Modelling Fruit Microstructure Using Novel Ellipse Tessellation Algorithm

H.K. Mebatsion[1], P. Verboven[1], Q. T. Ho[1], F. Mendoza[1], B. E. Verlinden[2], T. A. Nguyen[1] and B. M. Nicolaï[1,2]

Abstract: Modeling plant microstructure is of great interest to food engineers to study and explain material properties related to mass transfer and mechanical deformation. In this paper, a novel ellipse tessellation algorithm to generate a 2D geometrical model of apple tissue is presented. Ellipses were used to quantify the orientation and aspect ratio of cells on a microscopic image. The cell areas and centroids of each cell were also determined by means of a numerical procedure. These characteristic quantities were then described by means of probability density functions. The model tissue geometry was generated from the ellipses, which were truncated when neighbouring areas overlap. As a result, a virtual microstructure consisting of truncated ellipses fills up the entire space with the same number of cells as that of microscopic images and with similar area, orientation and aspect ratio distribution. Statistical analysis showed that the virtual geometry generated with this approach yields spatially equivalent geometries to that of real fruit microstructures. Compared to the more common algorithm of Voronoi tessellation, ellipse tesselation was superior for generating the microstructure of fruit tissues. The extension of the algorithm to 3D is straightforward. These representative tissues can readily be exported into a finite element environment via interfacing codes to perform *in silico* experiments for estimating gas and moisture diffusivities and investigating their relation with fruit microstructure.

keyword: Multiscale model, Gas transport, Water transport, Virtual tissue, Autocorrelation

1 Introduction

Physical, chemical and biological processes at the microscale, though very complicated, are fundamental to understanding the behaviour and characteristics of the fruit tissue [Aguilera (2005)]. The ultimate task becomes the determination of parameters in the leading transport equations on the macroscale given the known structure and heterogeneity of processes at the microstructure [Lammertyn, Scheerlinck, Verlinden, Schotsmans, and Nicolaï, (2001); Lammertyn, Scheerlinck, Jancsok, Verlinden and Nicolaï (2003); Veraverbeke, Verboven, Oostveldt and Nicolaï (2003)(1); Veraverbeke, Verboven, Oostveldt, and Nicolaï (2003)(2); Nguyen, Verboven, Scheerlinck, Vandewalle and Nicolaï (2004); Nguyen, Dresselaers, Verboven, D'hallewin, Culledu, Van Hecke and Nicolaï (2005);Ho, Verlinden, Verboven, Vandewalle, and Nicolaï (2005)]. The transition from the microstructure to macrostructure is then achieved by appropriate homogenization procedures. Because fruit tissue is composed of a very large number of cells, its properties depend not only on the characteristics of individual cells but also on the connectivity and interactions among cellular components. Statistical description and modelling of microstructures are an essential means to implement such an approach. During the past decades, several types of Voronoi based models have been used to represent microstructures of different materials such as aggregates of grains in polycrystals [Nygards and Gudmundson (2002)], composite materials [Ingraffea, Iesulauro, Dodhia, and Wawrzynek (2002); Raghavan and Ghosh (2004)], ceramic microstructures [Espinosa and Zavattieri (2000)] and fruit microstructure [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)]. The Poisson Voronoi diagrams (PVDs) have been extensively used and studied by different authors [Mattea, Urbicain and Roustien (1989); Roudot, Duprat and Pietri (1990); Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)]. Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006) generated Poisson Voronoi diagrams having similar statistical properties (area, orientation and aspect ratio) as that of apple parenchyma microstructures. However, the spatial statistics of the real cells and the representative PVDs were not evaluated.

[1] BIOSYST- MeBioS, Katholieke Universiteit Leuven, W. De Croylaan 42, B-3001 Leuven, Belgium
[2] Flanders Center of Postharvest Technology, W. De Croylaan 42, B-3001 Leuven, Belgium

Autocorrelation statistics are basic descriptive statistics for any data ordered in a sequence (or in space) as they provide basic information about the ordering of the data that is not available from other statistical measures such as mean and standard deviation of variables. Spatial autocorrelation statistics detect the presence of interdependence between data at neighbouring locations and derive the effect upon the values of the measurements. In other words, spatial autocorrelation statistics measure the amount that the measurements depart from the requirements of independence [Odland (1988)]; it may happen that values of a variable sampled at nearby locations are more similar than those sampled at locations more distant from each other. Spatial autocorrelation can occur at multiple spatial scales or vary with spatial orientation. For anisotropic materials, the spatial variability along the x-axis differs from that of the y-axis.

The aim of the presented work was to incorporate spatial statistics properties in to a virtual tissue algorithm. In this respect, this paper presents a novel approach to the generation of statistically and spatially equivalent virtual fruit microstructure using an ellipse tessellation algorithm.

2   Materials and methods

2.1   Microscopic images geometrical parameter calculation

The study was based on images of fruit cellular structures obtained from light microscopy. Sample preparation and image acquisition procedures outlined by Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006) were followed. Samples from apple (*cv.* Cameo), Onion and Conference pear parenchyma tissues were used. The geometrical characteristics of cells were estimated after the transformation of the digital images to representative polygons defined by points on the natural boundary of the cells. From the resulting coordinates, the centroidal points, areas, perimeters, aspect ratios and orientations of each approximated polygonal cells were determined using moment calculation and moment based ellipse-fitting algorithms [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)].

2.2   Virtual tissue generation

2.2.1   Voronoi diagrams

Voronoi diagrams are generated by introducing planar cell walls perpendicular to the lines connecting neighbouring points resulting in a set of convex polygons in 2D, filing the underlying space [Okabe, Boots, and Sugihara (1992)]. Both Poisson Voronoi diagram (PVD) and centroid based Voronoi diagram (CVD) were generated from random set of points and geometric centres, respectively [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)].

2.2.2   Ellipse tessellation

Microscopic fruit cell sections are elliptically shaped (aspect ratio greater than one) [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006); Schotsmans (2003)]. The ellipse-fitting algorithm approximates every cell of the microstructure with the corresponding ellipse that has the same second moment of area, from which aspect ratio and orientation of cellular images is determined.

An ellipse is a conic that can be described by an implicit second order polynomial:

$$F(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \qquad (1)$$

with an ellipse-specific constraint:

$$b^2 - 4ac < 0 \qquad (2)$$

where *a, b, c, d, e, f* are the coefficients of the ellipse and (x,y) are coordinates of sample points lying on it. The polynomial F(x, y) is called the "algebraic distance" of a point (x,y) to the conic $F(x, y) = 0$. By introducing vectors

$$a = [a,b,c,d,e,f]^T$$

$$x = \left[x^2, xy, y^2, x, y, 1\right]^T \qquad (3)$$

the above equation can be rewritten as:

$$F(a;x) = a \cdot x = 0 \qquad (4)$$

Fitting an ellipse to a general conic is to minimize the algebraic distance over the set of data points in the least square sense. To ensure an ellipse-specificity of the solution, because the vector *a* can be arbitrary scaled, the

following constraint equation can be considered instead of equation 2 [Fitzgibbon, Pilu and Fisher (1999)]:

$$4ac - b^2 = 1 \tag{5}$$

This is a quadratic constraint, which may be expressed in matrix form $a^T Ca = 1$ as

$$a^T \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} a = 1 \tag{6}$$

The minimization procedure can be carried out by defining *design vector* χ_i for the individual points (x_i, y_i) as

$$F(a; x_i) = \begin{bmatrix} x_i^2 & x_i & y_i & y_i^2 & x_i & y_i \end{bmatrix} \cdot [a \ b \ c \ d \ e \ f]$$
$$= \chi_i \cdot a \tag{7}$$

Then assembling the design vectors in to a *design matrix* D of size N x 6

$$D = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_N^2 & x_N y_N & y_N^2 & x_N & y_N & 1 \end{bmatrix} \tag{8}$$

Following conic sections fitting to scattered data procedure [Fitzgibbon (1997)] the constraint ellipse-fitting problem reduces to
Minimizing $E = \|D \ a\|^2$, subject to the constraint

$$a^T Ca = 1 \tag{9}$$

Introducing Lagrange multiplier λ and differentiating the following systems of simultaneous equations are obtained (for details refer to Fitzgibbon (1997)].

$$D^T Da - \lambda Ca = 0$$

$$a^T Ca = 1 \tag{10}$$

The solution of the minimization problem represents the best-fit ellipse for the given set of points.

