UNIVERSITY OF LONDON THESIS

Degree PhD Year 2005 Name of Author MAHESWARAN, S

Computational Modelling Of Epithelial Cell Interactions

Satheesh Maheswaran,
Biomedical Informatics Unit,
Eastman Dental Institute for Oral Health Care Sciences,
University College London

A thesis submitted for the degree of Doctor of Philosophy.

April 17, 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

This thesis describes and evaluates approaches to computational models of Epithelial cell interactions. It begins with a review of existing approaches and in particular includes a rational reconstruction of a model of squamous epithelial cell interactions previously described by Stekel [Stekel et al., 1995]. Suggestions for improving this model are made, including methods for analysing spatial clusters of cells using Delaunay triangulation and heterogeneity of epithelial tissue using connective component labeling.

Histological images of oral epithelium are used to develop a classification of basement membrane shape in normal and dysplastic tissues. The method combines Fourier descriptors for shape representation, Principal Component Analysis for data reduction and the closest mean and support vector machine algorithms for pattern recognition. This approach is suggested as a general technique for evaluating the output of simulation models which involve curvilinear features in a shape-based classification of the tissues modelled.

A new model of epithelial cell interactions is proposed by extending the Glazier-Graner framework for cell sorting [Graner, 1993, Graner and Sawada, 1993]. The model includes biological processes such as cell division, differentiation, adhesion and death. In particular, the roles of differential adhesion and cell division during the development of epithelium are discussed. The typical ordered structure of a healthy epithelium is shown to arise provided differential adhesion and cell division are modelled appropriately.

Contents

Acknowledgements

Like most research, this thesis owes a great deal to the environment in which it was conceived. I would like to thank my supervisors Peter Hammond and Paul Speight for providing enthusiastic encouragement and reading many versions of my thesis. Thanks are also due to many colleagues who kindly provided their time for thought-provoking discussions. I would also like to thank EPSRC and Eastman Dental Institute for providing financial assistance. Finally, I would like to thank my friends, too many to mention by name here, who supported me unreservedly through out good and bad times.

Chapter 1

Introduction

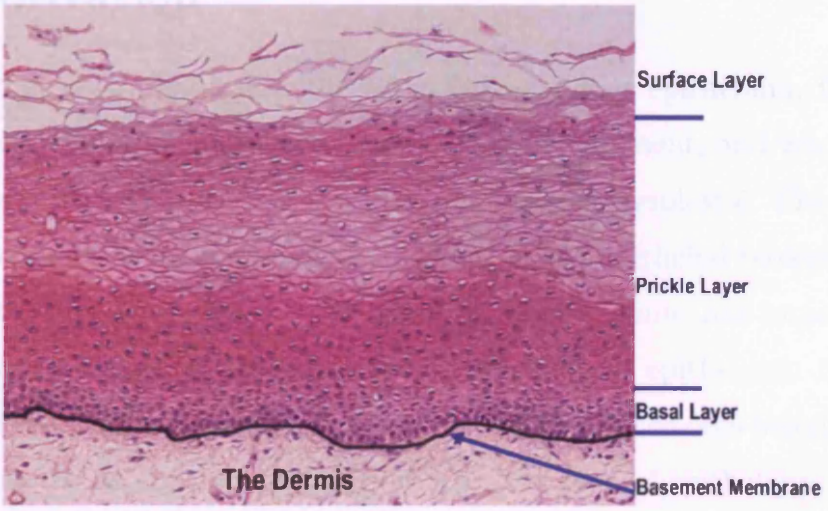
The emergence of multicellular tissue structure is an important process in embryonic development. During development, individual cells need to adhere to each other in order to function correctly as a tissue structure. In doing so, the cells develop ways of organizing themselves into compartments and communicating with each other. The compartment boundaries have fascinated biologists for some time and these boundaries play an important role in pattern formation. During the development of a multi-cellular organism, cells not only have to proliferate and differentiate but also have to sort themselves into different tissue structures. In addition, cells within tissues sometimes sort out or separate from each other. The Differential Adhesion Hypothesis (DAH) [Steinberg, 1963] has long been suggested as one possible hypothesis for cell sorting during embryonic development.

Epithelium is one of the primary tissues in animals as well as plants. There are at least seven different types of epithelia: (a)*squamous* epithelial cells are shaped like flat plates and may be found in cells lining the alveoli in the lung and renal glomerulus; (b)*cuboidal* epithelial cells are extensively involved in secretion and absorption and can be found in much of the renal tubular system; (c)*columnar* epithelial cells are tall and rectangular, highly polarized, and may be found in the small intestine; (d)*pseudostratified columnar* epithelial cells are usually found lining the airways such

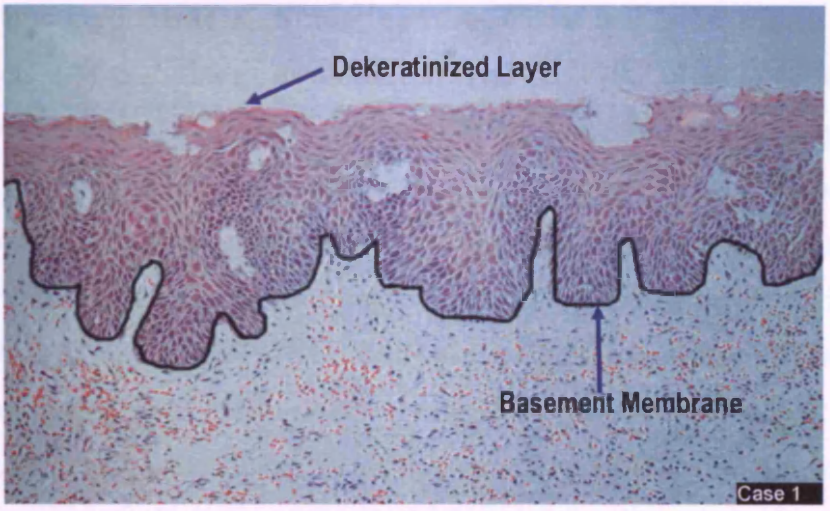
as the trachea. Here, the epithelium consists of a basal layer as well as a differentiated layer; (e)*stratified squamous epithelium* is where cells are multi-layered. Here the basal cells take over the proliferative function of the tissue, giving rise to differentiated cells which flatten and mature as they migrate through the epithelium. The stratified squamous epithelia are usually covered with fluid and can be found in areas such as the mouth, cornea and oesophagus; (f)*stratified cuboidal and columnar epithelia* are relatively sparsely distributed and can be found in sweat gland ducts; (g)*transitional epithelium* can be found lining the urinary tract in the bladder and upper urethra. Here the surface cells are large while the basal cells are smaller.

The epidermis consists primarily of keratinocytes. Scattered among the keratinocytes are a few other cell types like melanocytes, Langerhans cells, and Merkel cells. The quality of epidermis differs from place to place in the body and can also be altered by various disease states which influence the rate of cell division and differentiation. We focus on *Stratified squamous* epithelium. The epidermis consists of several layers. The layers are not distinctly different cells, but reflect visible changes along the process of keratinocyte maturation or keratinization. The *stratified squamous* epithelium is usually divided into 4 layers: the *basal layer*; the *prickle* or *spinous layer*, also called the *Malpighian layer*; the *granular layer*; and finally the *keratin* layer. Figure 1.1(a) shows a histological section of buccal epithelium.

During Epithelial Dysplasia (ED), which effects the palate, tongue and floor of the mouth, the structure of epithelium changes greatly. The shape of the basement membrane becomes more curvy and bulbous, basal cells detach themselves from the basement membrane and move into the differentiated layers. The biological hypothesis for such changes in the structure of epithelium include increased cell proliferation and changes in cell-cell adhesion [Thomas and Speight, 2001]. Figure 1.1(b) shows a histological section of a dysplastic epithelium.



(a)



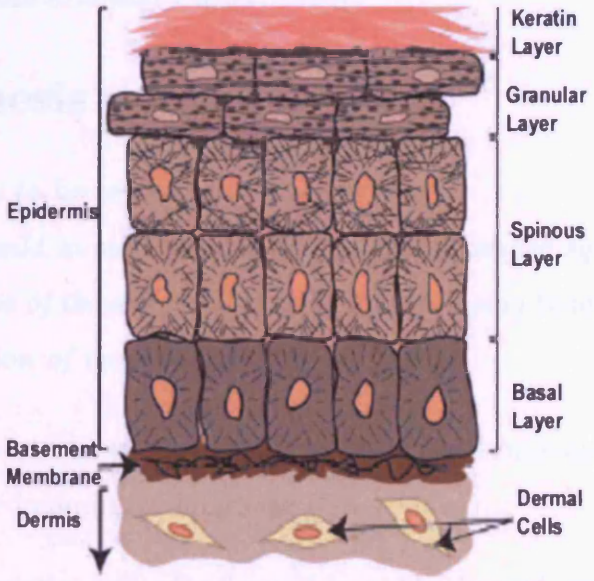
(b)

Figure 1.1: (a)The epithelium shown above is taken from buccal mucosa which is classified as a *stratified squamous* epithelium. The connective tissue lies beneath the epithelium, separated from the epithelium by the basement membrane (black line). The basal layer cells are attached to the basement membrane by hemi-desmosomes. When a basal cell divides, the daughter migrates upward to replenish outer layers of cells. Cells of the "prickle-cell" layer are attached to one another by desmosomes. These cells gradually move outward as new cells are formed from the basal layer. Cells in the granular-cell layer accumulate keratohyalin, visible as darkly stained granules before moving off to the keratinized layer and eventually get taken off. (b)A histological section of a dysplastic oral epithelium. As can be seen in the image, pronounced curved rete pegs is a highly indicative feature of severe cases as well as atrophy of the epithelium above retepegs. The basal cells are also completely distributed throughout the epithelium and the top layer is dekeratinized. Excessive numbers of lymphocytes can also be found in the connective tissue.

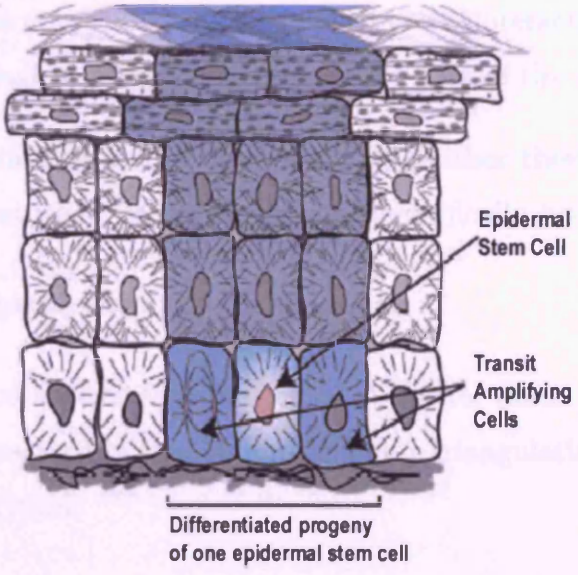
1.1 Motivation

Our motivation is two-fold: how the *Stratified squamous* epithelium, from here on referred to as the epithelium, is organised during development; and what causes the change in tissue structure when the epithelium becomes dysplastic. The morphogenesis of epithelium and its compartment has fascinated epithelial biologists for some time [Wright and Alison, 1984]. The stacking in epithelium has important consequences for cell-shedding mechanisms and the stability of epithelium. Since cells in the *keratin* layer are dead, stacking must be determined by the cells belonging to lower layers. During the establishment of an ordered structure of epithelium, cells formed in the basal layer, migrate into the Malpighian layer, flattening and moving laterally towards the top of epithelium. It has been known since the 1970s that epithelium is organised into columns of maturing layers as shown in Figure 1.2(a). Researchers initially anticipated that the entire basal cell layer consisted of stem cells [Mackenzie, 1997]. However, radiation dose-survival studies have suggested that only $2-7\%$ are stem cells in the basal layer [Mackenzie, 1997]. The current view of epithelial maintenance consists of periodically dividing stem cells with slow cycling times, giving rise to transit amplifying cells which populate most of the basal layer, dividing two or three times before moving upward and differentiating into mature cell layers 1.2(b).

We use computer models, from here on referred to as *in silico* models, to study the development of tissue architecture in epithelium. The *in silico* models may be built using existing biological results obtained from experiments and the literature. They then have to be validated before being used to test new biological hypotheses. The *in silico* models could simulate biological experiments before expensive and time consuming laboratory research is undertaken. We choose to concentrate our simulations at the cellular level using Individual Based Model approaches [Reynolds, 1987]. We focus on simulating biological processes which are widely believed to be the driving force in cell sorting and the formation of ordered structures and compartments in



(a)



(b)

Figure 1.2: (a)Diagrammatic representation of skin epithelial histology. Cells of the basal layer attach to an underlying basement membrane. Basal cells are mitotically active, but they lose this potential when they detach from the basement membrane and embark on the outward trek toward the skin surface. (b)Diagram of the epidermal proliferative unit. A slow-cycling epidermal stem cell occasionally divides, giving rise to a stem cell daughter and a transiently amplifying daughter. The transiently amplifying cell divides two to four times, and these progeny then leave the basal layer and execute a program of terminal differentiation. The description and figures are reproduced here from Alonso [Alonso and Fuchs, 2003]

biological tissues [Dahmann and Basler, 1999].

1.2 Hypothesis

The main hypothesis to be tested in this thesis is :

It is possible to build in silico models to simulate stratified squamous epithelial architecture on the basis of the differential cell adhesion using behavioural rules together with judicious selection of model parameters.

- By "possible" we mean implementing the biological rules on a personal computer using a standard computer language (C++/Java)
- The term differential cell adhesion refers to the hypothesis suggested by Steinberg [Steinberg, 1963] for cell sorting in cell populations.
- Behavioural rules governing cell-cell and cell-tissue interactions are derived from the existing biological literature and form the core of the *in silico* model.
- Model input parameters are chosen such that either they are supported from the biology literature or can be otherwise scientifically justified.

The subsidiary hypotheses in this thesis include :

- **Chapter 3:** The spatial organisation of epithelial tissue architecture may be quantitatively analysed by combining Delaunay triangulation and the connected component algorithm.
- **Chapter 4:** Curvilinear features in tissue shape profiles may be classified by using Fourier descriptors for shape description, Principal Component Analysis for dimensionality reduction and closest mean classification for discrimination.

1.3 Contributions

The contributions of this thesis include: a rational reconstruction of the squamous epithelial model proposed by Stekel [Stekel et al., 1995]; a method to quantitatively measure cell clusters in tissue; a method to evaluate output from *in silico* models using histological images; and, finally a model of epithelial cell-cell interaction using the Glazier-Graner cell sorting algorithm.

We rationally reconstructed Stekel's model to familiarise ourselves with the Individual Based Modelling (IBM) approach. We report on areas of concern in the model and suggest ways of overcoming some of these problems. Another contribution is that we present a method which combines Delaunay triangulation and connected component labelling to quantitatively analyse the ordering of a tissue structure. This method is applied to epithelial tissue to measure the level of heterogeneous ordering in the tissue architecture.

Additionally, we present a method to classify curvilinear features. This combines Fourier Descriptors for shape representation, Principal Component Analysis for data reduction and closest mean and support vector machine algorithms for pattern recognition. The method is applied to histological images of oral epithelium to discriminate between normal and dysplastic tissue structures. We believe this approach enables histological images to be used to evaluate the output from an *in silico* model of epithelial tissue.

A further contribution is the model of epithelial cell-cell interaction. We include in the model biological processes such as cell division, differentiation, adhesion and death. Glazier-Graner's cell sorting algorithm is used to model cell sorting in epithelium. Our model is based on the Differential Adhesion Hypothesis suggested by Steinberg [Steinberg, 1963] and we apply this to epithelium. The model is evaluated by simulating experiments and comparing the results with the existing biological literature. These are the methodological innovations at the heart of this thesis.

1.4 Structure of thesis

Chapter 2 provides a background to the Individual Based Model approach and its application to modelling cell-cell interactions.

