC_dis This represents the disorder of the Compacity, and is given by the equation

$$\text{WGC\_dis} = 1 - \frac{1}{1 + \frac{\text{C\_sd}}{\text{C\_av}}}.$$

Acknowledgements

This work was supported by a grant from the Norwegian Cancer Society. We thank Signe Eastgate and Ruth Punthervold for preparation of HE slides for the analysis. Discussions with professor A. Sudbø on the chapters of fractal geometry are acknowledged.

References

[1] G. Anneroth, L.S. Hansen and S.J. Silverman, Malignancy grading in oral squamous cell carcinoma. I. Squamous cell carcinoma of the tongue and floor of mouth: histologic grading in the clinical evaluation, *J. Oral Pathol.* **15** (1986), 162–168.

[2] M. Bryne, H.S. Koppang, R. Lilleng, T. Stene, G. Bang and E. Dabelsteen, New malignancy grading is a better prognostic indicator than Broders' grading in oral squamous cell carcinomas, *J. Oral Pathol. Med.* **18** (1989), 432–437.

[3] T. Bundgaard, S.M. Bentzen, J. Wildt, F.B. Sorensen, H. Sogaard and J.E. Nielsen, Histopathologic, stereologic, epidemiologic, and clinical parameters in the prognostic evaluation of squamous cell carcinoma of the oral cavity, *Head Neck* **18** (1996), 142–152.

[4] A. Karabulut, J. Reibel, M.H. Therkildsen, F. Praetorius, H.W. Nielsen and E. Dabelsteen, Observer variability in the histologic assessment of oral premalignant lesions, *J. Oral Pathol. Med.* **24** (1995), 198–200.

[5] T. Jorgensen, K. Yogesan, F. Skjorten, A. Berner, K.J. Tveter and H.E. Danielsen, Histopathological grading and DNA ploidy as prognostic markers in metastatic prostatic cancer, *Br. J. Cancer* **71** (1995), 1055–1060.

[6] C. Busch, A. Engberg, B.J. Norlen and B. Stenkvist, Malignancy grading of epithelial bladder tumors, *Scand. J. Urol. Nephr.* **11** (1977), 143–148.

[7] G. Anneroth, J. Batsakis and M. Luna, Review of the literature and a recommended system of malignancy grading in oral squamous cell carcinomas, *Scand. J. Dent. Res.* **95** (1987), 229–249.

[8] H.J. Gundersen, T.F. Bendtsen, L. Korbo, N. Marcussen, A. Moller, K. Nielsen, J.R. Nyengaard, B. Pakkenberg, F.B. Sorensen and A. Vesterby, Some new, simple and efficient stereological methods and their use in pathological research and diagnosis, *APMIS* **96** (1988), 379–394.

[9] E. Rivedal and T. Sanner, Promotional effect of different phorbol esters on morphological transformation of hamster embryo cells, *Cancer Lett.* **17** (1982), 1–8.

[10] R.K. Hunt, L. Bodenstein, J.S. Cohen and R.L. Sidman, Positional variations in germinal cell growth in pigment-chimeric eyes of Xenopus: posterior half of the developing eye studied in genetic chimerae and in computer simulations, *Proc. Natl. Acad. Sci. USA* **85** (1988), 3459–3463.

[11] L. Bodenstein and R.L. Sidman, Cell patterning in vertebrate development: models and model systems, *Curr. Top. Dev. Biol.* **21** (1987), 1–29.

[12] L. Bodenstein, A dynamic simulation model of tissue growth and cell patterning, *Cell Differ.* **19** (1986), 19–33.

[13] P.J. van Diest, K. Kayser, G.A. Meijer and J.P. Baak, Syntactic structure analysis, *Pathologica* **87** (1995), 255–262.

[14] K. Kayser and W. Schlegel, Pattern recognition in histopathology: basic considerations, *Methods Inf. Med.* **21** (1982), 15–22.

[15] J.P. Rigaut, D. Schoevaert-Brossault, A.M. Downs and G. Landini, Asymptotic fractals in the context of grey-scale images, *J. Microsc.* **189**(Pt. 1) (1998), 57–63.

[16] S.O. Mikalsen, E. Rivedal and T. Sanner, Morphological transformation of Syrian hamster embryo cells induced by mineral fibres and the alleged enhancement of benzo[a]pyrene, *Carcinogenesis* **9** (1988), 891–899.