These ellipses were used to quantify the orientation and aspect ratio of cells on a microscopic image. The cell areas and centroids of each cell were also determined by means of a numerical procedure. The model tissue geometry was generated from the ellipses, which were truncated when neighbouring areas overlap. By doing so, as many truncated ellipses as there are cellular images were generated filling the entire cellular space. Fig. 1 shows the mechanism of generating non-overlapping truncated ellipses. For every microscopic cell, an ellipse is fitted and for every fitted ellipse, the algorithm searches a region that is not in the intersection with the rest fitted elliptical regions. This yields a set of truncated ellipses, representing the fruit microstructure.
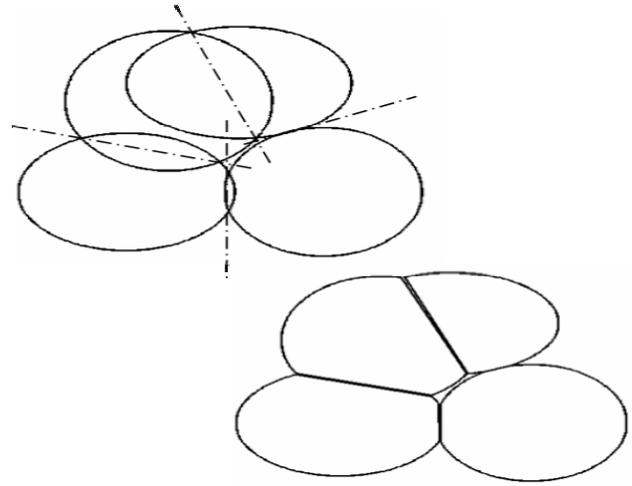

Figure 1 : The ellipsoidal virtual tissue generating mechanism.

2.3 Description of geometric characteristics

2.3.1 Identification of pore spaces

In plant tissue, free spaces called intercellular spaces (or pores) make up a large fraction of a tissue. Intercellular spaces are mainly classified into two types depending on their origin. Those originating from the separation of cells along their edges (called schizogenous intercellular spaces) and others originating from dying cells (lysigenous intercellular spaces) [Kuroki, Oshita, Sotome, Kawageo and Seo (2004)]. Based on their contents, intercellular spaces can also be classified as gas filled and liquid filled. Gas filled intercellular spaces, even though with small volume fraction, are thought to be the diffusion pathways through plant organs for respiration and photosynthesis. Diffusion in these gas-

filled spaces is about 10, 000 times faster than in water [Terashima, Miyazawa and Hanba (2001)]. Hence, accurate identification of pore spaces in the tissue is an important component of microstructural modelling.

Intercellular spaces are clearly visible on a micrograph as cavities of various shapes and sizes as can be seen in Fig.2. These spaces are irregular, elongated objects that are concave towards the centre [Lewicki and Pawlak (2005)]. However, it is not easy to make a distinction between intercellular spaces and cells from microscopic images alone. Even with careful experienced analysis, incurring an error is inevitable. As a result, the identification of intercellular spaces should be supported by the use of a microscope.
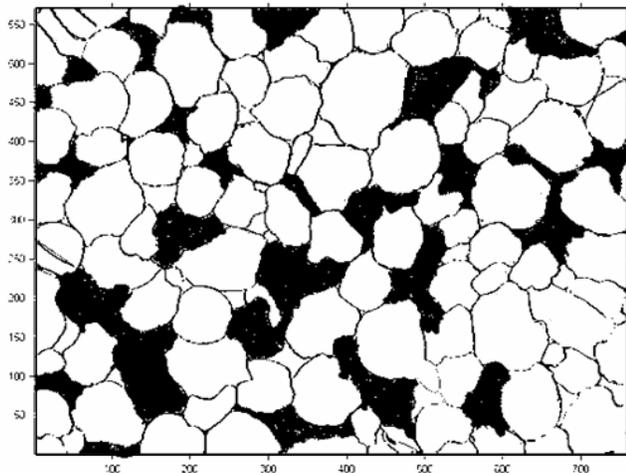

Figure 2 : Micrograph of apple (*cv.* Cameo) parenchyma tissue. Shaded areas are intercellular spaces. (dimensions in pixels; 1pixel=1.85185μm).

2.3.2 *Shape factor, roundness and aspect ratio of cellular materials*

There are different parameters such as shape factor, roundness and aspect ratio which have been used to describe the shape of cellular images [Schotsmans (2003); Lewicki and Pawlak (2005); Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)].

The shape factor, s_f of the cells is defined to be proportional to the ratio of the areas of cellular images to the squares of perimeters:

$$s_f = \frac{4\pi A}{P^2} \tag{11}$$

where A is the area, and P is the perimeter of the cell.

The shape factor is actually the measure of how far cellular images are from being circular or spherical in two and three dimensions, respectively. The shape factor values range between zero and one. The shape factor of circular objects is one.

The roundness of cellular images, on the other hand, is defined as the reciprocal of the shape factor. It is also a measure of the circularity of cells [Schotsmans (2003)]. As the shape factor, the roundness value of circular objects is unity and the higher the roundness, the more the shape of the cell deviates from circle.

The definition of aspect ratio of cellular images is defined as the ratio of the major and the minor axes of the hypothetical ellipse that has the same second moment area as that of cellular images [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)]. The aspect ratio gives more detailed information when used together with the orientation of the cells (from the positive x-axis) as to how the cellular images are situated in spatial domain than the shape factor, which says nothing about spatial arrangement of cells.

2.4 **Statistics**

2.4.1 *Statistical distributions*

Fitting distributions to data sets by assuming a pre-specified distribution (e.g. lognormal) with distribution parameters (e.g. mean and variance) obtained from the data were carried out. Both the assumed and the empirical cumulative distribution functions (CDFs) at each data point were obtained. If the maximum distance between the assumed CDF and the empirical CDF were small, the assumed CDF was accepted as good fit [Justel, Pena and Zamar (1997)]. In this study, the *Kolmogorov-Smirnov test* (K-S test) of Goodness of Fit (GoF) at the 5% level of significance was used. The skewness, and kurtosis values of the microscopic and ellipse tessellated virtual tissues were also calculated to compare the degree of deviation of the area distribution with that of the normal distribution. The skewness coefficient of the normal distribution is zero. The kurtosis coefficient for normal distribution is 3 and values that are greater than 3 are distributions which are slim and have long tails whereas fat distributions with short tails have kurtosis values less than 3.

The skewness coefficient was calculated as follows:

$$S = E\left[\left(\frac{X - \mu_x}{\sigma}\right)^3\right] \qquad (12)$$

where μ_x and σ are the mean and the standard deviation of the variables under consideration.