Chapter 3 gives details of the reconstruction of Stekel's model for squamous epithelium. We also discuss here our method for analysing spatial structures in the epithelium.

Chapter 4 discusses the proposed method for classifying curvilinear features. We provide an introduction to Fourier descriptors and the classification techniques before moving on to explain our method and its application.

Chapter 5 introduces our model of epithelial cell-cell interaction. Here, we explain our cell behaviour rules and cite the biological evidence to support them. We conclude the chapter by showing results of the simulation runs obtained using various combinations of input parameters.

Chapter 6 uses the biological literature on the development of compartments in epithelial tissue architecture to validate our model output. It has been suggested that differential cell adhesion may be the driving force in the emergence of compartments in Drosophila [Dahmann and Basler, 1999, McNeill, 2000]. We use our model to see if this may be the case in epithelium.

Chapter 7 brings the main conclusions of the thesis together. Here we also discuss possible future work and how our model could be improved.

Chapter 2

Background

In this chapter, we establish a research context by considering how this thesis fits into the existing literature. We first introduce some concepts related to the thesis and then provide a selective discussion on computational models of cell-cell interactions. We then discuss ways of evaluating the *in silico* model outputs using histological images. We conclude our discussion by outlining the morphology of epithelium and review implementations of cell sorting algorithms suggested by Glazier [Graner and Glazier, 1992]. Given the interdisciplinary nature of our work, the reader may find some concepts basic.

2.1 Individual Based Models (IBM)

Individual Based Models (IBM) simulate global consequences of local interactions of members of a population [Reynolds, 1987]. These individuals may represent cells, plants, animals or people in crowds. An individual based model typically consists of a framework in which rules of interactions are defined for the individuals and behavioural characteristics are tracked through time. IBMs are also known as agent based models and each agent corresponds to an individual in the simulated domain. IBMs are of three types : Continuous Individual Based Simulation, Discrete Individual-based Simulation and Cellular Automata (CA).

2.1.1 Continuous Individual Based Simulation

The best example which describes this type of IBM is the "boid" [1] model proposed by Reynolds [Reynolds, 1987] to simulate the flocking behavior of birds. Reynolds placed a large number of boids in a virtual environment and programmed them to follow three simple rules.

- Separation: Maintain a minimum distance from other objects and obstacles in the environment
- Alignment: Match velocities with other boids in the neighbourhood
- Cohesion: Move toward the perceived center of mass of boids in the neighbourhood

What is interesting, is to note that none of these local rules says "form the flocks" and the rules only refer to what an individual "boid" could do and see in its own activity. A flock can only form bottom-up as an emergent phenomenon and they do form every time [Reynolds, 1987]. Reynolds extended the concept of bird flocking (steering) behaviour to the entertainment industry by applying behavioural rules to autonomous characters in animation to perform a wide range of motion behaviour[2].

2.1.2 Discrete Individual Based Simulation

Discrete individual based models are used to model complex objects which can be generated from a collection of simple single objects. In this type of modelling, what is being formed is the space itself rather than individual elements. For example, a settlement can be generated by aggregating a collection of houses [Hillier and Hanson, 1984, Hillier, 1996]. The elementary objects in the model are square cells and the rule of addition of square cells is a full face-wise join. The square cells are not

[1] an autonomous bird like agent
[2] http://www.red3d/cwr/steer/gdc99/index.html

allowed to join at the vertices. During the aggregation process, objects may be added randomly, but under the constraint that each cell must retain at least one side free from other square cells. This process results in a dense and continuous aggregate of cells containing a number of void spaces irrespective of the sequence of placing the objects [Hillier and Hanson, 1984, Hillier, 1996]. Once again, this is an example of the emergence of a global structure from the application of local rules.

2.1.3 Cellular Automata (CA)

When simulating nature, are mathematical equations the best way? In the early 1980s, Stephen Wolfram suggested that computer programs may be used to build models to study nature. Wolfram made a detailed study of a class of such models known as Cellular Automata (CA) and discovered that even when the underlying rules are very simple, the behaviour produced by computer models can be highly complex and mimic many features seen in nature [Wolfram, 1983, Wolfram, 1994]. The results of Wolfram's work are found in many applications, such as artificial life, biology, cryptography and fluid dynamics. The CA consist of a regular uniform lattice and a state for each cell or site (on/off for example). The CA evolves in discrete time steps, which means the value of a variable of one site is affected by states of its neighbourhood at the previous time step. The variables at each site are updated simultaneously according to a definite set of local rules.

2.2 Modelling biological systems using IBM and CA approaches

In this section, we review uses of the IBM and CA approaches. We restrict our review to the models simulating cell-cell interactions. Drasdo [Drasdo et al., 1995] suggested an approach which describes cells as particles in a strongly viscous fluid for growing tissue cultures and non-vascular tumors. The model assumes that each

cell is spherical in shape directly after division and deforms to a "dumb-bell" shape during mitosis. The model assigns a total energy for an assembly of cells according to the interactions between the cells resulting from attractive interactions due to cell adhesion and repulsive contributions from the ability of cells to deform and compress. The model has been extended to describe blastula formation [Drasdo, 2000] and the cell kinetics in intestinal crypts [Drasdo and Loeffler, 2001].

Dubertret suggested a model based on topology and statistical mechanics for the renewal of epidermis [Dubertret and Rivier, 1997]. In this model, each basal cell is attached to the basement membrane through a polygonal facet and the model assumes that the basement membrane to be flat; the tissue is in equilibrium and the horizontal pressure on the basal layer is isotropic. The model predicts that cells with a smaller attachment than average are likely to detach and ascend in the epidermis. The larger the attachment of a basal cell to the basement membrane, the more likely it is to divide. Moreover, the model suggests that a cell surface tension and adhesion play a paramount role in the renewal of the epidermis.

Clem suggested a 3D simulation of the renewal of the epidermis based on the hypothesis that consecutive transformation through stages is mainly due to increased cell growth [Clem and Rigaut, 1995,Clem et al., 1997]. A cell in the model is described by position in the tissue, age, phase in the cell cycle and generation (number of mitoses since birth). The model consists of stem and differentiated cells and only stem cells are allowed to divide. The first step of their simulation consists of a static representation of nasal epithelium based on positions, sizes, shapes and orientations of the nuclei of the cells of the tissue identified in 2D histological sections [Clem et al., 1992]. Using the starting values of cell proliferation parameters, they simulated the tissue growth process, and reproduced its evolution from normality to hyperplasia[3].

CA approaches on regular lattices have been used to model cell kinetics in intestinal crypts [Loeffler et al., 1986]. Crypts are pear shaped, single cell layer thick pockets

[3]An increase in the number of cells.

in the intestinal wall, which are responsible for maintenance of high-cell turnover in the intestine [Potten and Loeffler, 1990]. The majority of cell division takes place in the lower half of the crypts, cells then migrate towards the top of the crypt before being released into the intestinal lumen. In the model, the individual crypt is represented by a cylindrical array of cells with a periodic boundary condition in the horizontal direction. Cell migration is controlled by cell division, which is governed by two rules: cell division direction is chosen randomly; and the oldest neighbour cell is shifted towards the top of the crypt to make way for the newly divided cell. Cells that reach the top of the crypt are assumed to be released into the lumen and removed from the model. Meineke presented a model for intestinal crypt simulation using Voronoi tessellation [Meineke et al., 2001].

2.3 Evaluating in silico model outputs using histological images

In silico models have to be validated before their results could be used with any confidence. Traditionally, researchers have validated their models using data from biological experiments, either their own or those reported in the literature. We suggest that in some cases *in silico* models may be evaluated using histological images. Later, we demonstrate this using histological sections of oral epithelium. Oral epithelial structures vary greatly from one image to the next. However, the main characteristics include the basal cell distribution within the epithelium, the curved nature of the rete pegs, the depth of the layer of keratin and the population of lymphocytes [Speight et al., 1996]. In order to familiarise the reader with the histological terms used throughout this thesis, we provide a detailed discussion of the characteristics of epithelium as shown in Figure 2.1. The pink region is the epithelial tissue that lies at the surface of the mouth. A thin mucus-type liquid covers the top of the epithelium, which is called the keratin or superficial layer. The dark cells at the base

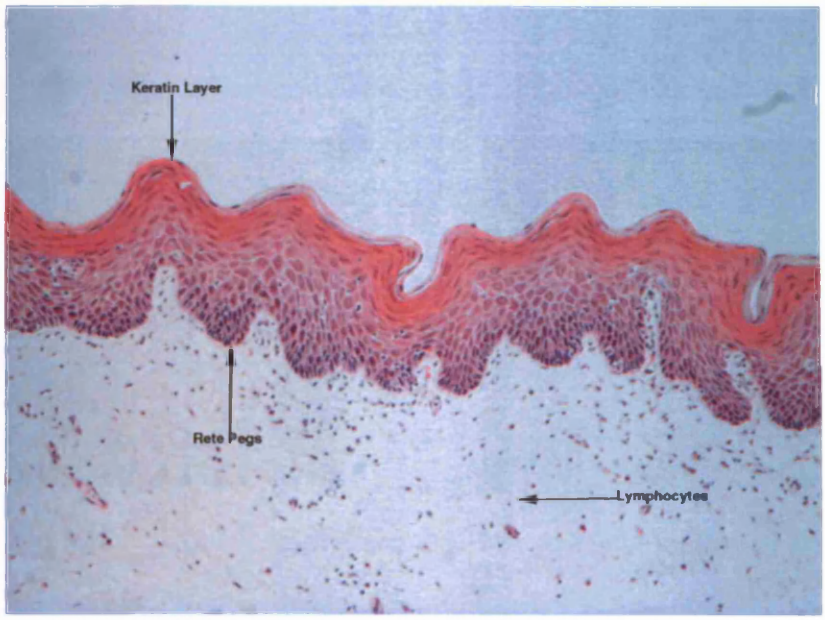
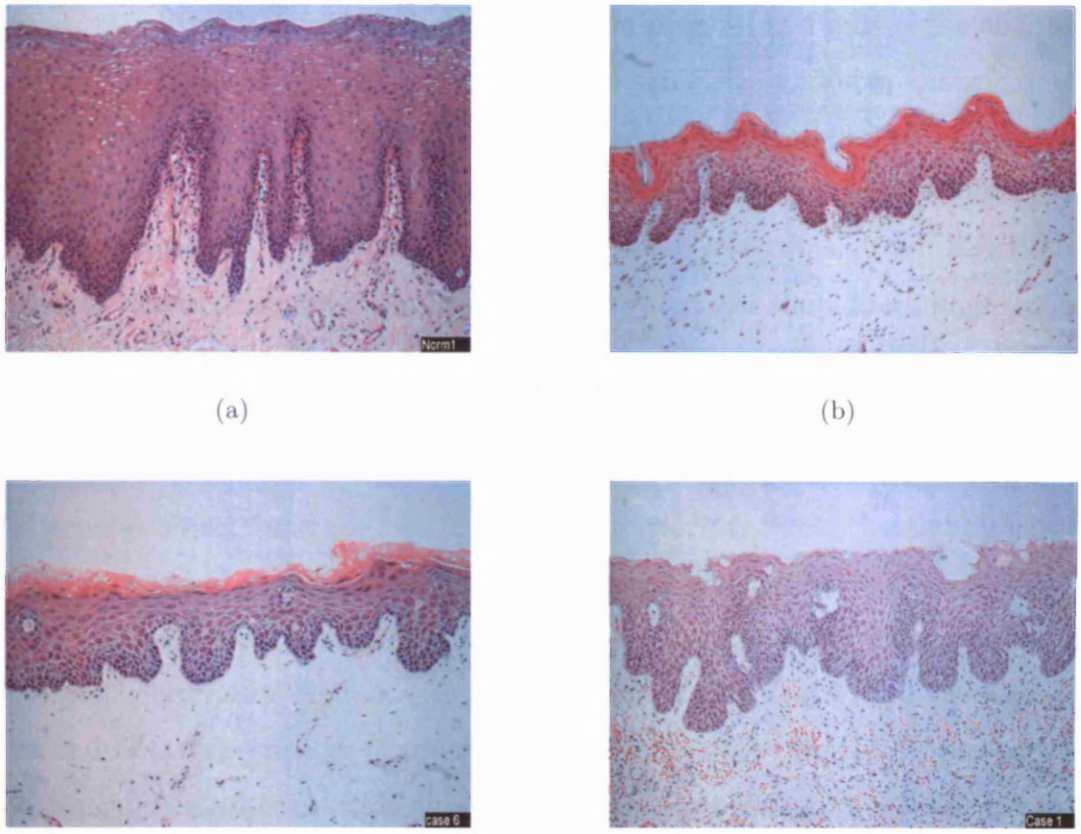


Figure 2.1: A histological section of oral epithelium

of the epithelium are the basal cells; in sections of healthy epithelium these are found at the base, but as tissue becomes dysplastic, such cells are found in the upper layers. The lymphocytes below the epithelial tissue are the small black spots in the image. The drop shaped rete pegs form the base of the epithelium. They become bulbous and highly curved in shape when an epithelium becomes dysplastic. Figure 2.2 shows histological sections of oral epithelium at different stages of dysplasia.

Oral pathologists have developed systems to diagnose oral carcinoma [Landini and Rippin, 1996, Aranjo, 1999] using descriptions of retepegs shape, distribution of basal cells and lymphocytes. One motivation of our work is to evaluate *in silico* models using histological sections of oral epithelium and build a shape model to represent the shape variation in normal and dysplastic rete pegs (basement membrane). The basement membranes can be represented as a collection of pixels in an image. In order to differentiate between the pixels, we need a way of describing them with a set of numbers called the object's *descriptors*. Previously unidentified objects can then



(a) (b)

(c) (d)

Figure 2.2: (a) Normal epithelium - The most distinctive feature of normal healthy epithelial tissue is a high concentration of basal cells lining the rete pegs and very straight rete pegs with few curved sections. (b) Mild Dysplasia - In comparison to Figure 2.2(a), the basement membrane is bulbous. Also the distribution of basal cells is far less concentrated in comparison to normal. (c) Moderate Dysplasia - In this case, the rete pegs are more rounded in shape and are curved. The distribution of basal cells covers more of the epithelium. (d)Severe dysplasia - Here, the rete pegs exhibit pronounced curvature.

be classified by matching the descriptors of the unknown objects against the known objects. However, in order to be useful descriptors they should have the following four properties.

First, two objects should have the same descriptors, if and only if the shapes are the same. Secondly if the two shapes have the same size and shape, their descriptors should be similar. Thirdly, they should have invariant properties. For example, rotation invariant descriptors would be useful for recognising objects, whatever their orientation. Other invariant properties include scale and translation. Fourthly, the descriptors should be compact. In other words, a descriptor should represent an object in an efficient way. The quantity of information used to describe this characterisation should be less than the information necessary to have the complete description of the object itself. However, there is no set of complete and compact descriptors to characterise general objects and the best recognition performance is obtained by carefully selecting a descriptor related to a particular type of application.

The descriptors may be divided into two types: *Region* and *Shape* descriptors. The *Region* descriptors characterise an arrangement of pixels within the area of an image and *Shape* descriptors characterise the arrangement of pixels in the perimeter or boundary. It is clear that for our application we would require a shape descriptor to describe the shape of the basement membrane. Fourier descriptors are a form of *Shape* descriptors, which allows the application of Fourier theory to shape description. The main idea of Fourier descriptors is to characterise a curve/contour as a set of numbers that represent the frequency content of the whole shape. Fourier based shape description has been around for some time [Persoon and Fu, 1977, Chellappa and Bagdazian, 1984] and has application in biology [Lestrel, 1997] and hand printed character recognition [Granlund, 1972].