[17] N. Aarsaether, J.R. Lillehaug, E. Rivedal and T. Sanner, Cell transformation and promoter activity of insulation oils in the Syrian hamster embryo cell and in the C3H/10T1/2 mouse embryo fibroblast test systems, *J. Toxicol. Environ. Health* **20** (1987), 173–185.

[18] E. Rivedal and T. Sanner, Retinoids have different effects on morphological transformation and anchorage independent growth of Syrian hamster embryo cells, *Carcinogenesis* **6** (1985), 955–958.

[19] C. Davies, E. Rivedal and T. Sanner, Fibrinolytic activity and morphological transformation of hamster embryo cells, *Carcinogenesis* **3** (1982), 621–624.

[20] L.A. Dethlefsen, J.M. Prewitt and M.L. Mendelsohn, Analysis of tumor growth curves, *J. Natl. Cancer Inst.* **40** (1968), 389–405.

[21] J.M.S. Prewitt, On some applications of pattern recognition and image processing to cytology, cytogenetics and histology, Thesis, University of Uppsala, Sweden, 1978.

[22] J.M. Prewitt and M.L. Mendelsohn, The analysis of cell images, *Ann. N.Y. Acad. Sci.* **128** (1966), 1035–1053.

[23] M.L. Mendelsohn, W.A. Kolman, B. Perry and J.M. Prewitt, Computer analysis of cell images, *Postgrad. Med.* **38** (1965), 567–573.

[24] B. Stenkvist, S. Westman-Naeser, J. Holmquist, B. Nordin, E. Bengtsson, J. Vegelius, O. Eriksson and C.H. Fox, Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations, *Cancer Res.* **38** (1978), 4688–4697.

[25] E. Bengtsson, O. Eriksson, J. Holmquist and B. Stenkvist, A software system to record and analyze digitized cell images, *Comput. Programs Biomed.* **7** (1977), 233–246.

[26] O. Eriksson, S. Westman-Naeser, B. Stenkvist, E. Bengtsson and J. Holmquist, A computerized recording system for cytopathologic diagnoses, *Acta Cytol.* **21** (1977), 266–270.

[27] P.E. Levi, E.H. Cooper, C.K. Anderson and R.E. Williams, Analyses of DNA content, nuclear size and cell proliferation of transitional cell carcinoma in man, *Cancer* **23** (1969), 1074–1085.

[28] H.K. Choi, T. Jarkrans, E. Bengtsson, J. Vasko, K. Wester, P.U. Malmstrom and C. Busch, Image analysis based grading of bladder carcinoma. Comparison of object, texture and graph based methods and their reproducibility, *Anal. Cell. Pathol.* **15** (1997), 1–18.

[29] T. Jarkrans, J. Vasko, E. Bengtsson, H.K. Choi, P.U. Malmstrom, K. Wester and C. Busch, Grading of transitional cell bladder carcinoma by image analysis of histological sections, *Anal. Cell. Pathol.* **8** (1995), 135–158.

[30] H.K. Choi, J. Vasko, E. Bengtsson, T. Jarkrans, P.U. Malmstrom, K. Wester and C. Busch, Grading of transitional cell bladder carcinoma by texture analysis of histological sections, *Anal. Cell. Pathol.* **6** (1994), 327–343.

[31] B. Stenkvist, E. Bengtsson, O. Eriksson, T. Jarkrans and B. Nordin, Image cytometry in malignancy grading of breast cancer. Results in a prospective study with seven years of follow-up, *Anal. Quant. Cytol. Histol.* **8** (1986), 293–300.

[32] B. Stenkvist, E. Bengtsson, B. Dahlqvist, G. Eklund, O. Eriksson, T. Jarkrans and B. Nordin, Predicting breast cancer recurrence, *Cancer* **50** (1982), 2884–2893.

[33] G.T. Toussaint, Pattern recognition and geometrical complexity, in: *Proceedings of the 5th International Conference on Pattern Recognition*, IEEE Catalogue NO 80CH1499–3, 1980, pp. 1324–1347.

[34] C. Dussert, M. Rasigni, J. Palmari, G. Rasigni, A. Llebaria and F. Marty, Minimal spanning tree analysis of biological structures, *J. Theor. Biol.* **125** (1987), 317–323.