The kurtosis coefficient was calculated as:

$$K = E\left[\left(\frac{X - \mu_x}{\sigma}\right)^4\right] \qquad (13)$$

2.4.2 *Autocorrelation Statistics*

In order to calculate autocorrelation statistics, it is necessary to express the spatial arrangement in terms of some function that assigns values to pairs of locations in order to represent their location with respect to one another by means of a "*weighting function*" [Cliff and Ord (1972); Odland (1988); Brown (1982); Kelejian and Prucha (2001)]. It is also possible to use a distance function to study the spatial arrangement based on clustering characteristics [Yotte, Riss, Breysse and Ghosh (2004)].

For data sets that are not arranged regularly or represent area units of different size and shape, weights could be defined as distance based weights [Brown (1982); Schabenberger and Pierce (2002)]. In this study, the weighting function was defined as the inverse of the distance between the geometric centres (centroids) of two cellular regions.

For a real valued stationary random field $u(x,y)$ over a spatial domain the autocorrelation function can be expressed as:

$$\rho(h) = \frac{n}{\sum\limits_i^n u_i^2} E(u_j - u_i) \qquad (14)$$

where:

$u_i = u(i) - E(u), E(u) =$ the expectation value of u with h is the separation distance between two regions in the spatial domain (in this case the distance between the geometric centres of two regions); n is the total number of regions in the domain and u is the variable of interest whose correlation statistics is to be measured. In our study, the term $u(i)$ represents the area of the individual cells.

For spatial datasets lacking global spatial association, the result of the autocorrelation function is difficult to interpret. Moreover, determining local spatial autocorrelation statistics from global spatial autocorrelations is not straightforward [Sokal, Oden and Thomson (1981)]. As a result, the need for other spatial variability measures becomes apparent. There are two types of numerical statistic tests for the measure of local spatial variability. These are *Moran's I* and *Geary's C* tests.

2.4.2.1 Moran's I test

Moran's I test is the most popular test for spatial correlation. The test procedure standardizes the variables by subtracting the mean and deflating by a specific factor (usually the variance). The *Moran's I* is defined as:

$$I = \frac{n}{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}} \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} u_i u_j}{\sum\limits_{i=1}^{n} u_i^2} \quad , \quad i \neq j \qquad (15)$$

where w_{ij}, is the inverse of the separation of geometrical centres of two regions.

Moran's I is merely the spatial auto covariance, standardized by the variance of the data sets and the weighting function. The weighting function defines the arrangement of regions on a map. It is the measure of connectivity of the set of regions

Moran's I has an expected value of $E(I) = -\left(\frac{1}{n-1}\right)$. Values of I that exceed the expected value indicate *positive spatial autocorrelation* in which values tend to be similar to neighbouring values. Values of I below the expectation indicate *negative spatial autocorrelation* in which neighbouring values tend to be dissimilar. If $I=0$, then u_i and u_j are independent.

The hypothesis of no spatial autocorrelation is rejected at α (usually 5%) significance level if

$$Z(I) = \frac{|I - E(I)|}{\sigma_I} \qquad (16)$$

is more extreme than the $Z_{\alpha/2}$ cut-off of a standard Gaussian distribution [Cliff and Ord (1970); Cliff and Ord (1981); Brown (1982); Odland1(1988); Schabenberger and Pierce (2002)]. σ_I is the estimate of the standard deviation.

$$\sigma_I = \left[\frac{n^2 S_1 - n S_2 + 3 S_0^2}{S_0^2 (n^2 - 1)}\right]^{\frac{1}{2}} \qquad (17)$$

where $S_0 = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}$, $S_1 = 2\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(w_{ij})^2$ and $S_2 = 4\sum\limits_{i=1}^{n}\left(\sum\limits_{j=1}^{n}w_{ij}\right)^2$

The p-value, the significance of the statistical tests, can be calculated from the standard normal cumulative probability table using the calculated Z value as:

$$p-value = prob\left(Z \geq Z(I)\right) \tag{18}$$

When the global statistics of spatial association is not significant, one can proceed to test the individual local spatial association coefficients by the method described in the next section.

2.4.2.2   Local indicators of spatial association (LISA)

The idea of local indicators of spatial association (*LISA*) is that, although there may be no spatial autocorrelation globally, there might be local pockets of positive or negative spatial autocorrelation in the data, so-called *hot spots and cold spots,* respectively [Schabenberger and Pierce (2000); Sokal, Oden and Thomson (1998)]. Spot definitions based on autocorrelation measures designate sites as unusual if the spatial dependency is locally much different from the other sites. The *LISA* for each observation also gives an indication of the extent of significant spatial clustering of similar values around an observation. The *LISA* permits the decomposition of the global coefficient of spatial association into separate parts, making it possible to identify individual locations that are major contributors of the global association [Sokal, Oden and Thomson (1998)].The *LISA* version of the Moran's *I* is

$$I_i = \frac{n}{\sum\limits_{i}^{n}u_i^2}u_i\sum\limits_{j}^{n}w_{ij}u_j \tag{19}$$

The term $\sum\limits_{j}^{n}w_{ij}u_j$ is formally known as the *spatial lag* for u at location i. In the absence of spatial autocorrelation, I_i has the expected value $E(I_i) = \frac{-\sum\limits_{i=1}^{n}w_{ij}}{n-1}$. The interpretation of *LISA* is that if I_i is smaller than $E(I_i)$, the sites connected to site i (called pivot site) have attribute values dissimilar from attributes at site i. A high (low) value at site i is surrounded by a low (high) values. On the other hand, if I_i is greater than $E(I_i)$, sites connected to site i show similar value. A high (low) value at site i is

surrounded by a high (low) values. The map of *LISA* indicates which regions of the domain behave differently from the rest [Sokal, Oden and Thomson (1998)]

2.4.2.3   Geary's C test

Geary's C test is an alternative statistics for spatial autocorrelation constructed by using another measure of covariation. The test uses the sum of the square difference between pair of data values as a measure of covariation instead of the joint covariance and standardize in a slightly different way to get a correlation defined as follows:

$$C = \frac{n-1}{2\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}w_{ij}}\frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}w_{ij}(u_i-u_j)^2}{\sum\limits_{i=1}^{n}u_i^2} \tag{20}$$

The expectation value of *Geary's C* is 1 (one) for independence among neighbouring values. It is inversely related to *Moran's I* and it does not provide identical inference as it emphasizes the differences in values between pairs of observations, rather than the covariation between pairs. *Moran's I* gives more global indicator, whereas the *Geary's C* coefficient is more sensitive to differences in small neighbourhoods. Moreover, the efficiency of *Moran's I* tests is generally better than that of *Geary's C* [Odland (1988)].

The significance of the Geary's *C* spatial statistics can be calculated in the same way as the Moran's *I* statistics. The standard deviation under normality assumptions given as [Cliff and Ord (1972); Cliff and Ord (1981)]:

$$\sigma_c = \left[\frac{(2S_1+S_2)(n-1)-4S_0^2}{2(n+1)S_0^2}\right]^{\frac{1}{2}} \tag{21}$$

All computations were carried out in the Matlab version 7.0.1-programming environment (The Mathworks, Natick, MA).