2.4 The Morphology of Epithelium

Stratified squamous epithelium may be divided in to two types: *non-keratinizing* and *keratinizing*. The *non-keratinizing* normally refers to wet surfaces which are subjected to wear and tear. The *keratin* is formed when the squamous epithelium covers a dry surface, such as the skin. The structural organisation of both types of epithelium are similar and are usually divided into several zones: the *basal layer*; the *prickle* or *spinous layer*, also called the *Malpighian layer*; the *granular layer*; and finally the *keratin* layer. Figure 2.3 provides a diagrammatic representation of the epithelium.

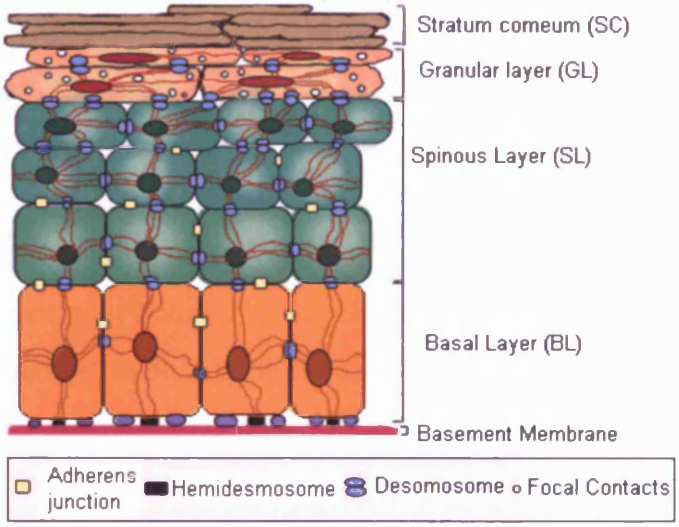


Figure 2.3: The epidermis is a stratified squamous epithelia, which consists of several layers. Resting on the basement membrane is the basal layer(BL) consisting of proliferating Stem and Transit amplifying cells. The basal layer give rise to differentiated cell layers of the spinous layer, granular layer and the stratum corneum. This image and text is reproduced from [Fuchs and Raghavan, 2002]

The basal cells form the proliferative compartment of epithelium, from which cells migrate, differentiate as they progress and eventually get desquamated at the surface. Most cells within the basal layer are the rapidly dividing progeny of stem

cells, referred to as transit amplifying cells [Jones and Watt, 1993, Jones and Watt, 1995]. The transit amplifying cells undergo a limited number of divisions before they withdraw from the cell cycle and commit to differentiation and begin to move towards the surface of the skin [Watt, 1998, Watt, 2001]. The epidermis is in a constant state of dynamic equilibrium, where it replenishes itself every two weeks [Fuchs and Raghavan, 2002].

The cells then migrate into the *Malpighian layer*, the thickness of which can vary a great deal. The cells in the *Malpighian* layers are attached to each neighbour by cytoplasmic processes called *desmosomes*, and the attachment between cells and the basement membrane are called *hemidesmosomes*. Since epithelium consists of a self renewing population of epithelial cells, these contacts between cells and membranes are constantly broken down and reformed as cells migrate. The *granular layer* is the transitional zone between the viable *keratinocytes* and the dead cells on the surface. The *keratin* layer is formed of flat keratinized plates.

Intercellular adhesion and its function on the development of tissue structures has been known for some time [Wilson, 1907]. Gumbiner [Gumbiner, 1996] suggested that the cell-cell adhesion system should be thought of as a mechanism that helps basic genetic information into the complex patterns of cells in tissues. Cell adhesion mediated morphogenesis is observed in several biological systems; for example Drosophilia wing development [McNeill, 2000, Dahmann and Basler, 1999], and the apical ectodermal ridge in the developing vertebrate limb [Kimmel, 2000]. In order for the epithelium to function as a tissue, epithelial cells must have the right shape and structure to pack together with their neighbours. This cellular architecture is maintained by cell-cell and cell-matrix adhesion and is important in its barrier and protective functions. This order is perturbed in several human genetic disorders including degenerative blistering diseases as well as carcinomas.

In epithelium, the adhesion is maintained by two types of adhesion junctions: adherens junctions (AJs) and desmosomes, which together constitute the intercellular

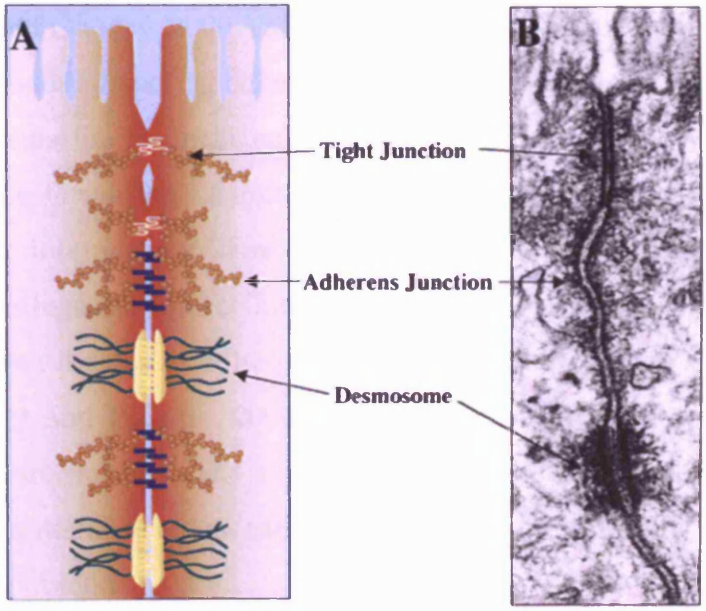


Figure 2.4: (A) Diagram of the three major types of intercellular junctions in epithelial cells. Tight junctions are composed of transmembrane proteins linked to the actin cytoskeleton and constitute a physical barrier between the apical and basolateral regions of the cells. (B) Electron micrograph depicting the ultrastructure of adherens junctions, desmosomes and tight junctions between two intestinal epithelial cells. The description and diagram is taken from [Perez-Moreno et al., 2003]

.

adhesion complex as shown in Figure 2.4. The transmembrane receptors[4] in the complex mediate the binding at the extracellular surface and determine the intracellular response. The associated cytoplasmic proteins of the transmembrane receptors structurally link them to the cytoskeleton [5], thereby establishing lines of communication to other cell-cell junctions and to cell-matrix junctions. This linkage of cell-cell and cell-matrix junctions allows the formation of epithelial tissue from individual epithelial cells [Green and Gaudry, 2000, Vasioukhin et al., 2000, Runswick et al., 2001]. The AJs' functions include: directing cellular organisation, movement within epithelia and transmitting information from the environment to the interior of cells [Tsukita et al., 2001]. During the formation of epithelial sheets, epithelial cells constantly change their intercellular interactions. For example, during wound healing epithelial cells momentarily downregulate intercellular adhesion and increase cell proliferation. The association between defects in adherens junctions proteins and human cancer is well understood [Pelfer and Polakis, 2000]. The AJs are cadherin[6]-dependent adhesive structures which are linked to the actin microfilament network within the cytoskeleton. E-cadherin, a member of the cadherin family, exists primarily in epithelia at the sites of cell-cell contacts.

The adheren junctions are also integrated into a variety of other cellular processes through associations with other types of intercellular junctions and membrane receptors. Epithelial sheets display AJs alternating with desmosomes [Vasioukhin et al., 2000, Green and Gaudry, 2000] as shown in Figure 2.4. Studies have revealed that the establishment of AJs is a prerequisite for the formation of desmosomes and other

[4]Transmembrane receptors are integral membrane proteins, which reside and operate typically within a cell's plasma membrane, but also in the membranes of some subcellular compartments and organelles. Binding to a signalling molecule or sometimes to a pair of such molecules on one side of the membrane, transmembrane receptors initiate a response on the other side. In this way they play a unique and important role in cellular communications and signal transduction. This description was taken from wikipedia.org

[5]The cytoskeleton is both a muscle and a skeleton, and is responsible for cell movement, and the organization of the organelles within the cell.

[6]A homophilic cell adhesion protein, which binds to other cadherin molecules

junctions [Gumbiner et al., 1988]. The desmosomes are specialized cadherin mediated cell-cell junctions which attach to the intermediate filament network providing internal mechanical strength to epithelial cells [Fuchs and Cleveland, 1998].

During embryogenesis, boundaries usually develop between cell populations. It has long been suggested that differential cell affinities play a role in orchestrating the formation of tissue boundaries and cadherin plays a cental role in this process. There are now more than 20 cadherin proteins, which are differentially expressed in complex patterns. Takeichi et al [Nose et al., 1988] transfected two of the cadherins, E- and P-cadherin into a separate group of cells which normally express little or no cadherin. The transfected cells preferentially adhered to cells expressing the same cadherin proteins and developed epithelial sheets, whereas untransfected cells did not express any cadherin and did not form epithelial sheets.

2.5 Cell sorting and computer models

Steinberg and colleagues performed a series of experiments to study various cell mechanisms suggested by Holtfretter and others [Curtis, 1960,Curtis, 1961], which underly tissue affinities. Steinberg suggested the Differential Adhesion Hypothesis(DAH) as the driving mechanism behind cell sorting as observed in embryonic development [Steinberg, 1963, Steinberg and Takeichi, 1994, Steinberg, 1996, Steinberg and Foty, 1997]. The DAH postulates that cell surface adhesive properties cause tissues whose cells are mobile to behave as living liquids. For a system to show behaviour postulated by DAH, they must be composed of many subunits which cohere while being motile. In liquids, the subunits are molecules and the mobility is thermal; in cell populations, the subunits are living cells and their mobility may be active or passive. The syndrome of behaviors displayed by embryonic cell populations which duplicate the behaviors of immiscible liquids include: irregularly shaped tissue fragments with shapes toward a sphere, the spreading of one tissue mass over the surface of another [Steinberg, 1996].

Glazier and Graner suggested a cell sorting algorithm to model DAH using a cellular automaton approach [Wolfram, 2002]. The algorithm simulated the sorting of a mixture of two types of biological cells using a modified version of the large-Q Potts model with differential adhesion [Graner and Glazier, 1992, Graner, 1993, Graner and Sawada, 1993, Ball, 2000]. Savill [Savill and Sherratt, 2003] extended the Glazier-Graner model to study the control of stem cell clusters by Notch-mediated lateral induction. The simulations were based on the Glazier-Graner [Graner and Glazier, 1992] framework. The model included cell properties like age, size, type, adhesion strength to neighbouring cells and Delta and Notch concentration on its membrane. Savill [Savill and Sherratt, 2003] concluded that the regulation of differentiation is the most likely mechanism for cluster control in epithelium and stem cells must adhere more strongly to each other than they do to differentiated cells. Developing the model further, he showed that lateral-induction mediated by the Notch signalling pathway is a natural mechanism for cluster control. Savill's work focussed on a single stem cell cluster. The ability of the basal layer of the epidermis to produce a single stem cell giving rise to multiple stable clusters is not modelled [Jones and Watt, 1995].

Savill [Savill and Hogeweg, 1997] presented a 3D hybrid cellular automata with a partial differential equation model to study the morphogenesis in simple cellular systems. He applied the model to the cellular slime mould *Dictyostelium discoideum,* "from single cells to crawling slug" and demonstrated that using local interactions, basic morphogenesis can be obtained with three processes: production of and chemotaxis to *cAMP*, and cellular adhesion. The cellular adhesion is modelled using the Glazier-Graner model and the amoebae are represented as many connected automata instead of a point like object. Several extensions and modifications were made to the Glazier-Graner model: three dimensions, coupling to a partial differential equation and chemotaxis along a positive spatial *cAMP* gradient.

Izaguirre [Izaguirre et al., 2004] suggested a multi-model software framework called *COMPUCELL* for simulation of morphogenesis by extending the Glazier-Graner model.

It models the interaction of cellular mechanisms such as cell adhesion, division, haptotaxis [7] and chemotaxis [8]. A combination of stochastic local rules and a set of differential equations and partial differential equations model cellular processes. Izaguirre used *COMPUCELL* to simulate the formation of the skeletal architecture in the avian limb bud.

Stott [Stott et al., 1999] suggested a simulation of benign avascular tumour growth using the Glazier-Graner model. He simulates tumour growth including cell sorting, the latter occurring due to variation of adhesion between different types. The simulated tumor grew exponentially, then formed three concentric shells as the nutrient supplied to the core by diffusion decreases: the outer shell consisting of proliferating cells, the middle of quiescent cells and the centre a necrotic core where there is insufficient nutrient to sustain life. The simulated tumour eventually approaches a steady state, where the increase in growth is equal to the loss due to the disintegration of the cells in the necrotic core.

Turner [Turner and Sherratt, 2002] developed a discrete model of malignant invasion by extending the Glazier-Graner model. The model simulates a population of malignant cells experiencing interactions due to intracellular adhesion while secreting enzymes and experiencing a haptotactic gradient. The model investigates the influences of changes in cell-cell adhesion on the invasion processes and demonstrates that the morphology of the invading front is influenced by changes in the adhesive strength between cells. Turner also demonstrates with the model that cell-cell adhesion is less of an influence compared with cell-medium adhesion. The simulation is then extended by including cell proliferation. Turner concludes that increased proliferation rate does not necessarily result in an increased depth of invasion into the extracellular matrix, and may indeed cause the invasiveness to be reduced.

Mombach [Mombach et al., 1993], [Mombach et al., 1995], [Mombach and Glazier,

[7] A directed response of cells in a gradient of adhesion.

[8] The movement of a cell or organism in response to a chemical gradient.

1996] studied cell sorting using the Glazier-Graner model. He presented a three-dimensional lattice simulation of embryonic cell movement driven by differential adhesion mechanisms [Mombach, 1999]. The simulation is evolved by a Metropolis algorithm [9] and he found that variation on the relative concentration of the cell types in an aggregate affects the dynamics of sorting.

[9]The Metropolis algorithm was invented by Metropolis et al in 1953 for sampling an arbitrary probability distribution. This algorithm has been very successful and influential in many areas of Monte Carlo method.

Chapter 3

Modelling Epithelial Cell Interactions

3.1 Introduction

Our aim is to construct an *in silico* model of epithelial cell-cell interactions. Once validated, the model can be perturbed to simulate and hence enhance our understanding of the role of cellular interactions in disease processes. Any such model would have to be compared with *in vivo* or *in vitro* experiments to determine its validity. The *in silico* models can then simulate novel biological experiments on a computer. *In vitro* biological experiments are used to inform and enhance simulation models. Conversely, *in silico* models could simulate *in vitro* biological experiments before expensive and time consuming laboratory research is undertaken. Here we report on our attempts to extend Stekel's [Stekel et al., 1995] rule based model of the squamous epithelium.

Stekel [Stekel et al., 1995] proposed a rule based model to simulate a multilayered sheet of epithelial cells. The model consists of several cell layers, where the deepest undulating layer is known as the basal layer. The basal layer is separated from the underlying connective tissue matrix by the basement membrane. The model simulated biological processes such as cell division, differentiation, death and adhesion. The kinetics of the simulation were modelled through pressure arising from overlapping cells. Stekel initialised the model in a variety of ways, e.g with a single cell and

a line of cells and showed that the model can reach a steady state in terms of cell population distribution. The model employed five cell types: Stem, Basal, Intermediate, Mature, and Dead. Only Stem and Basal cells on the basal layer are capable of dividing in the model. Intermediate, Mature and Dead cells are produced through cell differentiation. Chemical factors which affect the cell differentiation of individual cell types are produced by cells within the model. We produced a rational reconstruction of Stekel's model in order to gain experience of agent based modelling. Collision detection algorithms were used to improve computational efficiency. We extend the model by suggesting a method to analyse spatial ordering of cells from the model output using techniques from computational geometry and image processing.

3.2 The Stekel's 2D Model For Squamous Epithelium

Stekel suggested a model for squamous epithelium by suggesting a rule based simulation. The cell shape was modelled as a circle of fixed radius. All types of cells are assumed to be of the same size. Cell movement results from physical interaction with other cells. The basement membrane consists of a continuous line of cells forming a lower boundary of the simulated epithelium. On division, each stem cell on the basement membrane gives rise to a stem and a basal cell. Stekel assumed the existence of an interaction between stem and basal cells and the underlying connective tissue in terms of a downward force which acts on Stem and Basal cells only. This force is known as the Connective Tissue Attractive Force(CTAF). Another force in the model, known as the Connective Tissue Repulsive Force(CTRF), acts on Intermediate, Mature and Dead cells.