[35] D. Cheriton and R.E. Tarjan, Finding Minimum spanning trees, *SIAM J. Comput.* (1976), 724–742.

[36] J.P. Kruskal Jr., On the shortest spannning subtree of a graph of the travelling salesman problem, *Proc. Natl. Acad. Sci. USA* **7** (1956), 48–50.

[37] J. Beardwood, J.H. Halton and J.M. Hammersely, The shortest path through many points, *Proc. Cambridge Philos. Soc.* **55** (1959), 299–327.

[38] S. Verblunsky, On the shortest path through a number of points, *Proc. Amer. Math. Soc.* **2** (1951), 904–913.

[39] S. Ulam, Patterns of growth of figures: mathematical aspects, in: *Proceedings of Symposia in Applied Mathematics, XIV: Mathematical Problems in the Biological Sciences*, American Mathematical Society, 1962, pp. 64–65.

[40] R. Marcelpoil, Normalization of the minimum spanning tree, *Anal. Cell. Pathol.* **5** (1993), 177–186.

[41] G. Haroske, V. Dimmer, D. Steindorf, U. Schilling, F. Theissig and K.D. Kunze, Cellular sociology of proliferating tumor cells in invasive ductal breast cancer, *Anal. Quant. Cytol. Histol.* **18** (1996), 191–198.

[42] G. Bigras, R. Marcelpoil, E. Brambilla and G. Brugal, Cellular sociology applied to neuroendocrine tumors of the lung: quantitative model of neoplastic architecture, *Cytometry* **24** (1996), 74–82.

[43] E. Raymond, M. Raphael, M. Grimaud, L. Vincent, J.L. Binet and F. Meyer, Germinal center analysis with the tools of mathematical morphology on graphs, *Cytometry* **14** (1993), 848–861.

[44] F. Darro, A. Kruczynski, C. Etievant, J. Martinez, J.L. Pasteels and R. Kiss, Characterization of the differentiation of human colorectal cancer cell lines by means of Voronoi diagrams, *Cytometry* **14** (1993), 783–792.

[45] J. Ammermuller, W. Mockel and P. Rujan, A geometrical description of horizontal cell networks in the turtle retina, *Brain Res.* **616** (1993), 351–356.

[46] K. Kayser and H. Stute, Minimum spanning tree, Voronoi's tesselation and Johnson–Mehl diagrams in human lung carcinoma, *Pathol. Res. Pract.* **185** (1989), 729–734.

[47] D. Sulsky, S. Childress and J.K. Percus, A model of cell sorting, *J. Theor. Biol.* **106** (1984), 275–301.

[48] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.

[49] C. Dussert, M. Rasigni, J. Palmari, G. Rasigni, A. Llebaria and F. Marty, Minimal spanning tree analysis of biological structures, *J. Theor. Biol.* **125** (1987), 317–323.

[50] A.G. Percus and O.C. Martin, Finite size and dimensional dependence in the Euclidian travelling salesman problem, *Phys. Rev. Lett.* **76** (1996), 1188–1191.

[51] R. Marcelpoil, Methodologie pour l'etude de la sociologie cellulaire: Application a l'etude du tissu prostatique mormal et pathologique, These, Université Joseph Fourier, 1993.

[52] H. Schulerud, G.B. Kristensen, K. Liestol, L. Vlatkovic, A. Reith, F. Albregtsen and H.E. Danielsen, A review of caveats in statistical nuclear image analysis, *Anal. Cell. Pathol.* **16** (1998), 63–82.

[53] P. Ranefall, K. Wester, A.C. Andersson, C. Busch and E. Bengtsson, Automatic quantification of immunohistochemically stained cell nuclei based on standard reference cells, *Anal. Cell. Pathol.* **17**, 111–123.

[54] P. Ranefall, L. Egevad, B. Nordin and E. Bengtsson, A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei, *Anal. Cell. Pathol.* **15** (1997), 145–156.

[55] C. Dussert, G. Rasigni and A. Llebaria, Quantization of directional properties in biological structures using the Minimal Spanning Tree, *J. Theor. Biol.* **135** (1988), 295–302.

[56] J. Hove, S. Mo and A. Sudbø, Hausdorff dimension of critical fluctuations in abelian gauge theories, *Phys. Rev. Lett.* **85** (2000), 2368–2371.