3   Results and Discussion

3.1   *The spatial analysis of Voronoi tessellations*

The virtual tissues that were generated using Poisson Voronoi tesselation algorithms had similar geometrical statistical properties of the fruit microstructure [Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006)]. To see whether the virtual tissues resemble that of the real microstructure spatially, *Moran's I* and
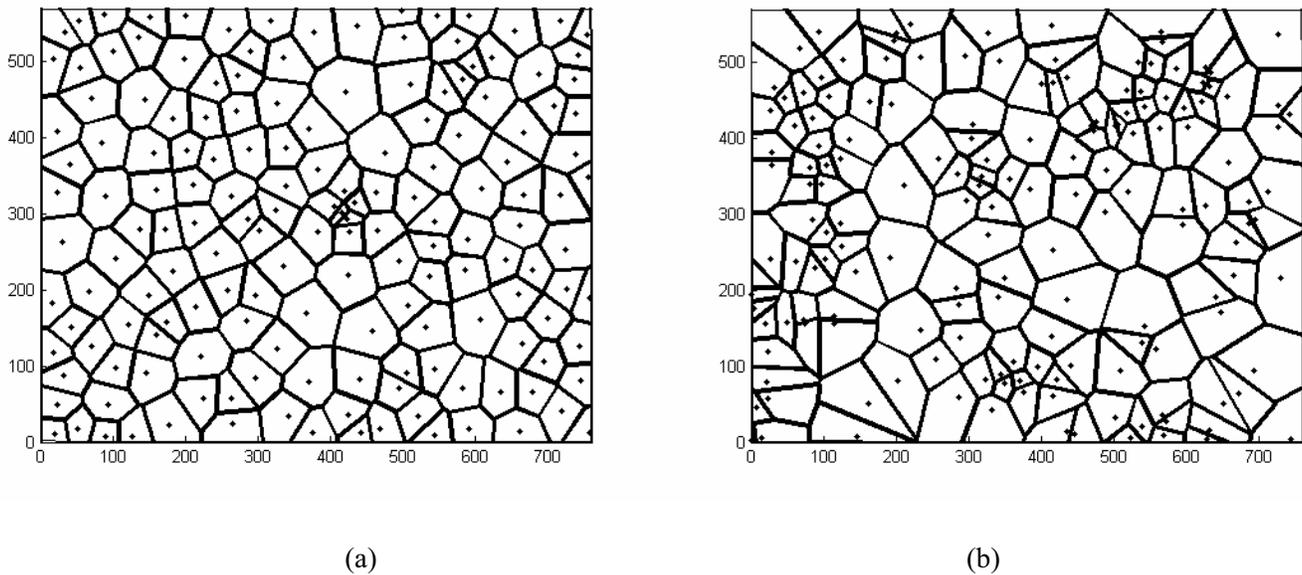
(a) (b)

Figure 3 : Voronoi tessellations of apple parenchyma tissue: (a). Centroid based Voronoi diagram (CVD); (b). Poisson Voronoi diagram (PVD).

Geary's C autocorrelation tests were performed. All the spatial tests showed that both PVD and CVD were far from representing the fruit microstructural tissues. The cellular image, the CVD and the PVD shown in Fig.1 and Fig.3a & 3b had *Moran's I* values of -0.0076, 0.0387, and 0.0485, respectively. A white random (completely uncorrelated) field that has the same number of cells as that of the image could have a *Moran's I* value of -0.0061. The cellular image, therefore, tended towards no spatial correlation. The *Moran's I* values of the Voronoi tissues, on the other hand, were greater than that of the random field, and according to the rule of *Moran's I* test, there were positive spatial autocorrelations between neighbouring Voronoi cells, i.e. nearby values had similar areas. Nevertheless, the degree of similarity was different. More PVD cells were clustered in a certain region than that of the CVD and the microscopic images. This difference could also be detected visually using figure 3a & 3b.

The statistical results on *Moran I* values also support the above conclusion. There was no indication that cellular images were spatially correlated. The Z values of the image, the CVD and PVD were calculated to be equal to 0.1543 (p=0.44), 4.7859 (p=$79*10^{-8}$), and 5.9538 (p=182×10^{-11}), respectively. Since the Z value from the table of the Normal distribution at 5% level of significance is 1.96 and that of the CVD and the PVD are greater than this value, we conclude in favour of the presence of positive spatial correlation in Voronoi tissues. On the other hand, there was no evidence to reject the hypothesis of no spatial correlation in cellular images (Z=0.1543<1.96 or p=0.440>0.05).

The *Geary's C* values of the cellular image, the CVD and the PVD were 1.0084, 0.9343 and 0.8593, respectively. The *Geary's C* value of the cellular images was greater than 1, which is considered to be the value for a white random field, while the corresponding values of the Voronoi geometries were less than 1. This indicated that there was a positive spatial correlation within the Voronoi cells but not with that of the cellular image. The statistical analysis using the Z scale shows that there was no evidence to reject the hypothesis of no spatial correlation in cellular images and CVD cells with Z scale values 0.137 (p=0.444) and 1.0282 (p=0.152), respectively. The Z scale value of PVD, on the other hand, was 2.2528 (p=0.02), which is greater than the Z value of from the table of Normal distribution, and hence we conclude in favour of the presence of positive spatial correlation.

Like the results of *Moran's I* test, PVD seemed to have more clustered cells than the CVD and the cellular images. To investigate the location of the clustered cells and its deviation from randomness, the *LISA* test could be carried out. Fig. 4a-4c show the *LISA* spatial representation of the image, the CVD and the PVD, respectively. The concentration of high LISA values in certain

regions of the PVD (e.g., around x=600 and y=450) indicate clustering of cells with similar area. This is defined in Fig. 3b. The *LISA*, *Moran's I* and *Geary's C* tests clearly showed that the PVD, which had similar geometrical properties as the microscopic images, was far from representing the microscopic tissue spatially.

3.2 *The spatial analysis of ellipse tesselation*

Fig. 2&5a show an apple microscopic image (*cv.* Cameo) and its equivalent ellipse tessellation virtual tissue. The shaded regions represent the intercellular regions. Fig. 5b shows the detailed magnified form of the virtual tissue. The statistical comparison between area, aspect ratio and orientation distributions of the microscopic images and virtual tissues are represented in Fig. 6a-6c. There was a good agreement between all the geometrical property distributions investigated. The skewness and kurtosis of the area distribution of the microscopic images and the ellipse tessellated virtual tissue gave comparable results.

The skewness values of the cellular images and the virtual tissues were 3.71 and 3.80, respectively. Moreover, the kurtosis coefficients of the cellular images and the ellipse virtual tissue were calculated to be 0.98 and 1.00, respectively. Geometrically speaking, all of the above results suggest that the ellipse tessellated virtual is a good representative of the micrograph.

The spatial comparison using *Moran's I* and *Geary's C* tests also showed that the virtual tissue was a good approximation of the real microscopic image. *The Moran's I* values of ellipse tessellated virtual tissue is -0.0071 showing negative spatial association. Concerning the Z scale of the ellipse tesselation, there is no enough evidence to reject the null hypothesis of no spatial correlation (with $Z = 0.1019 < Z_{\alpha/2}$ or $p = 0.46 > 0.05$). Moreover, the *Geary's C* value is calculated to be equal to 1.0093 supports the presence of spatial similarity between the micrograph (with C=1.0084) and the ellipse tessellated virtual tissue. The *LISA* test clearly shows the presence of comparable trend in local spatial associations between microscopic images and ellipse tessellated counterparts (Fig.4a & 7a).

3.3 *Advantages of ellipse tesselation over Voronoi tesselation*

Mebatsion, Verboven, Verlinden, Ho, Nguyen and Nicolaï (2006) implicated the drawbacks of the Voronoi tesselation algorithms to generating virtual equivalents of regular anisotropic tissues such as onion. Such geometries were far from being considered as random and getting their geometrically equivalent virtual tissues was difficult. Moreover, when very large cells were surrounded by very small ones the Voronoi tesselation algorithm fails to generate the virtual equivalents. Because of the characteristics of Voronoi tessellations algorithm, edges must be equally spaced from the generating points and this promotes smaller cells to become larger and vice versa. Both of the above drawbacks of Voronoi tesselation algorithms were overcome by using the new ellipse tessellation algorithm. Fig. 8a & 8b show the microscopic and ellipse based virtual tissues of onion epidermis. As the figures show except for some regions where three or more fitted ellipses met, the ellipse tesselation algorithm generated virtual tissues that resemble the regular anisotropic onion micrograph. Since the algorithm merely search for two non-overlapped regions at a time, it fails to resolve the rare event of finding three or more intersecting fitted ellipses.