The connective tissue repulsive force causes all cells apart from Stem and Basal cells to move away from the basement membrane. It is likely these assumed forces give rise to the formation of retepegs in the model, although there is no evidence

in the literature to suggest that such forces exist in real biology. Basal cells on the basement membrane are also capable of division giving rise to two basal cells. When a basal cell is sufficiently distant from the basement membrane, it differentiates into an intermediate cell. Intermediate cells are not capable of dividing. However, if at any stage an intermediate cell comes into contact with the basement membrane, the cell differentiates back into a basal cell and thereafter is capable of division.

3.2.1 Cell Division

Cell division is modelled as a stochastic process. The parent cell retains its original position. The daughter cell is placed such that the distance between the center of the parent and daughter cell is two cell radii. The cells are programmed to divide "upwards" away from the basement membrane at an angle between 0 and 90 degrees in a random fashion as shown in Figure 3.1. Cells divide randomly to the right and left of the perpendicular to the horizontal line through the centre of the cell about to divide. However, this is biased so that cells divide away from the nearest stem cell or stem cell clusters [Stekel et al., 1995]. This bias is computed according to the stem cell factor (see section 3.2.2) perceived by the cell.

3.2.2 Chemical Factors

Cells are assumed to secrete chemical factors in the model. The chemical factor produced by a cell is distributed throughout the model on the basis of an inverse square law. Cells respond to the various chemical factors either as a threshold or on a linear basis. Stem cells secrete chemicals, which is referred to in Stekel's model as the stem cell factor. If the Stem cell factor at a point falls below the threshold level, the basal cells differentiate and become Stem cells. The chemical factors produced by Stem and Basal cells act on Intermediate cells to prevent their differentiation to mature cells and is known as the Stem and Basal cell factor. All chemical factors are computed at a point, ie: from the center of a target cell. The Stem and Basal factor

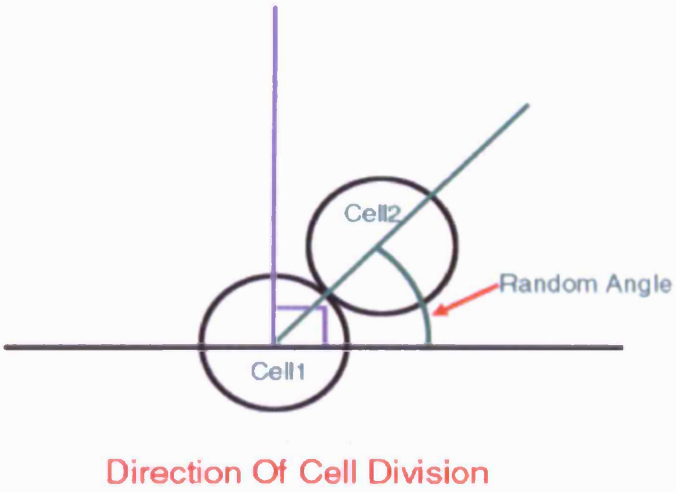


Figure 3.1: Direction of cell division as proposed by Stekel [Stekel et al., 1995]. After division, the parent cell (cell 1) retains its position. The daughter cell (cell 2) is placed such that the distance between the center of the two cells is twice the cell radius.

decays with respect to distance and is computed as shown in Equation 3.1.

$$SBF(y) = \sum_{x \varepsilon S} \frac{1}{\partial(x,y)^2} + \frac{1}{10} \sum_{x \varepsilon B} \frac{1}{\partial(x,y)^2} \tag{3.1}$$

where S and B represent the sets of stem and basal cells respectively and $\partial(x, y)$ is the Euclidean distance between cells x and y.

3.2.3 Cell Differentiation

Cell differentiation is determined by age, position and neighbouring cells. However, Stekel gave no explanation of how neighbouring cells are defined. We define neighbours as cells within some predefined distance from the center of the target cell. Stem cells do not differentiate if they are in contact with the basement membrane. However, if a stem cell is positioned off the basement membrane and is surrounded by

intermediate cells it differentiates to become an intermediate cell. The surrounding intermediate cells are those within a predefined distance from the centre of a target cell.

Basal cells differentiate into Stem cells if the chemical signal perceived by the basal cell falls below a threshold defined for the Stem cell factor. If a basal cell is positioned away from the basement membrane it differentiates and becomes an intermediate cell. An intermediate cell differentiates to become a mature cell when the cell's age is greater than a pre-determined threshold. When an intermediate cell comes into contact with the basement membrane, it differentiates into a basal cell.

Mature cells that lie above the intermediate cell layer arise from the differentiation of intermediate cells. This process of differentiation from intermediate cell to mature cell also depends on cell age. Stem and Basal cells in the model do not age. Only differentiated cells age. The cell age is computed on the basis of biological iteration (see section 3.2.7). The age of a cell is one of the variables taken into consideration during cell maturation and cell death. Figure 3.2 shows possible differentiation paths available for cells in the Stekel model.

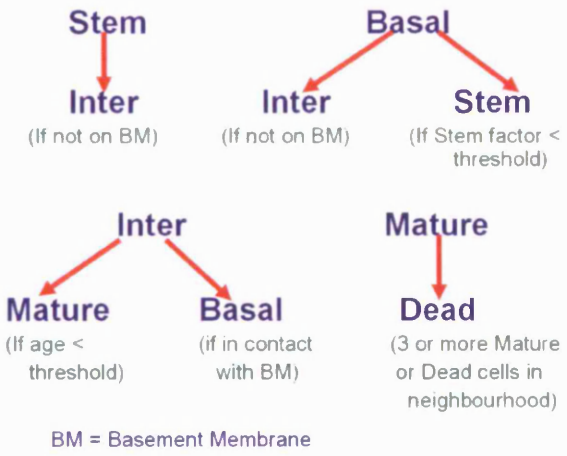


Figure 3.2: Possible cell differentiation paths in the Stekel model.

3.2.4 Cell Death

A mature cell becomes a dead cell when there are three or more neighbouring mature or dead cells [Stekel et al., 1995]. The desquamation factor, as it is known, is computed on the basis of an inverse square law and is modelled to be secreted by mature and dead cells. A dead cell is removed from the model when its age is greater than a specified threshold and the desquamation factor is below a predefined threshold value [Stekel et al., 1995].

3.2.5 Cell Movement

The model employs both physical and chemical forces. The physical force, also referred to as the compressive force, is proportional to the amount of overlapping area between cells. We use the formula suggested by Stekel ([Stekel et al., 1995], page 286) to compute the overlap. Cells in the model also experience the two chemical forces mentioned earlier. The connective tissue attractive force acts on Stem and Basal cells by moving them downwards towards the basement membrane. The connective tissue repulsive force acts on Intermediate, Mature and Dead cells by moving them upwards away from the basement membrane. Cells respond to this connective tissue repulsive force by moving upwards towards the top of the tissue being modelled. According to Stekel [Stekel et al., 1995], these forces are modelled to simulate the chemical factors secreted by the underlying connective tissue. The connective tissue attractive force and connective tissue repulsive force are assumed to be uniform throughout the model.

3.2.6 Defining the Basement Membrane

Stekel modelled basement membrane shape according to rules defining positions of stem and basal cells. A scoring system is used to decide if a cell is on the basement membrane. If a cell detects that another cell below is within 3 cell radii from the centre of the target cell, then 2 is added to its score. A score of 1 is added if the

cell below is between 3 and 5 cell radii from and at an angle less than $\frac{\pi}{4}$ from the vertical as shown in Figure 3.3. If a cell scores less than 6 then it assumed to be on the basement membrane [Stekel et al., 1995]. The basement membrane shape is then defined as the curve outlined by stem and basal cells.

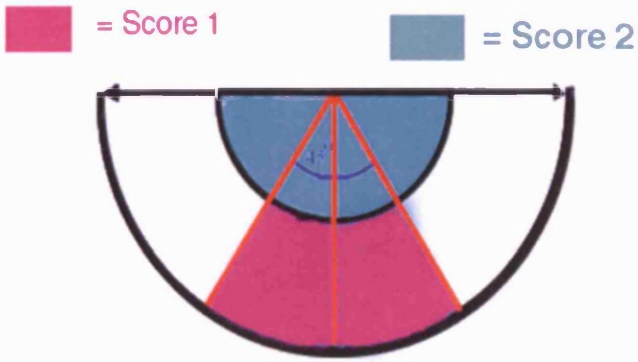


Figure 3.3: Stekel's algorithm for deciding on which cells are on the basement membrane. Cells directly below the target cell and within 3 cell radii score 2. Cells within 5 cell radii and at an angle $\frac{\pi}{4}$ from the vertical score 1. If the total score is greater than 6, the cell is assumed not to be on the basement membrane

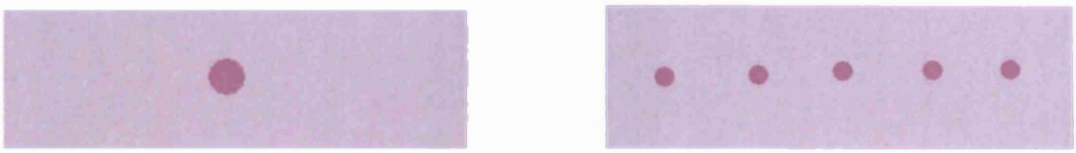
3.2.7 Rational reconstruction of the Stekel Model

We initialise our version of the Stekel model with a single stem cell (Figure 3.4(a)) and with multiple stem cells (Figure 3.4(b). In our implementation, we deal with biological processes in the following order.

- Compute which cells are on the basement membrane
- Compute Stem and Stem Basal chemical factors
- Compute cell differentiation

- Compute cell age, cell death and cell desquamation
- Check for cell mitosis

The above steps are referred to as a biological iteration. We then model the kinetics of the cells by computing the compressive and chemical forces on each cell and move them according to the forces acting on them.



(a) (b)

Figure 3.4: (a) Initialise the simulation with a single stem cell [Stekel et al., 1995]. (b) Initialise the simulation with multiple stem cells [Stekel et al., 1995]

Compressive Force	0.03
Stem Factor	0.005
Stem Basal Factor	0.05
Desquamation Factor	0.007
Inter to Mature age	15
Mature to Dead age	40
CTRF	3
CTAF	1
Stem probability	0.6
Basal probability	0.8

Table 3.1: List of input parameters used in the simulation

All results shown here were obtained using the input parameters shown in Table 3.1. The compressive force parameter defines the force between two cells due to overlap. The stem factor, inter to mature age and mature to dead age influences cell differentiation as explained in section 3.2.3. The stem and basal probabilities refer to

likelihood of cell division. Stekel [Stekel et al., 1995] did not publish or provide evidence for the input parameters used in the simulations. Initially we were motivated by producing outputs similar to Stekel's results. Our input parameters were obtained through trial and error and do not reflect in anyway biological measurements. The model is sensitive to the choice of input parameters.

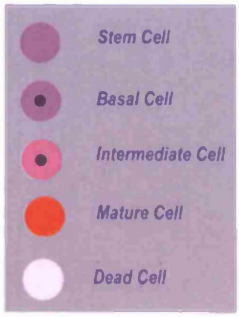


Figure 3.5: Colour key for different types of cells in the model

Example output data from our re-implementation of the Stekel model are shown in Figures 3.6, 3.7 and 3.9. In Figure 3.6, the model was initialised with a single stem cell and executed for 200 biological iterations. Like Stekel [Stekel et al., 1995], we also initialised our model with multiple stem cells and the results are shown in Figure 3.7. A movie named Stekel can be found in the CD attached and a description of the movie is provided in Appendix B. The movie is obtained by recording the simulation output over time.

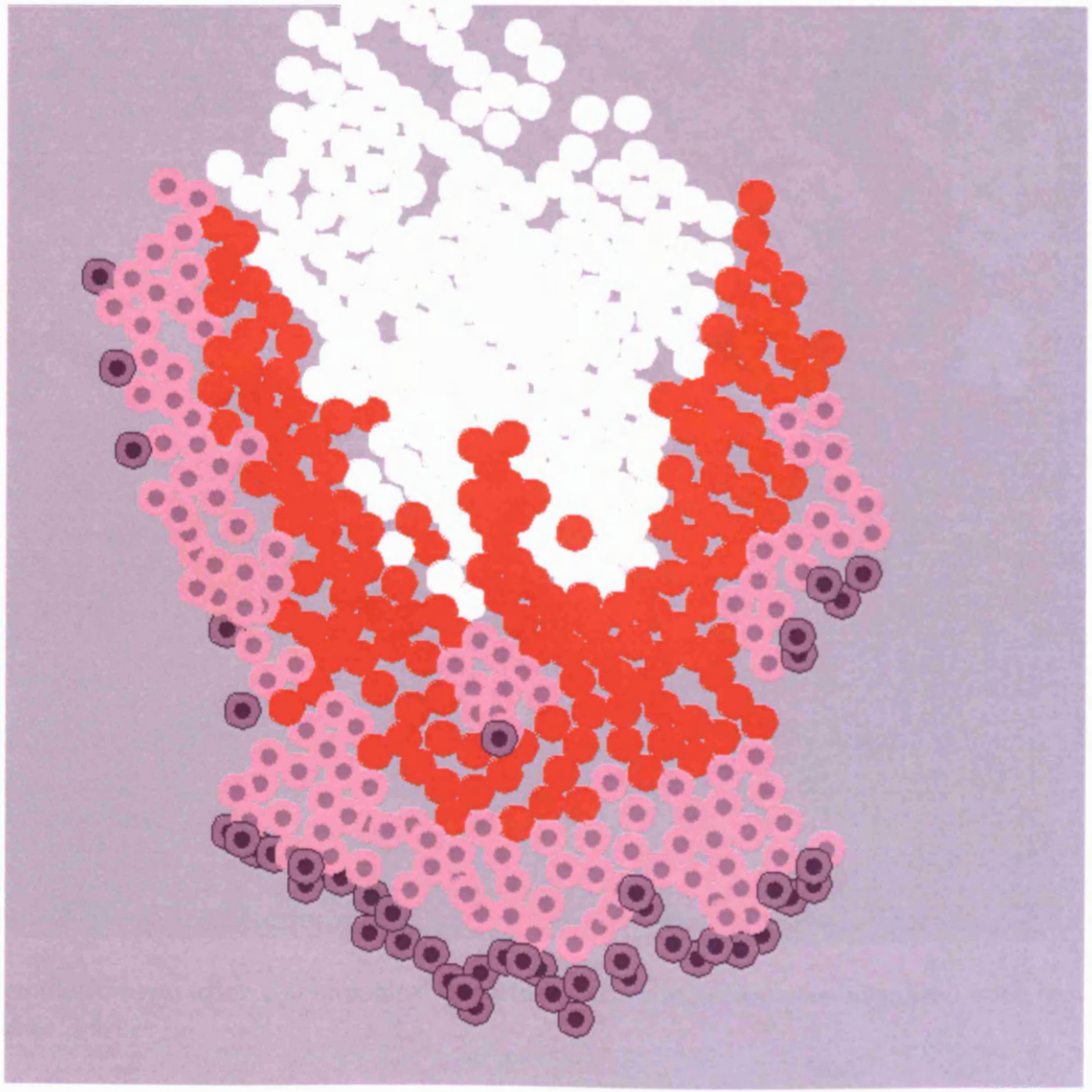


Figure 3.6: A simulation run after 200 biological iterations. The simulation was initialised with a single stem cell

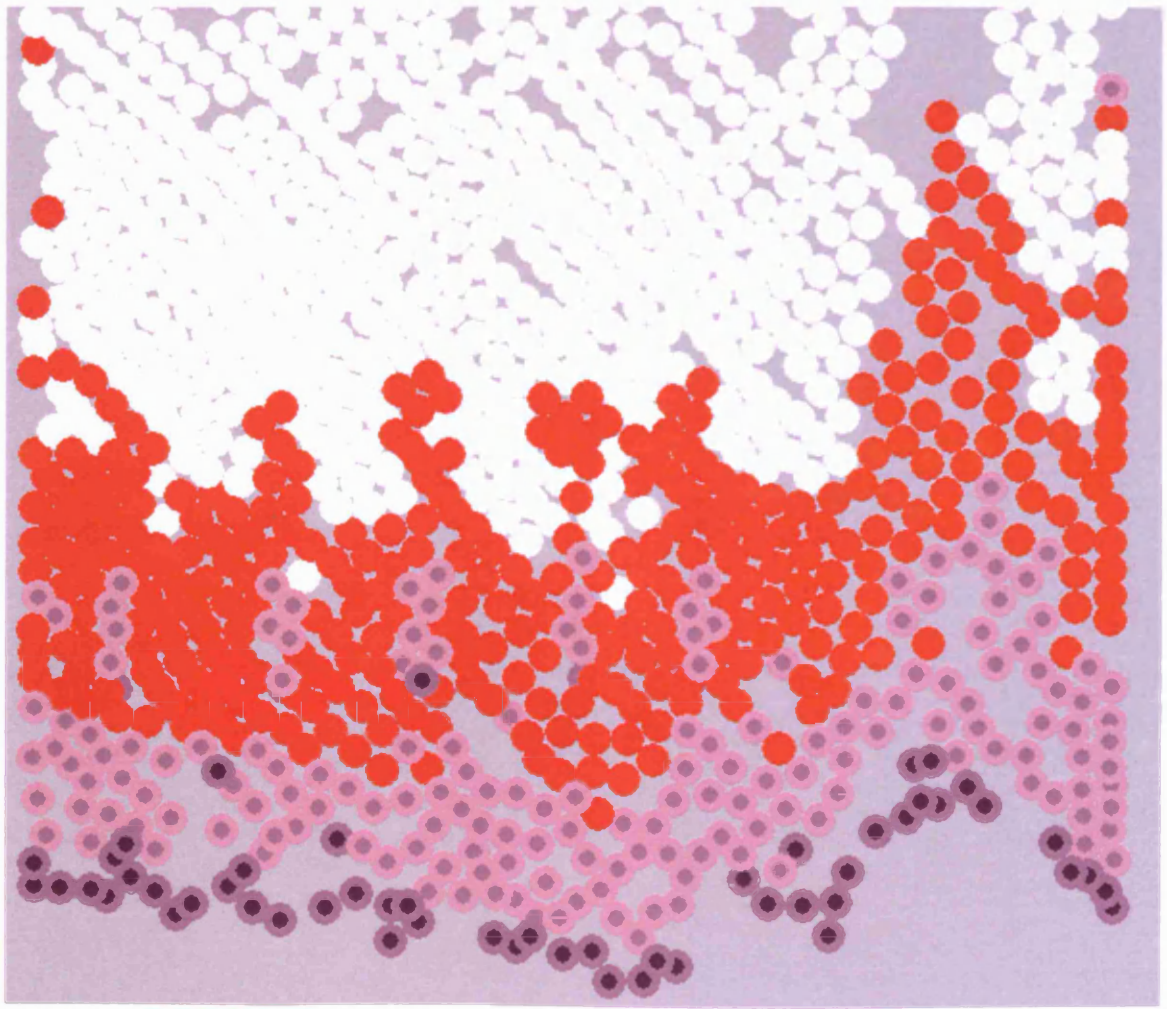
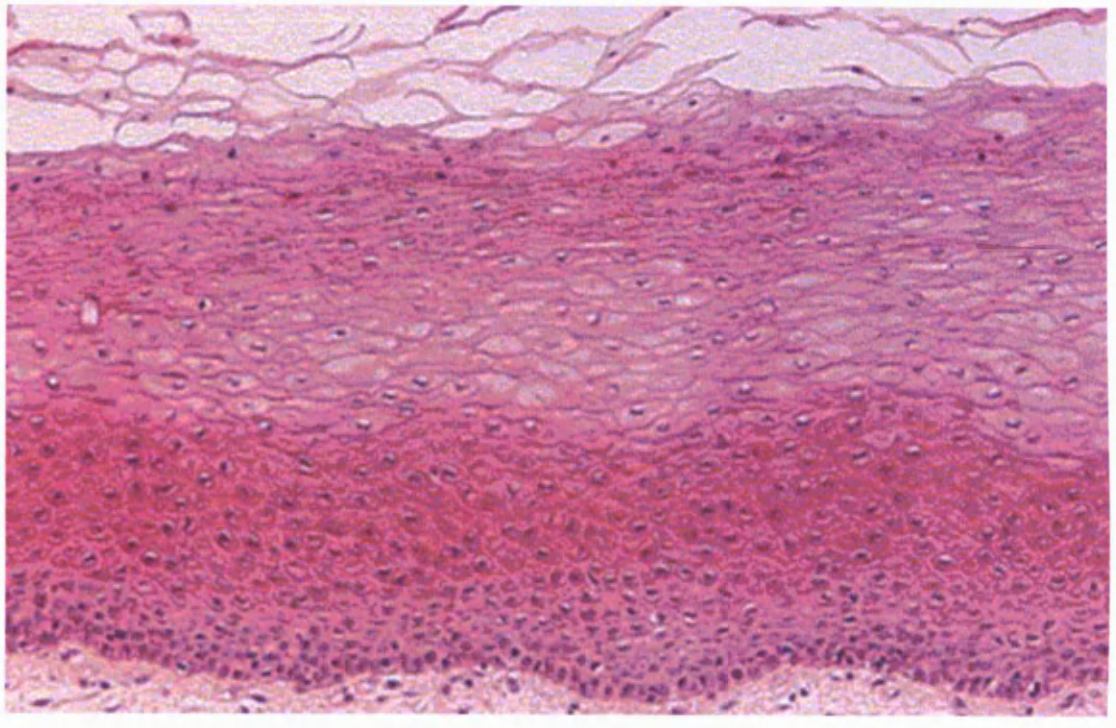
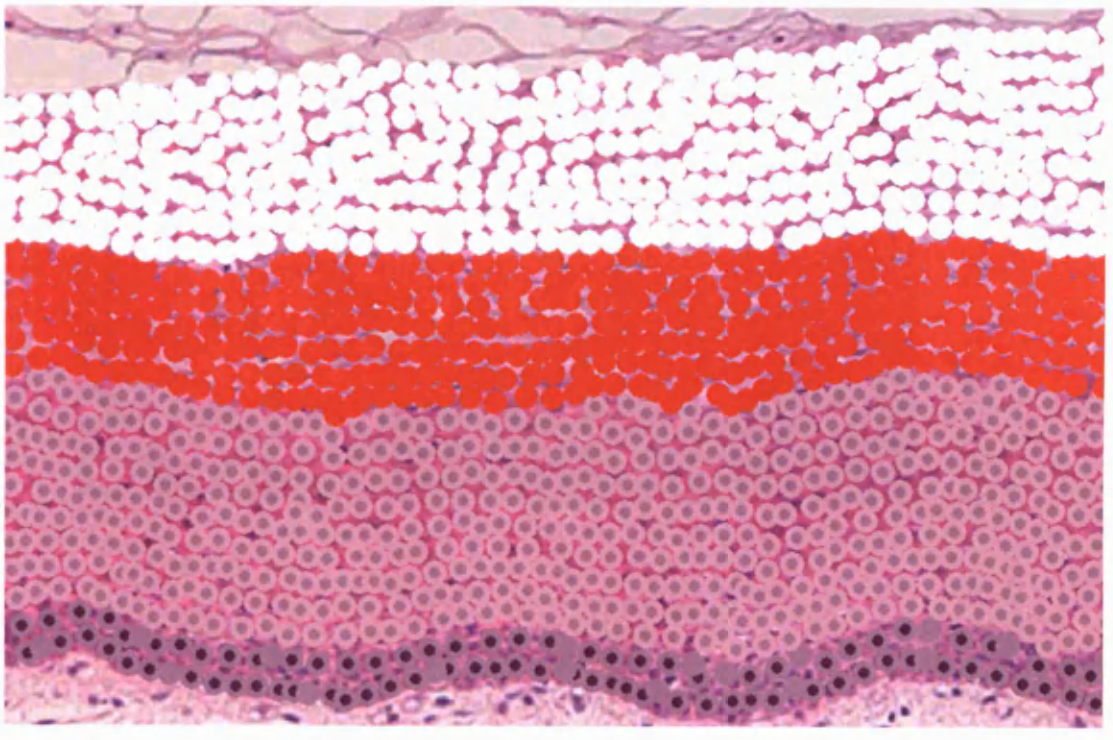


Figure 3.7: A simulation run after 200 biological iterations. The simulation was initialised with multiple stem cells as shown in Figure 3.4(b)

In one of our experiments as shown in Figure 3.8, we initialised the model with a large number of cells derived from an epithelial image. We used the buccal epithelial image in Figure 3.8(a) to derive initial cell positions and types, as shown in Figure 3.8(b). We then ran the simulation for 200 biological iterations and the result is shown in Figure 3.9.



(a)



(b)

Figure 3.8: (a) A histological image of buccal epithelium. (b)Cell positions and thickness of the layers are derived from the buccal epithelial image.

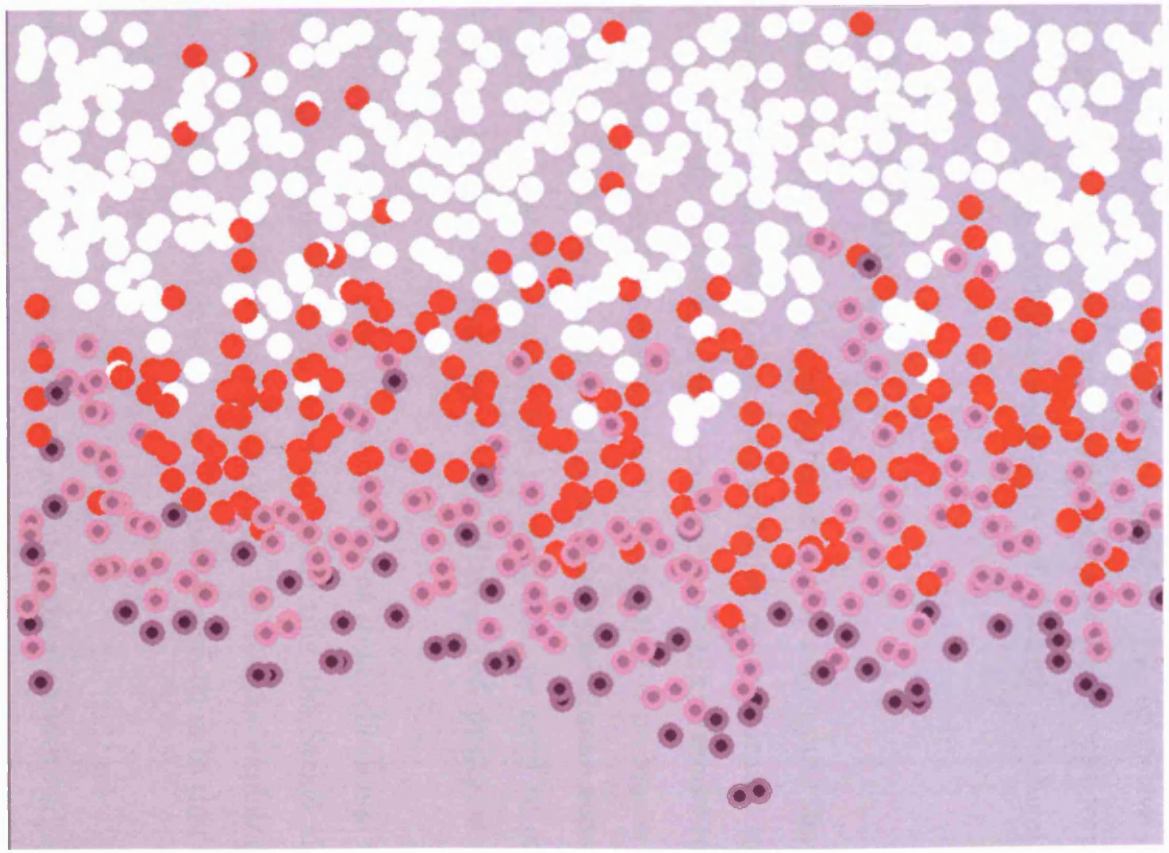


Figure 3.9: A simulation run after 200 biological iterations. The simulation was initialised with a histological image as shown in Figure 3.8(b). The simulation shows a lack of cohesion in the epithelial cell mass.

3.3 Criticism of Stekel's Model

After re-implementing Stekel's model [Stekel et al., 1995], we identified a number of inadequacies. The major problem is the lack of biological evidence to support the assumptions about the direction of cell division, the connective tissue attractive force and repulsive forces, and the determination of which cells form the basement membrane. We discuss criticisms related to cell division, chemical forces and cell shape here.

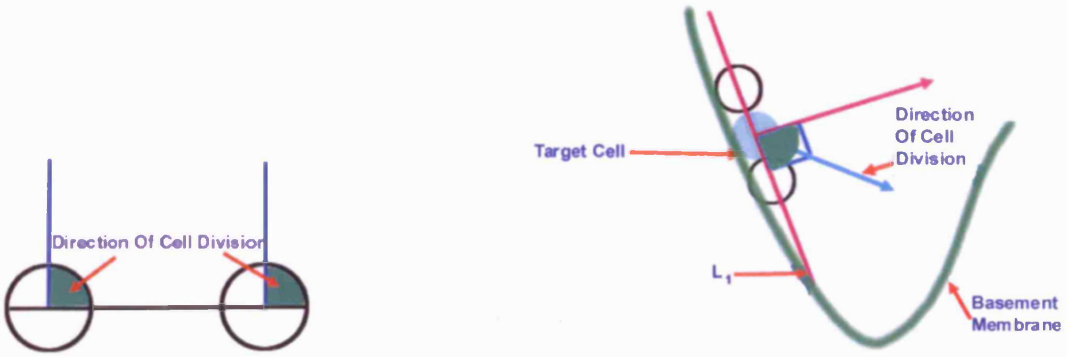
3.3.1 Cell Division

The algorithm provided by Stekel [Stekel et al., 1995] to compute the direction of cell division works when the model is started with two cells separated by some distance or with a line of cells. However as the number of cells in the model increases and the basement membrane becomes undulated, this algorithm is no longer sufficient. Cells are programmed to divide away from the basement membrane according to Stekel. However when the basement membrane is undulated, we need to compute the line from which the cell moves away by taking adjacent cell positions into account as shown in Figure 3.10.

Robert [Lavker and Sun, 1983] suggested that as stem cells divide they move along the basement membrane layer. In order to model this, the Stekel algorithm for cell division would have to be modified significantly. For reasons explained in section 3.5 we choose not to do this, but propose an alternative algorithm to that of Stekel [Stekel et al., 1995].

- Compute a line L_1 through the cell about to divide and its adjacent neighbours on either side
- Construct line perpendicular to L_1 through the centre of target cell as shown in Figure 3.10

- Compute a random angle between L_1 and its perpendicular.
- Place the new cell on tangential to its parent.



(a) (b)

Figure 3.10: (a) Direction of cell division when the basement membrane is flat. (b) Direction of cell division when the basement membrane becomes undulated

3.3.2 Chemical Forces

We have found no biological evidence to suggest the existence of the connective tissue attractive and repulsive forces introduced by Stekel. We also failed to find any biological explanation for the undulation of the basement membrane shape. Figure 3.11 shows an output from the simulation, where the Connective Tissue Attractive Force (CTAF) and Connective Tissue Repulsive Force (CTRF) are set to 0. As can be seen from Figure 3.11, the retepegs are not well formed in comparison to the output shown in Figure 3.7 and layering of different type of cells are not as uniform as that shown in Figure 3.7.