On the other hand, as is shown in Fig.9a & 9b, conference pear tissue with small cells surrounded by large ones was accurately described by our new tesselation algorithm. The pore geometries of ellipse tesselation could easily be generated from the microscopic information (assigned as pores in the microscopic images) rather than random generation of pores in the Voronoi tessellations.

4   Applications

The geometry of each ellipse-tessellated diagram, described by the corresponding point coordinates could be transferred into a finite element code (Femlab, COMSOL, Stockholm, Sweden) via an interfacing Matlab code (The Mathworks, Natick, MA). The cell walls were determined by shrinking the cell geometry until the desired cell wall thickness was obtained. The cells, the pores and the cell walls were then exported as separate bodies so that different conditions could be specified. The Femlab mesh generator performed meshing of the tessellation automatically. These meshed representative virtual tissues could be used to perform *in silico* experiments for the simulation of gas and moisture transport in pome fruit tissues under different simulations conditions. Fig. 10 shows the ellipse tessellated diagram in the Femlab environment.
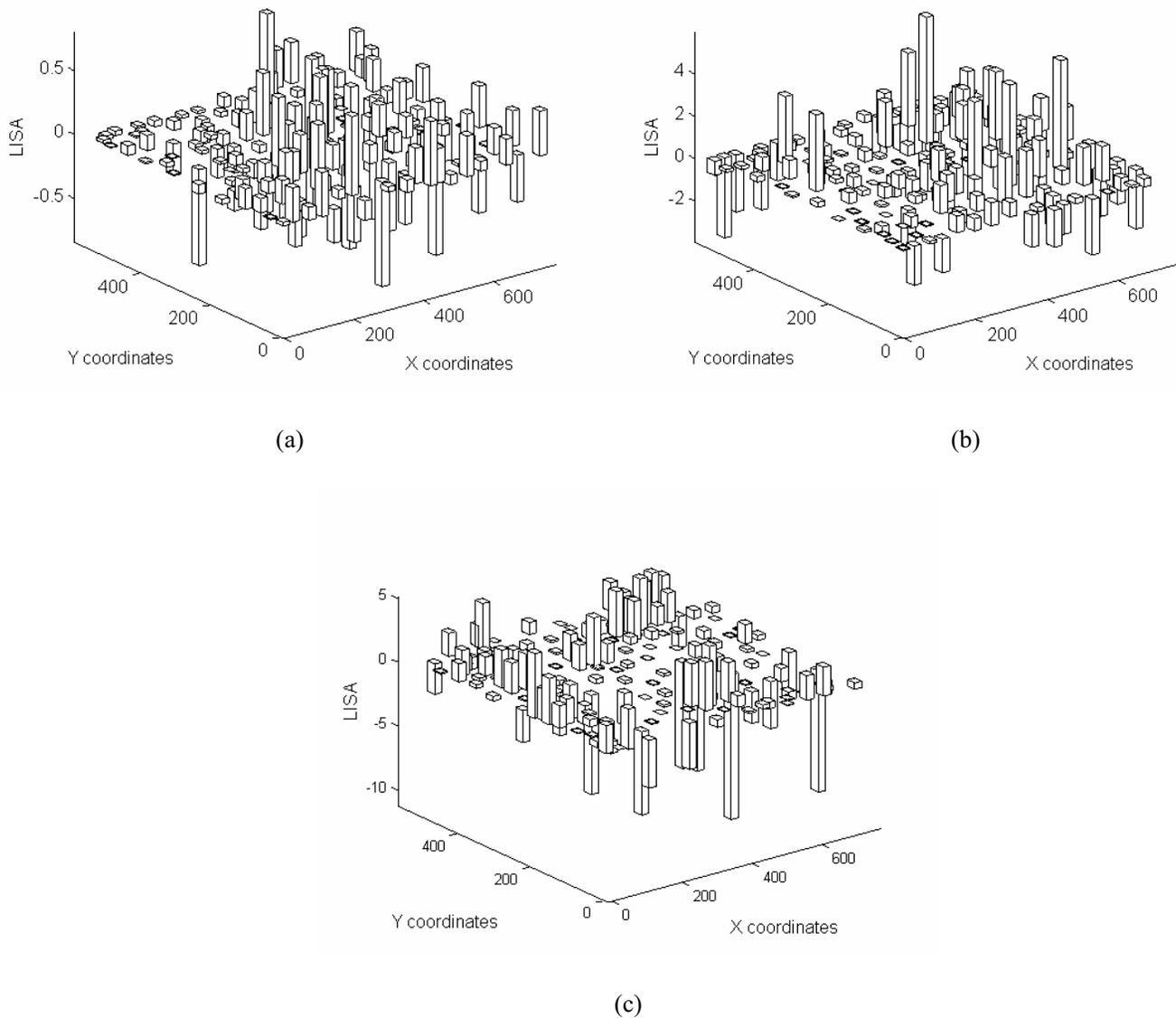
(a)

(b)

(c)

Figure 4 : Spatial comparison between images and Voronoi tessellations of apple parenchyma tissue: (a). Image; (b). Centroid based Voronoi diagram (CVD); (c). Poisson Voronoi diagram (PVD).

5   Conclusions

A novel ellipse tesselation algorithm for generation of fruit was developed. Compared to the more common algorithm of Voronoi diagrams, both the centroid based Voronoi diagrams (CVDs) and the Poisson Voronoi diagrams (PVDs), ellipse tesselation was superior for generating the microstructure of tissue. The difficulty of generating virtual tissues for regular microscopic fruit tissues (such as onion) and tissues with large cells surrounded by smaller ones or vice versa was overcome by our new tessellation algorithm. The ellipse tessellated virtual tis-

sues were proved spatially and statistically representative to the fruit microstructure.

The shape of fruit cells were described in terms of their aspect ratio (defined as the ratio of the major and minor axes of an idealized ellipse that has the same area moments as that of the cells) rather than the shape factor and roundness which were the measure of how cellular geometries deviate from being circular. The use of aspect ratio was better to define the shape of fruit cells than the shape factor or roundness as the measure of ellipsicity includes the measure of circularity.
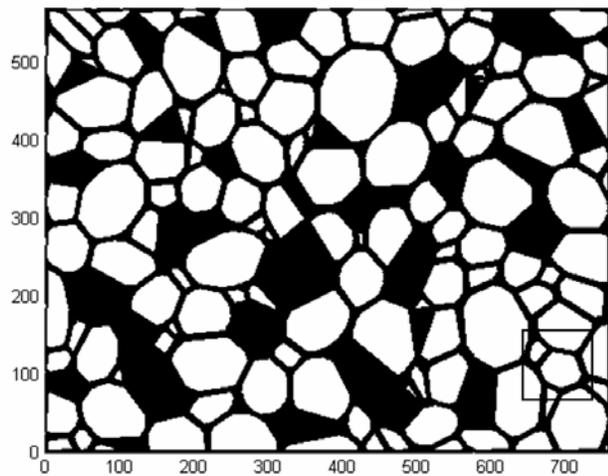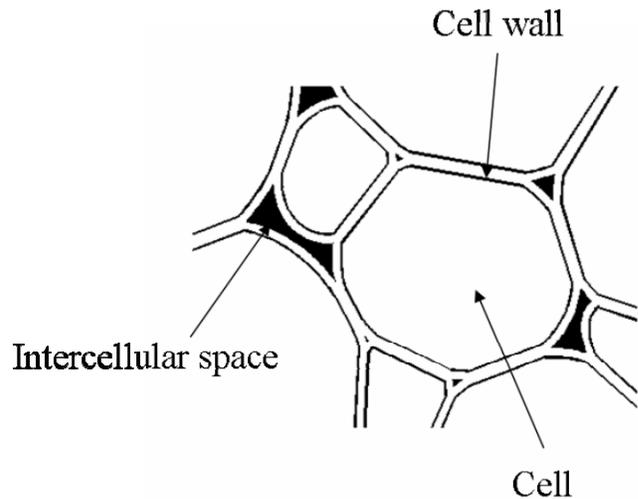
(a)   (b)

Figure 5 : Ellipse tessellated virtual apple parenchyma tissue: (a) virtual tissue; (b) magnified view of a virtual tissue.

Such a microstructural model has large potential in multiscale modelling of gas and moisture transport in fruits. It avoids a tiresome classical continuum modelling approach where transfer properties that appear in the macroscopic transport equation are determined experimentally. As such, the heterogeneous properties of tissue and complex microstructural structures are rather recognized and separate model parameters and boundary conditions are defined to perform *in silico* experiments. This leads to a better understanding of the physics of the process .The transition from microstructure to macrostructure is then achieved by appropriate averaging procedures.

Notation

a, b, c, d, e, f	ellipse coefficients
$F(x, y)$	polynomial function
a, x	vectors
C, D	matrices
λ	Lagrangian multiplier
s_f	shape factor
A	area (μm^2)
P	perimeter (μm)
S	skewness coefficient
μ_x	mean of a variable
σ	standard deviation
K	kurtosis coefficient
$u(x,y)$	stationary random field
$\rho(h)$	autocorrelation function
$u_i = u(i) - E(u)$	deviation from the mean
n	number of regions
I	Moran's coefficient
I_i	local Moran's coefficient
C	Geary's coefficient
Z	Z scale coefficient

References

Aguilera, J. M. (2005). Why food microstructure? *J. Food Eng.,* vol. 67, pp 3-11.

Brown, M. A. (1982). Modelling the spatial distribution of suburban crime. *Economic Geography*, vol. 58(3),pp 247-261.

Cliff A.D.; Ord K. (1972).   Spatial autocorrelation:
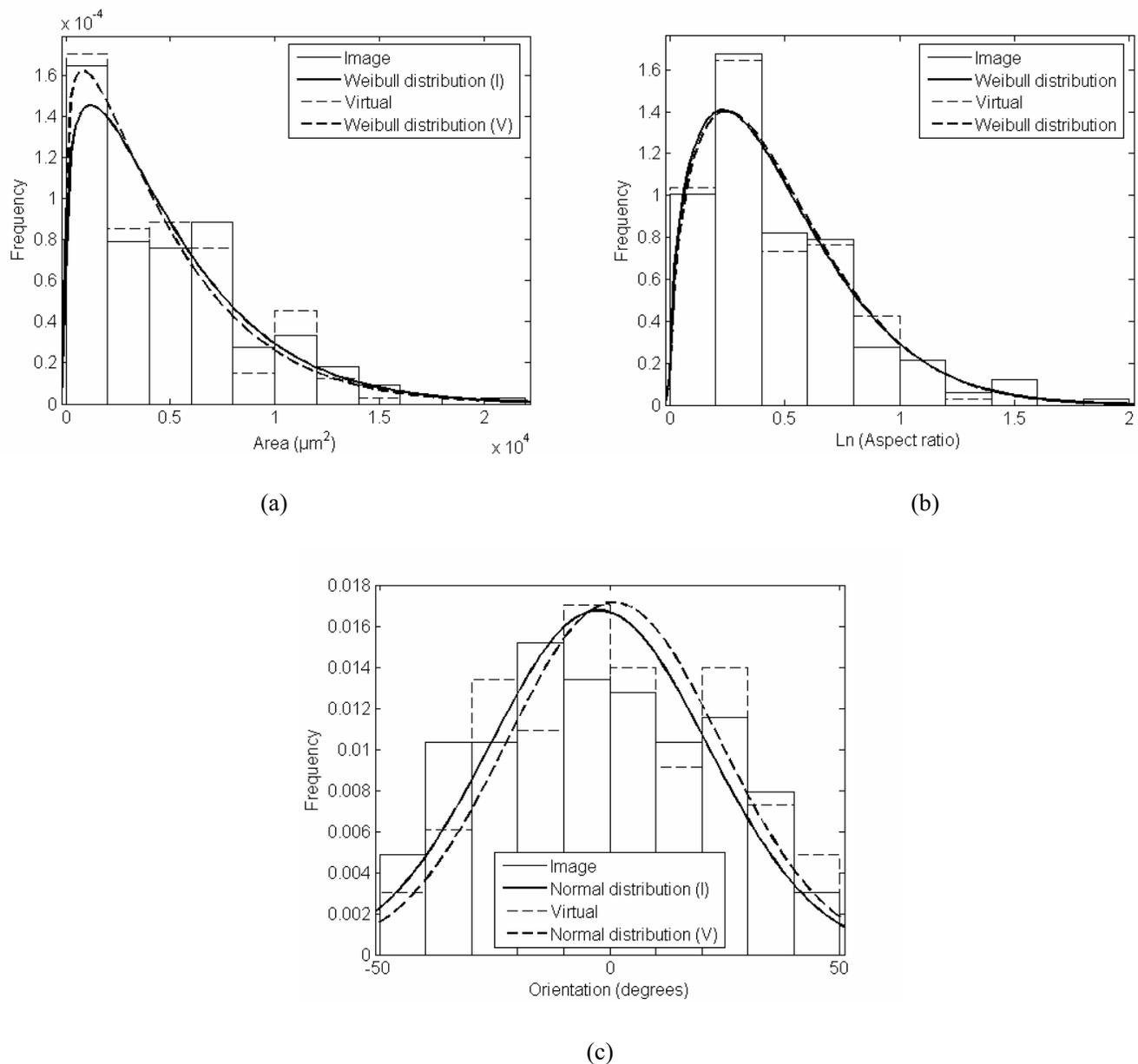
(a)

(b)

(c)

Figure 6 : Geometrical properties of cellular images and virtual apple parenchyma tissue: (a) area distribution; (b) aspect ratio distribution; (c) orientation distribution.

A review of existing and new measures with applications. *Economic Geography. International Geographical Union Proceedings. Commission on Quantitative methods*, vol. 46, pp 269-292.

Cliff, A.D.; Ord, J.K. (1981). Spatial processes: *Models & applications.* London, UK

Espinosa, H.D.; Zavattieri, P. (2000). Modeling of ceramic microstructures: Dynamic damage initiation and evolution. *Amer. Inst. Physics,* pp 333-338.

Fitzgibbon, A.; Pilu, M.; Fisher, R. B. (1999). Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intel.,* Vol. 21(5), pp 476-480.

Fitzgibbon, A. (1998). ”Stable segmentation of 2D curves” *PhD thesis*, University of Edinburgh, UK, 125p.