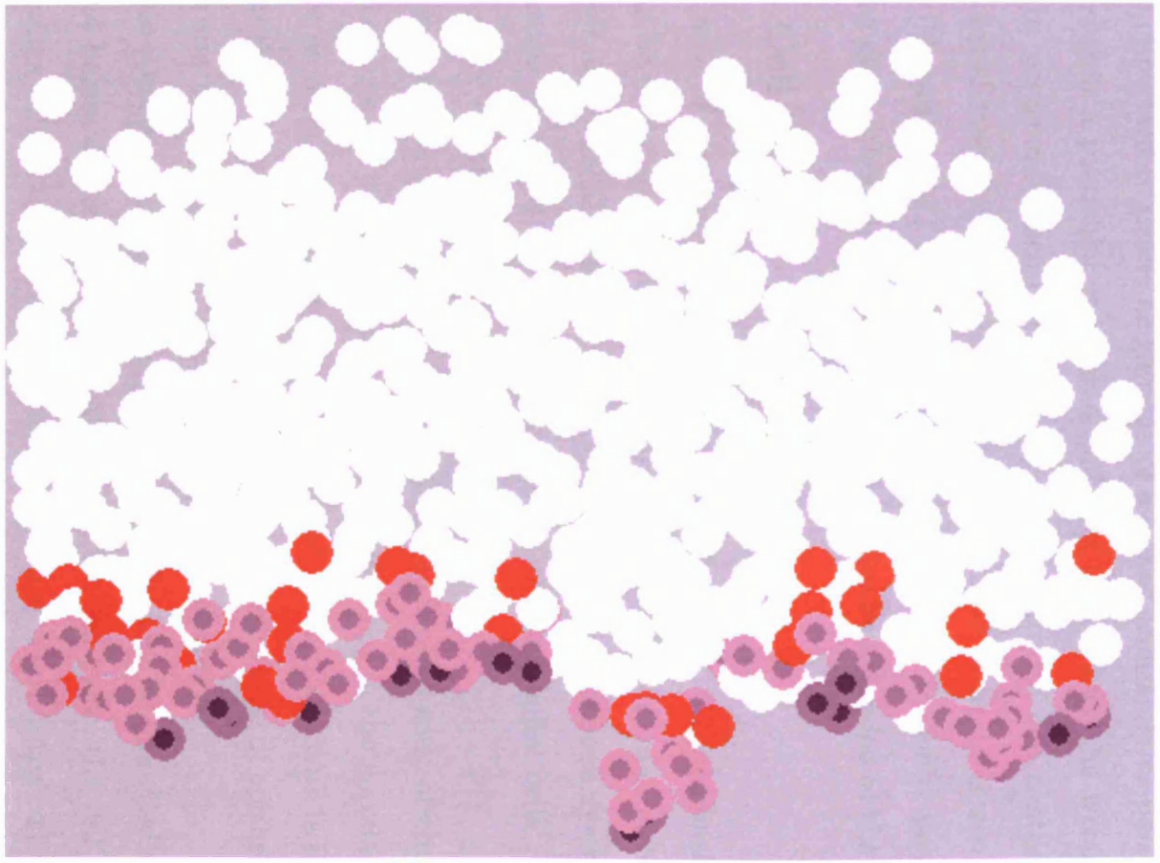


Figure 3.11: A simulation output, where the CTAF and CTRF are set to 0

3.3.3 Defining Neighbours

A healthy epithelial tissue is an ordered structure consisting of several layers of cells of different types. Epithelial signalling may be mainly due to cell-cell contacts. Therefore, it is important to define neighbours of a cell as accurately as possible. Defining all cells within a particular area as neighbours is probably the wrong way to approach the problem. Since cell signalling here is through cell contact we propose to use a tessellation technique [Okebe et al., 1992] in order to describe cell shape. Meineke [Meineke et al., 2001] used Voronoi tessellation to define neighbours when simulating gut epithelium. We believe this may also be extended to the Stekel model.

3.3.4 Cell Shape

Another area of concern is the modelling of cell shape. We demonstrate how cell shape can be modelled when using an individual based model approach. We filmed a mono layer of epithelial cells over 14 hours using a digital camera system from Leica Microsystems to record the motion sequence of oral epithelial cells[1]. Images were acquired every 10 minutes over the 14 hour period.

Although we were not able to follow an individual cell during filming and observe its shape change, we saw significant changes in cell shape in the population observed. Therefore modelling all cells as of equal size and of circular shape is far too simple. Epithelial tissue is an ordered structure of cells and the shape of epithelial cells plays an important part in how tightly they are packed together.

Drasdo [Drasdo et al., 1995] suggested that cells may be modelled as a dumbbell shape just before division. At the beginning of the cell cycle, the cell radius will be R_{min}. Radius will grow until the cell mass is doubled. During the cell division phase of the cell cycle, the distance d between the two centers of the symmetrical dumbbell shape is grown until $d = 2 \times R_{min}$. At the same time the radius R shrinks to keep the

[1]The Filming experiment was carried out at the Department of Orthodontics, Eastman Dental Institute with help from Dr Mark Lewis.

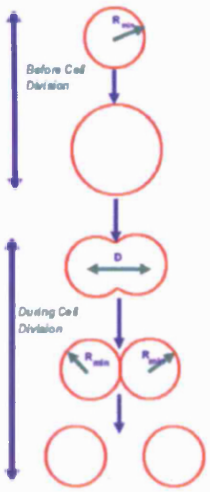


Figure 3.12: A diagrammatic representation of how cell shape could be modelled. Figure is reproduced from Drasdo et al [Drasdo et al., 1995]

surface area constant. Figure 3.12 is reproduced from Drasdo [Drasdo et al., 1995] and provides a pictorial representation of how cell shape could be modelled.

Schaller [Schaller et al., 2003] presented a solution to modelling cell shape using weighted Delaunay triangulation for modelling cell shape in tumour growth. Honda et el. [Honda et al., 1996] proposed an explanation of the topological organisation of corneal epithelium based on the assumption that cells migrate upwards to occupy less crowded regions in tissue. They show that the structure shown by corneal epithelium can be described through spontaneous organisation of cells. This is one of the first models to suggest that epithelial structure can be a consequence of a self-organised process. A different approach taken by Dubertret et al. [Dubertret and Rivier, 1997] models epithelial cell shape as a two-dimensional topological foam and is based on topology and statistical mechanics. Dubertret's approach predicts a strong correlation between the size of the attachment of the basal cells to the basement membrane and their biological fate (division or detachment from basement membrane). The models

proposed by Honda [Honda et al., 1996] and Dubertret [Dubertret and Rivier, 1997] to explain epithelial organisation are not applicable to our approach of individual based modelling since they do not model cells and their properties individually.

We believe cell shape should be modelled by simulating the mechanisms which cause it to change. In other words, we argue that the cause is what we should be modelling and not the effect. The cause of cell shape change may take place at a molecular level. Moreover, it is beneficial to address the problem using a method which requires little processing time until we have enough biological information about what causes the cell to change its shape. If we were to pursue the individual based model approach, modelling the cell shape from the molecular level will not be possible since the individual molecules would have to be modelled instead of the cells as proposed here. Drasdo's [Drasdo et al., 1995] approach appears to be more suitable for modelling at the cellular level and for incorporating into Stekel's model. We demonstrate a model of cell shape in a later chapter.

3.4 Simulating a stable normal epithelium

Healthy epithelial tissue does not alter in thickness significantly after its initial development. An example of this is our skin, which does not get infinitely thicker as we age. Hence the tissue must enter a dynamic equilibrium, where the rate of division equals the rate of death. Initial conditions in a simulation of epithelial tissue structure may vary from a few cells to a large collection of cells derived from an image of the epithelium. Therefore, it is important to define a stage in the model where the epithelium moves from development to one of maintenance.

Stekel [Stekel et al., 1995] used the raw number of cells of different types in the model to decide if dynamic equilibrium had been reached. He claimed that when cell numbers in the model do not change significantly, the model is in a state of dynamic equilibrium. An output from our implementation of Stekel's model as shown in Figure 3.13 demonstrates that, although cell numbers do not change significantly

there is no cell type ordering in the structure. Therefore, the epithelium shown in Figure 3.13 cannot be considered to represent a "normal" epithelium in a state of equilibrium. Thus, it is essential to take the distribution of cell type into consideration when defining a stable epithelium. A dysplastic epithelium exhibits a disordered structure in comparison to a healthy epithelium. We seek ways to define a dynamic equilibrium state using spatial positions of the cell as well as cell numbers. Our method employs Delaunay triangulation and connected component labelling.

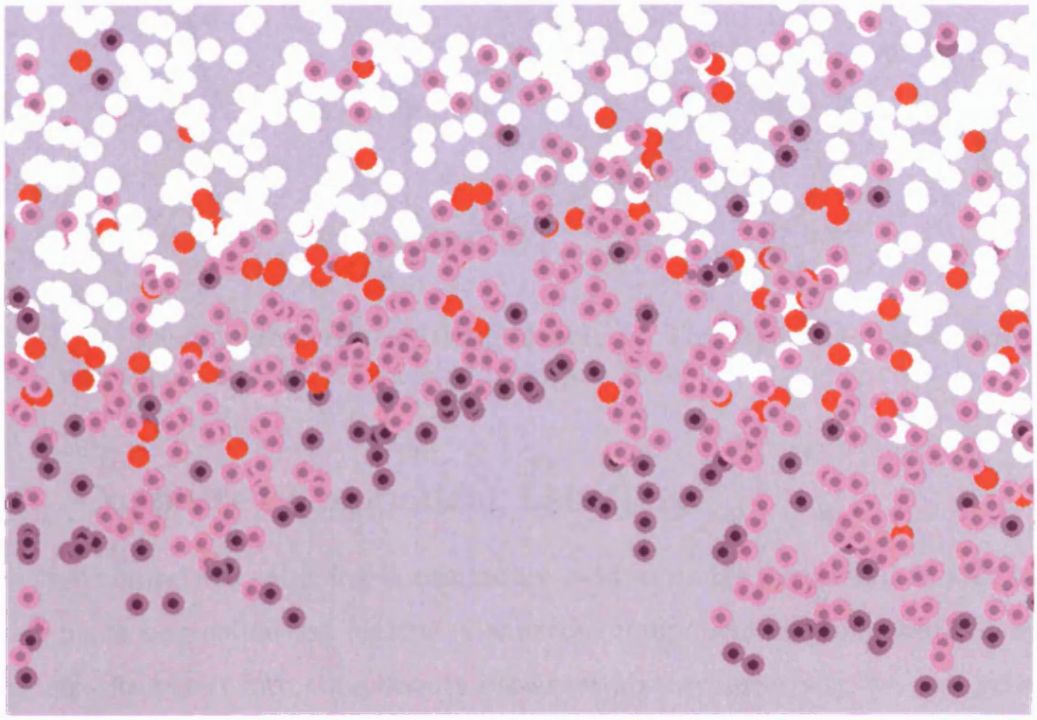
3.4.1 Delaunay Triangulation

There exists several ways to build a triangulation from a set of points. Figure 3.14 shows 3 different ways of generating triangles for a given set of points. A Delaunay triangulation of a vertex set is a triangulation of the vertex set with the property that no vertex in the vertex set falls in the interior of the circumcircle (a circle that passes through all three vertices) of any triangle in the triangulation [Okebe et al., 1992].

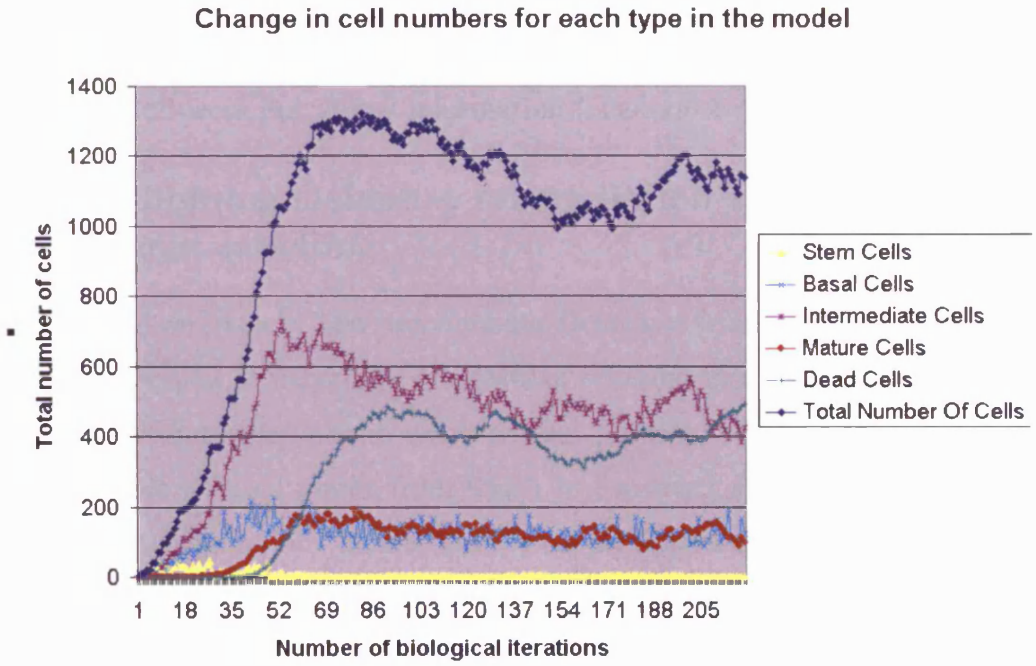
All triangles in the Delaunay triangulation for a set of points will have empty circumscribed circles. That is, no points lie in the interior of any triangle's circumcircle. Figure 3.15 shows the triangulation for the points shown in Figure 3.14. We can immediately see that the first triangulation is Delaunay, since all of its circumcircles are empty. There exists a Delaunay triangulation for any set of points in two dimensions. It is always unique as long as no four points in the point set are co-circular.

Delaunay triangulations are used to build topological structures from unorganized (or unstructured) points. Applications of Delaunay triangulation include computational geometry, simulation of the growth of crystals, mesh generation, computation of neighbouring points and visualisation of surfaces. We use the Visualisation Toolkit library [2] to build our Delaunay triangulation structures.

[2] http://public.kitware.com/VTK/



(a)



(b)

Figure 3.13: (a) We modify the compressive force in the model to 0.15, a five fold increase from the value chosen in Table 3.1 to obtain a disordered structure. (b) We demonstrate that even though cell numbers don't change, the epithelium generated in Figure 3.13(a) can not be considered normal due to the disordered nature of the cell layers. By disordered we mean no clear continuous layer or boundaries.

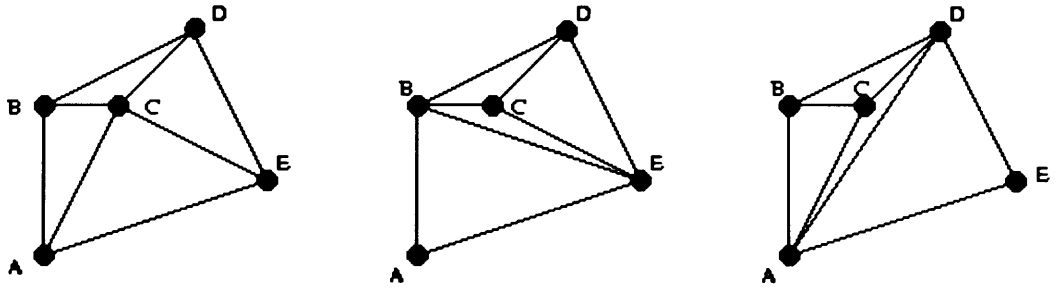


Figure 3.14: Triangulation of unstructured points. This figure is reproduced from "http://www.geom.uiuc.edu/ samuelp/delproject.html"

3.4.2 Connected Component Labelling

Connected component labelling is commonly used in image processing to join neighbouring pixels into connected regions. Connected component labelling scans an image and groups its pixels into components based on pixel connectivity, i.e., all pixels in a connected component share similar pixel intensity values and are in some way connected with each other [Davies, 1990]. The extraction and labelling of disjointed and connected components in an image is central to many automated image analysis applications. The main computational task is to take local information (bond connections) and work out global information (clusters of sites).