Ho, Q.T.; Verlinden, B.E.; Verboven, P.; Vandewalle, S.; Nicolaï, B.M. (2005). Simultaneous measurement of oxygen and carbon dioxide diffusivities in pear fruit tissue using optical sensors. *J. Sci. Food and Agric.* (Ac-
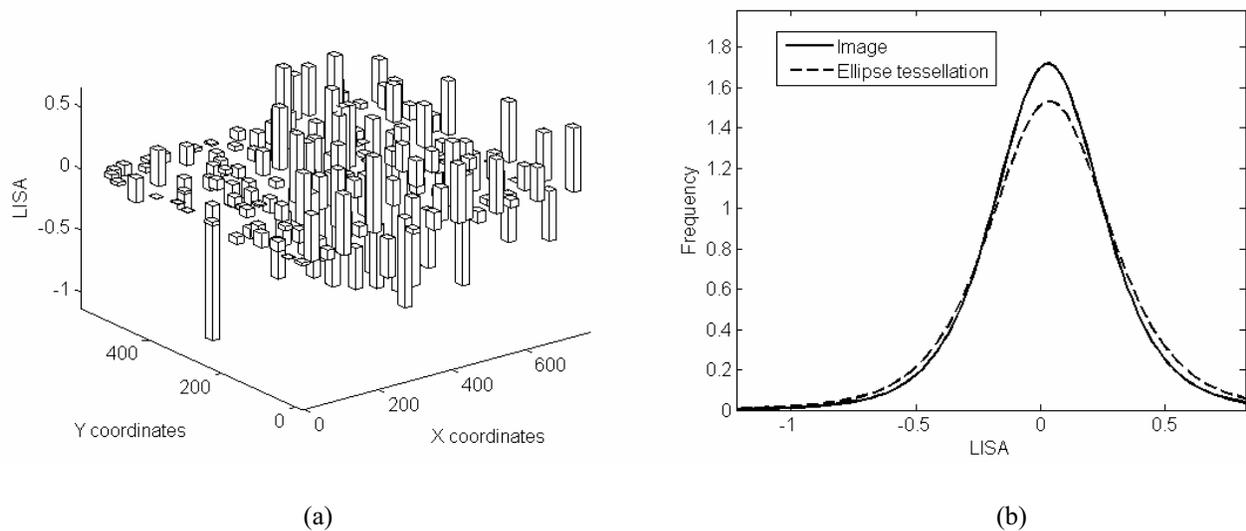
(a)   (b)

Figure 7 : LISA test of the microscopic and ellipse tessellated apple parenchyma tissue: (a) LISA of Ellipse tessellated virtual tissue; (b) LISA frequency distribution.
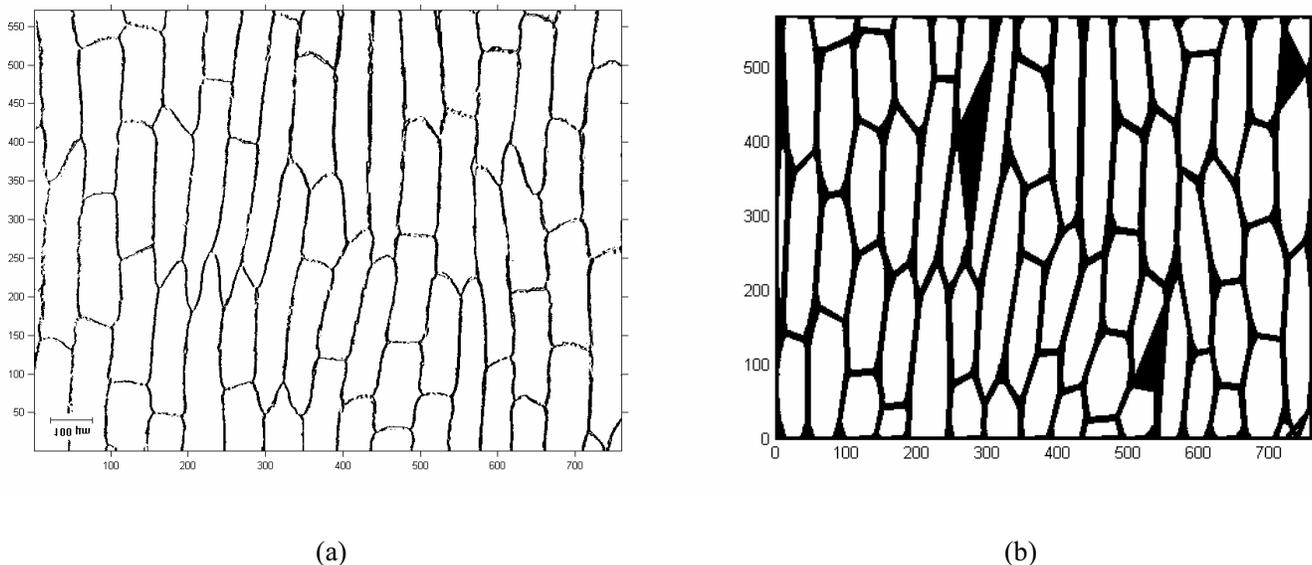

(a)   (b)

Figure 8 : a) Micrograph of Onion epidermis tissue; b) Ellipse tessellated virtual onion epidermis tissue.

cepted).

Ingraffea, A.R.; Iesulauro, E.; Dodhia, K.; Wawrzynek, P.A. (2002). A multiscale modeling approach to crack initiation in aluminium polycrystals. *Fifth World Congress on Computational Mechanics*, Vienna, Austria.

Justel, A.; Pena, D.; Zamar, R. (1997). A multivariant Kolmogorov-Smirnov test of goodness of fit. *Statistics and Probability Letters,* vol. 35, pp 251-259.

Kelejian, H. H; Prucha, I. R. (2001). On the asymptotic distribution of the *Moran I* test statistics with applications. *Journal of Econometrics*, vol. 104, pp 219-257.

Kuroki, S.; Oshita S.; Sotome,I.; Kawageo, Y.; Seo, Y. (2004). Visualization of 3-D network of gas filled intercellular spaces in cucumber fruit after harvest. *Postharvest Biol. & Technol*., vol. 33, pp 255-262.

Lammertyn, J.; Scheerlinck, N.; Jancsok, P., Verlinden B.E.; Nicolaï, B.M. (2003). A respiration-diffusion

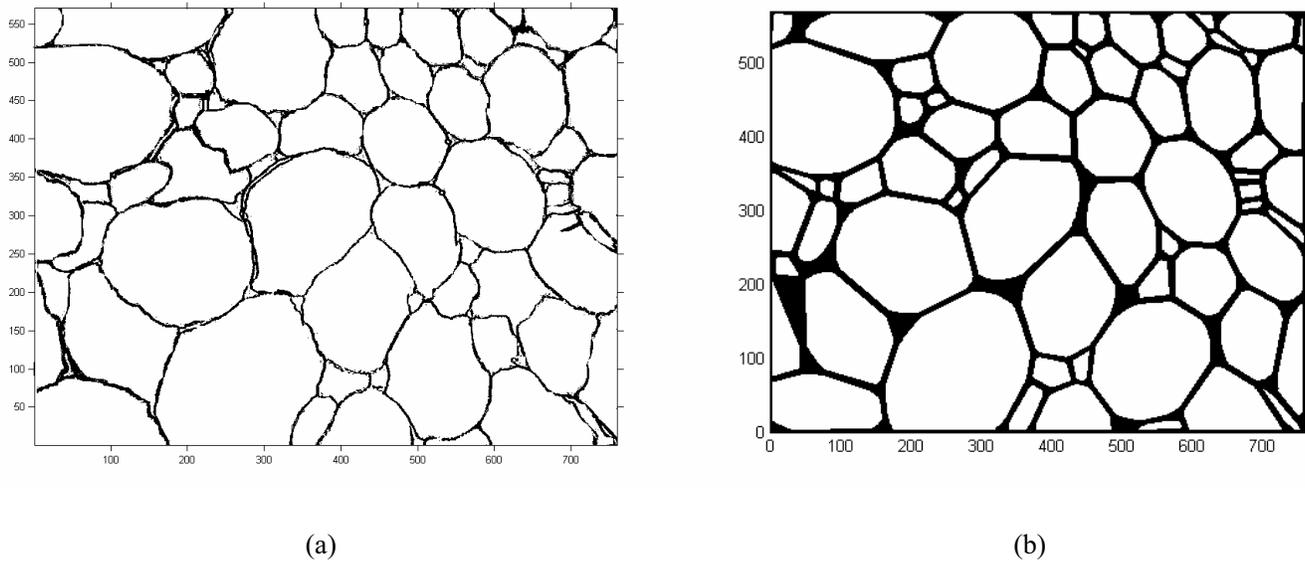(a)   (b)