3.4.3 Combining Delaunay triangulation and Connected component labelling

In this section we explain how we combine Delaunay triangulation and Connected component labelling to determine the level of ordering in a set of cells, real or simulated, using information about cell type and position. We use the centres of the cells as the unstructured points from which to construct a topological structure using Delaunay triangulation. In this topological structure, the centres of cells become the vertices of the Delaunay triangles. Once we have built a topological structure of cells using Delaunay triangulation we assign a number as the label to each cell as

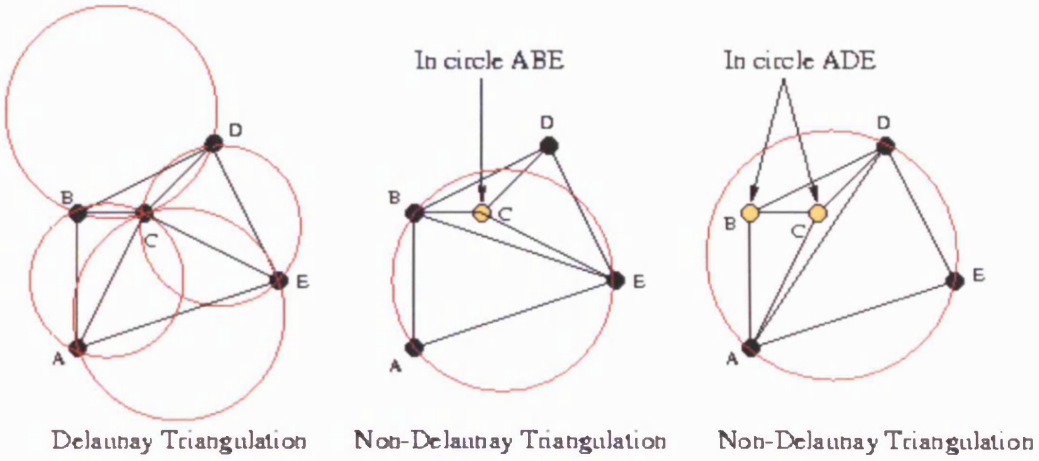
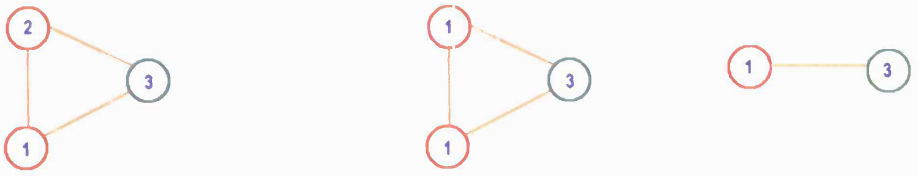


Figure 3.15: Figure shows possible triangulations of all arrangements. We can immediately see that the first triangulation is Delaunay. This diagram is reproduced from "http://www.geom.uiuc.edu/ samuelp/delproject.html"

shown in Figure 3.16(a). If the two vertices of an edge in the Delaunay triangle have different labels but the cells are of the same type, all instances of higher numerical labels of the vertices are then replaced with the lowest one until there are no further changes 3.16(b).

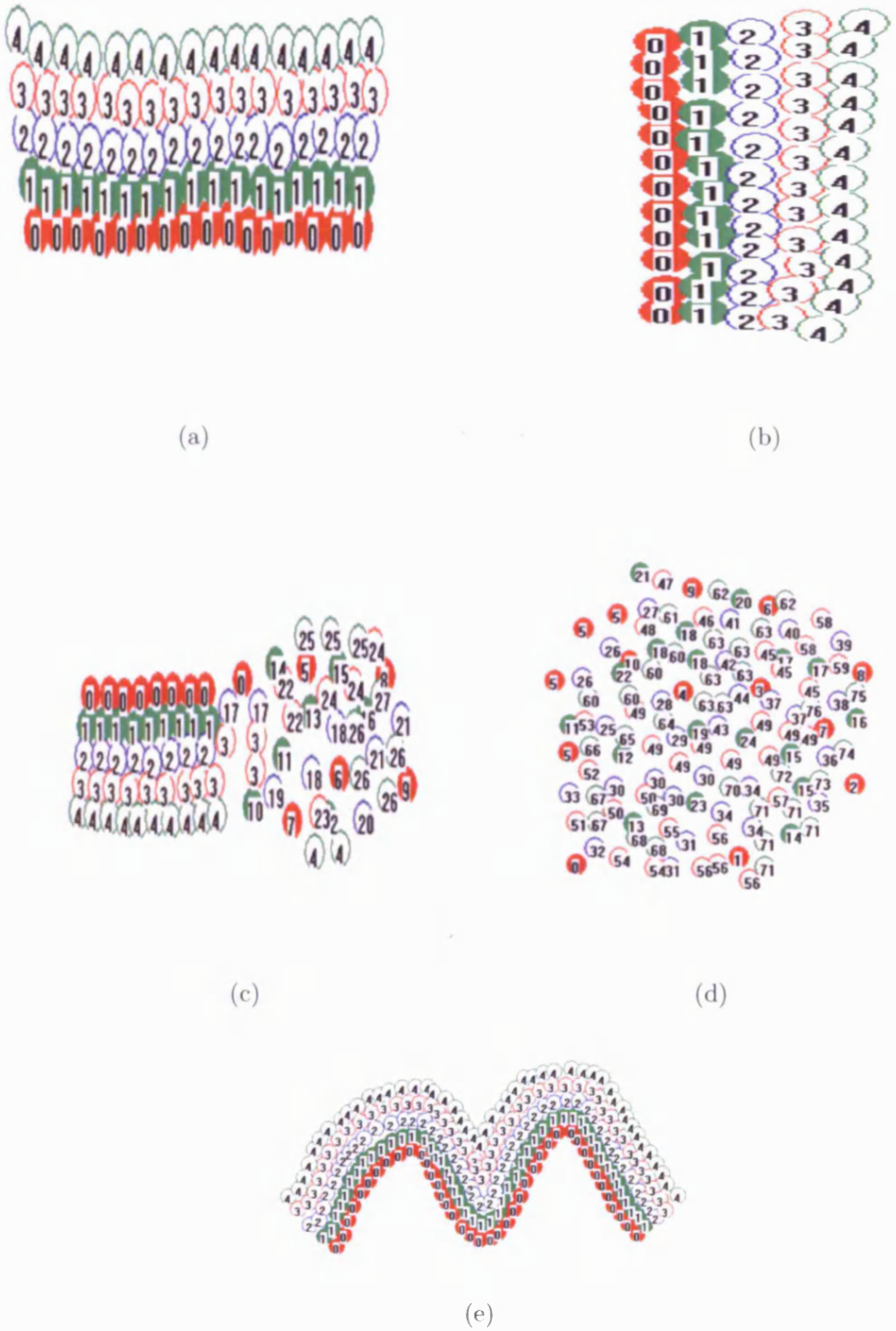
This is repeated for every Delaunay triangle in the topological structure. The end result of this process is to identify clusters of cells with the same label. Figure 3.16 shows edges connecting vertices with the same label being collapsed in order to identify different clusters. The greater the disorder in the system the more labels are required to label clusters. Figure 3.17 shows results of applying this algorithm to various distributions of cells. We apply our method to a model output as shown in Figure 3.18. As can be seen from Figure 3.18, we require large number of labels (above 100) to label the structure. An ordered structure of epithelium consisting of 5 different cell types would require 5 labels to identify the layers as demonstrated in Figure 3.17(e). Hence the number of labels required to label the epithelium may be used together with cell numbers to decide if the model has reached a state of dynamic



(a) (b) (c)

Figure 3.16: (a) An example of Delaunay triangulation for two different type of cells. (b)Same label is associated with two cells, since they are of same type and connected via an edge. (c) Edges connecting vertices with the same label are collapsed

equilibrium.



(a) (b)

(c) (d)

(e)

Figure 3.17: (a) An ordered structure of 5 different types of cells, which is labelled using 5 labels as expected. (b) An ordered structure of 5 different types of cells, which is labelled using 5 labels as expected. (c) A structure of 5 different types of cells, where one half represents some level of ordering and the other half is not ordered. Higher number of labels is required to label connected structures. (d) A structure of 5 different types of cells with no ordering in the structure. Larger number of label is required to label the structure. (e) A structure of 5 different types of cells which looks similar to a healthy epithelium. Only 5 labels are required to label the structure.

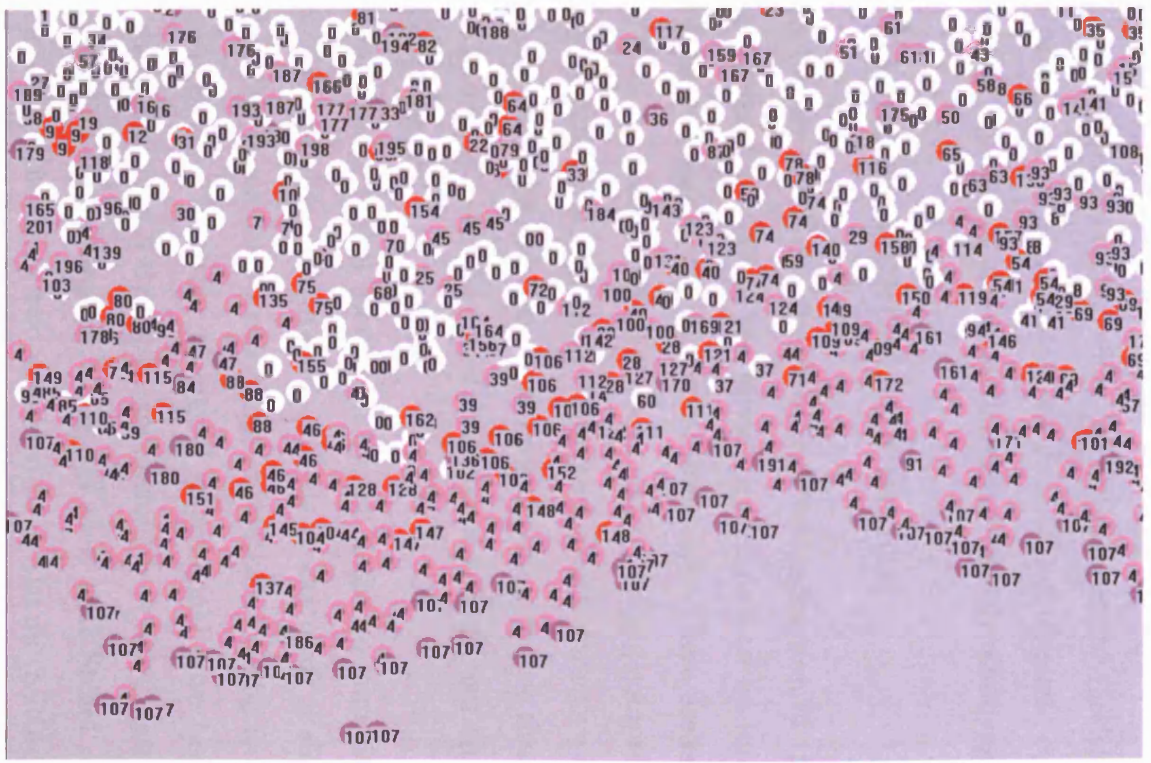


Figure 3.18: We apply our connected component labelling method to a sample of Stekels' model output. As can be seen, we require a large number of labels to label the clusters in the output. This could be used as a measure to quantitatively evaluate the level of ordering in the model output.

The number of different labels required to label an epithelial tissue structure in this way will give us insight into the level of ordering in the tissue. We believe this approach together with numbers of cells in the system as described by Stekel [Stekel et al., 1995], gives a more useful test to see if the model has reached a state of dynamic equilibrium. Stekel claimed that when cell numbers in the model do not change very much, the model can be considered to be in a state of dynamic equilibrium. We have shown here that this is not adequate to describe the dynamic equilibrium.

3.5 Discussion and Conclusion

In this chapter, we have provided a rational reconstruction of the Stekel's model of epithelium. Several aspects of the model were shown to be inadequate or at least without strong evidence in the biological literature. These areas included the direction of cell division, computing neighbour cells, modelling compressive forces and cell shape. In particular, the suggested connective tissue attractive and repulsive forces appear to have no biological basis. Such assumptions can be made to produce desirable graphical output from the model, but do not enhance our current understanding of biological processes.

Stekel proposed a model motivated by the desire to simulate a stable, healthy epithelium. He initialised the execution of his model using a single cell or a line of a few cells. This is similar to modelling an epithelial tissue from a developmental biology perspective. Two other biological stages could have been tested by Stekel. They are modelling how a dynamic equilibrium is reached and modelling how the tissue organisation is maintained. It is our view that Stekel only tested the first phase, even though the latter two appeared to be the motivating factor.

Our implementation of the Stekel model had 10 input parameters. We chose arbitrary values as input parameters. It would be more favorable to use input values obtained from the biological experiments literature or new experiments. In order to use these values from biological experiments, we must make sure that the modelling

scales involved are the same. When modelling a biological problem, the models are built using the known biological processes. It is our opinion that in a good model results from simulation would suggest new biological experiments, which would ultimately enhance the understanding of the biological problem being modelled. Therefore, it is vital to keep the information flow in both directions between the biological problem and the model. However, prior to the flow of information from the model to suggest new biological experiments, it is important that the model is validated. The model should be validated using *in vitro* experiments, *in vitro* images or *in vivo* experiments. This will improve confidence in the model output significantly.

Chapter 4

Using *in vitro* images to evaluate in silico models

4.1 Introduction

Landini [Landini and Rippin, 1996] previously used fractal techniques to distinguish between normal and dysplastic structures for diagnostic purposes by measuring curvature along the basement membrane and classifying using Linear Discriminant Analysis(LDA). In contrast, we capture curvilinear features on a global scale. We present a method which combines Fourier Descriptors, Principal Component Analysis(PCA) and closest mean classification techniques to distinguish labelled classes of tissue images using the curvilinear shape of features found in the image set. We employ Fourier Descriptors to represent boundary features and PCA to reduce dimensionality. The labelled classes are then used in a supervised fashion to train the closest mean algorithm to distinguish unseen examples from the different classes. We demonstrate this combination of techniques on histological images of normal and dysplastic (abnormal) tissue samples taken from two sites of the oral mucosa, the tongue and inner cheek (buccal surface).

The data set of images consisted of randomly chosen images of normal epithelium, 30 from tongue mucosa and 40 from buccal mucosa. We also used 10 randomly chosen images of dysplastic tissue. Dysplasia causes the shape of the basement membrane,

the boundary between the epithelium and extracellular matrix, to change irrespective of its location in the mouth. Hence, we use this shape change to build our model. An example of buccal epithelium and a manually traced basement membrane is shown in Figure 4.1. These tracings were used to build a statistical model of the shape variation.