Figure 9 : a) Micrograph of Conference pear parenchyma tissue; b) Ellipse tessellated virtual conference pear parenchyma tissue.
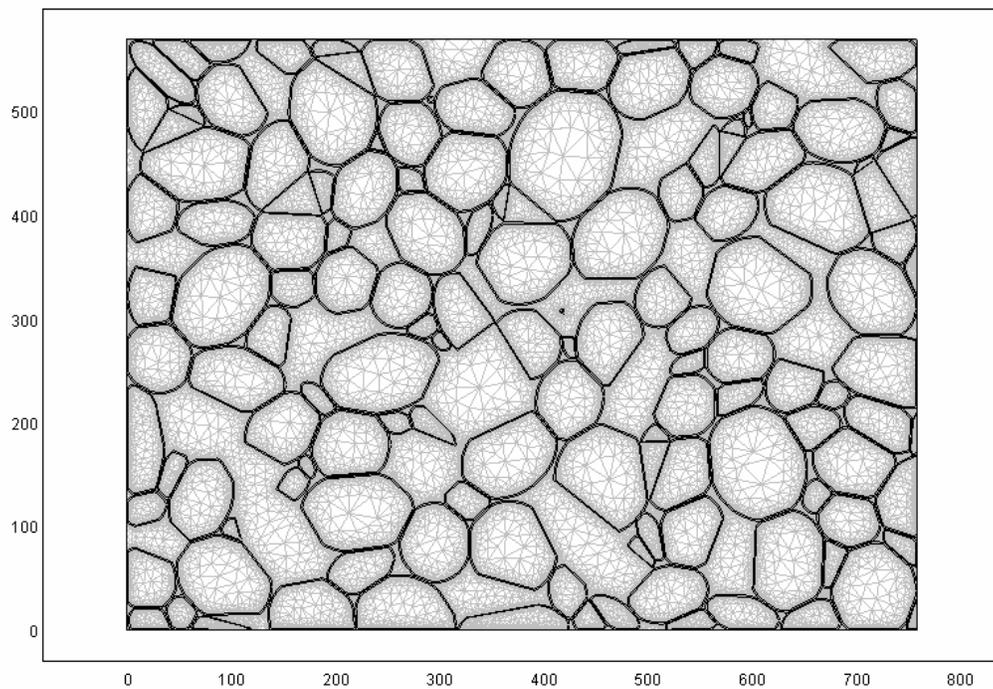

Figure 10 : Meshed ellipse tessellated virtual apple parenchyma tissue (*cv.* Cameo).

model for 'Conference' pears: model development and validation. *Postharvest Biology and Tech.* vol. 30, pp 29-42.

Lammertyn, J.; Scheerlinck, N.; Verlinden, B.E.; Schotsmans, W.; Nicolaï, B.M. (2001). Simultaneous determination of oxygen diffusivity and respiration in pear skin and tissue. *Postharvest Biology and Tech.* vol. 23, pp 93-104.

Lewicki, P .P.; Pawlak, R.P. (2005). Effect of osmotic dewatering on apple tissue structure. *J. Food Eng*. vol.

66, pp 43-50.

Mattea, M.; Urbicain, M.J.; Rotstein, E. (1989). Computer model of shrinkage and deformation of cellular tissue during dehydration. *Chem. Eng. Sci.*, vol. 44, pp 2853-2859.

Mebatsion, H. K.; Verboven, P.; Verlinden, B. E.; Ho, Q.T.; Nguyen, T.A.; Nicolaï, B. M. (2006). Microscale modelling of fruit tissue using Voronoi tessellations. *Comput. Electron. Agric.* (in press).

Nguyen, T.A.; Dresselaers, T.; Verboven, P.; D'hallewin, G.; Culledu, N.; Van Hecke P.; Nicolaï, B. M. (2005). Finite element modelling and MRI validation of 3D transient water profiles in pears during postharvest storage. *J. Sci. Food and Agric.* vol. 86, pp 745-756.

Nguyen, T.A.; Verboven, P.; Scheerlinck, N.; Vandewalle, S.; Nicolaï, B.M. (2004). Estimation of effective diffusivity of pear tissue cuticle by means of a numerical water diffusion model. *J. Food Eng.*, vol. 72, pp 63-72.

Nygards, M.; Gudmundson, P. (2002). Micromechanical modelling of ferritic/ pearlitic steels. *Material Science and Eng.* vol. 325, pp 435-443.

Odland, J. (1988). Spatial autocorrelation. *Scientific geography series*. Sage publication, USA.

Okabe, A.; Boots, B.; Sugihara, K. (1992). Spatial tessellations: concepts and applications of Voronoi diagrams. *Wiley Series in Probability and Mathematical Statistics*, John Wiley and Sons, Chichester, UK.

Raghavan, P.; Ghosh, S. (2004). Adaptive multiscale computational modelling of composite materials. *CMES: Computer Modeling in Engineering & Sciences* vol. 5 (2), pp 151-170.

Roudot, A. C.; Duprat, F.; Pietri, E. (1990). Simulation of a penetrometric test on apples using Voronoi-Delaunay tessellation. *Food Structure*, vol. 9, pp 215-222.

Schabenberger, O.; Pierce, F. (2002). Contemporary statistical models. CRC press, USA.

Schotsmans, W. (2003). Gas diffusion properties of pome fruit in relation to storage potential. *PhD thesis*, Katholieke Universiteit Leuven, Leuven, Belgium, 186 p.

Sokal, R.R; Oden, N.L; Thomson, B. A. (1998). Local spatial autocorrelation in biological variables. *Biol. J. Linn. Soc.,.* vol. 65, pp 41-62.

Terashima, I.; Miyazawa, S.; Hanba, T.Y. (2001). Why are sun leaves thicker than shade leaves? Consideration based on analyses of CO2 diffusion in the leaf. *J. Plant Res.,* vol. 114, pp 93–105.

Veraverbeke E.A.; Verboven, P.; Oostveldt, P.V.; Nicolaï, B.M. (2003). Prediction of moisture loss across the cuticle of apple (Malus sylvestris subsp. mitis (Wallr.)) during storage: Part 1. Model development and determination of diffusion coefficients. *Postharvest Biol. Tec.,* vol. 30, pp 75-88.

Veraverbeke E.A.; Verboven, P.; Oostveldt, P.V.; Nicolaï, B.M. (2003). Prediction of moisture loss across the cuticle of apple (Malus sylvestris subsp. mitis (Wallr.)) during storage: Part 2. Model simulation and practical applications. *Postharvest Biol. Tec.,* vol. 30, pp 89-97.

Yotte, S.; Riss, J.; Breysse, D.; Ghosh, S. (2004). PMMC cluster analysis. *CMES: Computer Modeling in Engineering & Sciences* vol. 5 (2), pp 171-187.