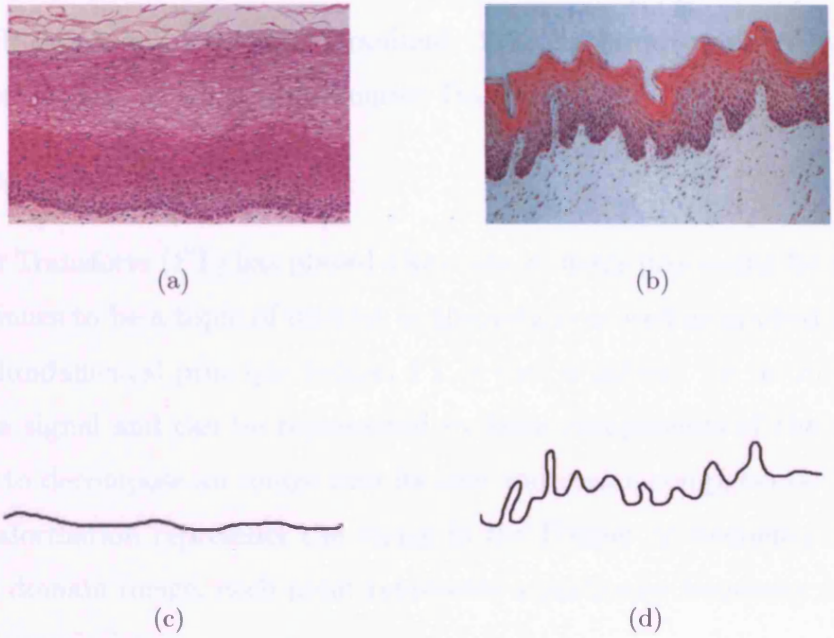


Figure 4.1: (a) A histological section of normal buccal epithelium. (b) A section of dysplastic epithelium. (c) Shape of the normal basement membrane from 4.1(a). (d) Shape of dysplastic basement membrane traced from 4.1(b).

4.2 Feature Extraction, Boundary Representation and Principal Component Analysis

Feature extraction and object recognition are active research areas in the field of computer vision and image processing. Object recognition is largely based on matching descriptions of shapes. There are numerous shape description techniques, such

as analysis of scalar features, Fourier descriptors, moment invariants and boundary chain coding. A detailed description of these techniques can be found in Nixon [Nixon and Aguado, 2002]. The application of these techniques is well understood when applied to images and has been developed to describe shapes irrespective of scale, rotation and translation. The application of one of these techniques, Fourier descriptors is described here. The Fourier descriptor is an all purpose shape description technique [Persoon and Fu, 1986, Granlund, 1972, Leicester et al., 1998]. Fourier Descriptors(FD) are obtained from Fourier Transforms.

4.2.1 Fourier Transform

The Fourier Transform (FT) has played a key role in image processing for many years, and it continues to be a topic of interest in theoretical as well as applied work in this field. The fundamental principle behind FT is that a pattern (or an image) can be treated as a signal and can be represented by basic components of the signal. The FT is used to decompose an image into its sine and cosine components. The output of the transformation represents the image in the Fourier or frequency domain. In the Fourier domain image, each point represents a particular frequency contained in the spatial domain image.

If a function $f(x)$ can be represented by its Fourier series, then $f(x)$ is uniquely determined by its Fourier coefficients, c_n. Inversely, if the Fourier coefficients C_k of its Fourier series are known, $f(x)$ can be reconstructed from the set of c_n. The Fourier series establish an unique correspondence between f(x) and its Fourier coefficients. This correspondence is expressed by the Fourier transform of $f(x)$:

$$F(u) = \int_{-\infty}^{\infty} f(x) exp(-j2\pi ux) dx \tag{4.1}$$

Since our application deals with images, we require the Discrete Fourier Transform (DFT). In order to convert Equation 4.1 into discrete form, $f(x)$ is sampled into N samples, which may be given as

$$f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \cdots, f(x_0 + (N-1)\Delta x)$$

where Δx is called the sample step in spatial domain, then $f(x)$ can be expressed as

$$f(x) = f(x_0 + x\Delta x), x = 0, 1, 2, \cdots, N-1$$

The DFT is then given as

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp\left(-\frac{j2\pi ux}{N}\right), \qquad u = 0, 1, 2, \cdots (N-1) \tag{4.2}$$

The sample step Δu in frequency domain and the sample step Δx in spatial domain are related by the expression

$$\Delta u \geq \frac{1}{N\Delta X}$$

For a two variable function $f(x, y)$ defined in $0 \leq x, y < N$, the DFT is given by

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{x=0}^{N-1} f(x, y) \exp\left(-\frac{j2\pi(ux + vy)}{N}\right), \qquad u, v = 0, 1, 2, \cdots (N-1) \tag{4.3}$$

Fourier coefficients, from here in referred to as the Fourier descriptor values produced by FT provide a description of shape. Zhang et el [Zhang and Lu, 2002] provide a comprehensive review of shape description and image retrieval using Fourier descriptors. The Fourier descriptors C_k are determined using Equation 4.4.

$$C_k = \sum_{n=0}^{N-1} f(x) \exp\left(\frac{j2\pi kn}{N}\right), \qquad 0 \leq k \leq N-1 \tag{4.4}$$

Properties of Fourier Descriptors

In reality images are obtained under various camera settings and they need to be normalised before analysis can be carried out. FD values provide a description of shape, which may be made invariant under scale, rotational and translational using the properties of Fourier descriptors.

- Translation The translation property of the Fourier transform is given by

$$f(x - x_0, y - y_0) \Longleftrightarrow F(u, v).exp[-j2\pi(ux_0 + vy_0)/N]$$

  This indicates that a shift in spatial domain results in phase change in frequency domain.

- Rotation The rotation property of the Fourier transform is given by

$$f(r, \theta + \theta_0) \Longleftrightarrow F(\omega, \phi + \theta_0)$$

  In other words, rotating $f(x, y)$ by an angle of θ_0 in spatial domain rotates F(u, v) by the same angle in frequency domain.

- Scaling For two scalars a and b, the scale property of Fourier transform is given by

$$f(ax, by) \Longleftrightarrow) \frac{1}{ab} F(u/a, v/b)$$

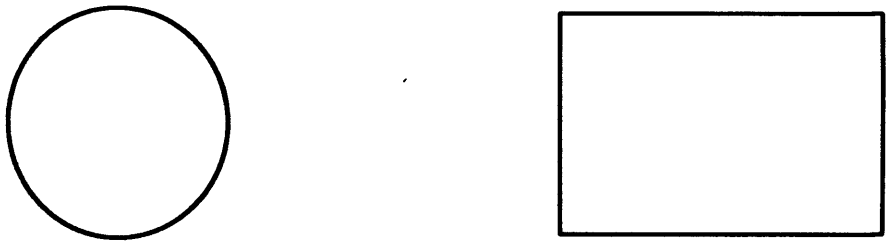
It indicates the scaling of $f(x, y)$ with a and b in x and y directions in spatial domain (time domain in 1-D case) causes inverse scaling of magnitude of $F(u, v)$ in frequency domain. It also reduces the number of $F(u, v)$ by $1/a$ and $1/b$ in u and v directions in frequency domain. In general terms, enlarging an object in an image gives rise to lower frequencies in spectral domain while shrinking an object in an image gives rise to higher frequencies in spectral domain. We use these properties of

Transformation	Boundary	Fourier Descriptors
Translation	$\tilde{f}(x) = f(x) + f_0$	$\tilde{C}(k) = C(k) + f_0\delta(k)$
Scaling	$\tilde{f}(x) = \alpha f(x)$	$\tilde{C}(k) = \alpha C(k)$
Rotation	$\tilde{f}(x) = f(x)e^{j\theta_0}$	$\tilde{C}(k) = C(k)e^{j\theta_0}$

Table 4.1: Properties Of Fourier Descriptors

Fourier descriptors to normalise our data. Table 4.1, shows the properties of FDs for a 1-D case.

The lower frequency descriptors describe the general shape of the object, and the higher frequency descriptors store information about finer details in the image. In order to illustrate the use of Fourier descriptors, we represent a circle and rectangle using Fourier descriptors. The circle and rectangle are shown in Figure 4.2.1 and the resulting Fourier descriptors are shown in Figure 4.3. We show only the first 100 Fourier descriptors of both shapes since this is done for illustrative purposes and not with accuracy in mind.



(a) (b)

Figure 4.2: (a) An image of a Circle. (b) An image of a rectangle

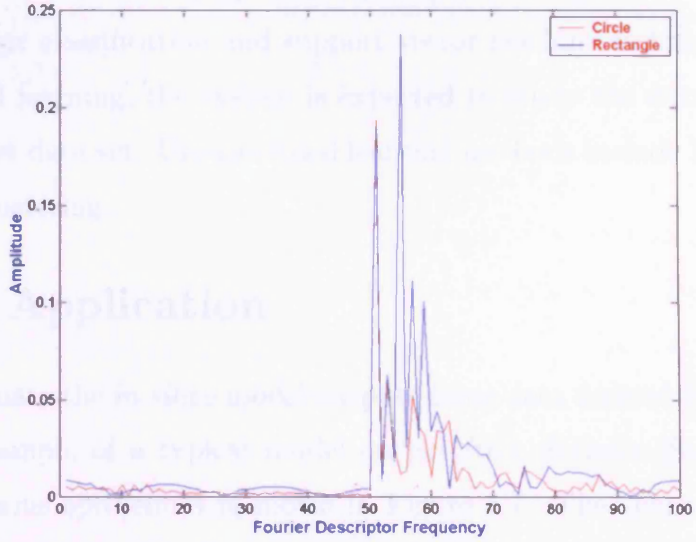


Figure 4.3: Fourier descriptor representation of the circle and rectangle shown in Figure 4.2.1.

Principle Component Analysis (PCA) is a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. A detailed explanation of PCA can be found in Appendix A.

4.3 Classification

In order to distinguish between two kinds of objects using their shape, we require both shape representation and a classification algorithm. Classification techniques use either supervised or unsupervised learning methods. Supervised learning methods are where the system is trained to classify a feature vector. A supervised learning algorithm uses a training set which consists of labelled classes to predict labels for

the unseen test data set. Techniques in supervised learning include closest mean, k-nearest neighbour classification and support vector machine techniques. In the case of Unsupervised learning, the system is expected to study the data set and predict labels for the test data set. Unsupervised learning methods include Bayesian learning and k-Means clustering.

4.4 The Application

We wish to evaluate the *in silico* model outputs using data derived from human tissue images. An example of a typical model output from Stekel's [Stekel et al., 1995] model of squamous epithelium is shown in Figure 3.7. The shape of the basement membrane changes when the tissue becomes dysplastic. In order to evaluate if the model correctly simulates a "normal" basement membrane, help could be sought from an expert or we may use existing labelled images. We suggest a method, which aims to capture the variation between normal and dysplastic tissue images.

The boundary of the basement membrane in images of normal and dysplastic tissue were traced manually. Examples of such tracings were shown earlier in Figure 4.1(c) and Figure 4.1(d). The tracings are then chain coded to avoid any possibility of over hangs in the shape of basement membrane. The definition given in Equation 4.4 is used to compute the Fourier descriptors for the tracings of normal and dysplastic tissues. As mentioned earlier, since images could be obtained under various camera settings it is essential that we make the Fourier descriptors invariant under scale, translation and rotation. We accomplish this by using the properties of Fourier descriptors defined in Jain et el [Jain, 1989]. The FDs are made rotation invariant by ignoring the imaginary phase angle as shown in Table 4.1. Scale invariance is achieved by dividing all FDs by the first component and translational invariant is obtained by defining all FDs from the same reference point. We choose the first FD as our reference point.

An exponential decay type of a curve in Fourier space represents a "smooth" line

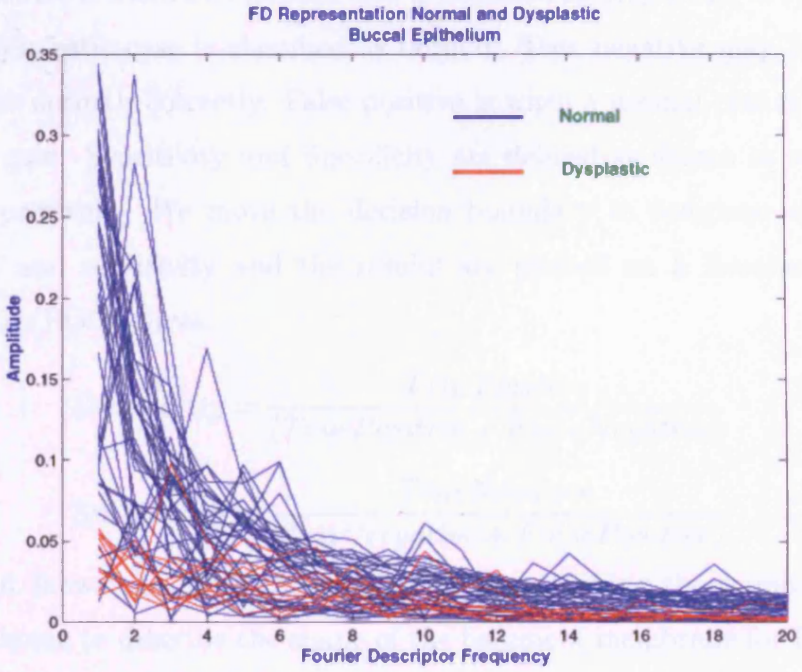
in feature space. The "random" behavior of descriptors in Fourier space represents a "wavy" line in cartesian space [Jain, 1989]. Hence we expect our Fourier descriptors for the normal basement membrane to be distributed along an exponential decay type of curve, and dysplastic membrane set of Fourier descriptors to be randomly distributed in Fourier space. Figure 4.4 illustrates normal and dysplastic basement membrane representations in Fourier space which indeed display an exponential decay for normal basement membranes and in the case of dysplastic membranes, Fourier descriptors can be seen to be randomly distributed.

We use closest mean classifier as our classification tool. The closest mean classification involves the computation of a mean feature vector for the training set in each class. Given a test feature vector, the Euclidean distance measure given in Equation 4.5 is used to compute which mean feature vector the test sample is closest to, and the test sample is assigned a label accordingly.

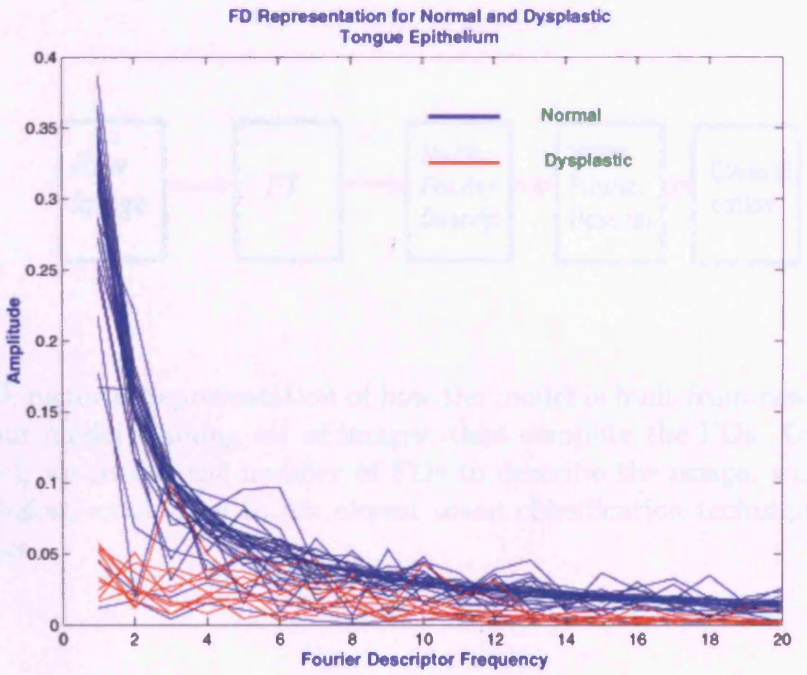
$$d = \left(\sum_{i=0}^{N} |f_m^i - f_d^i|^2 \right)^{\frac{1}{2}} \tag{4.5}$$

where f_m is the mean shape Fourier descriptors, and indexed by $f_m = [f_m^1, f_m^2, \ldots, f_m^N]$, the test feature vector is indexed by $f_d = [f_d^1, f_d^2, \ldots, f_d^N]$ and N is the length of the feature vector.

Our classification is performed using the 10 fold cross validation technique as described by Kohavi et el [Kohavi, 1995]. We randomly select 90% of normal and dysplastic images as our training set and the rest of the images are used as the test cases. For example, if there are 40 normal and 10 dysplastic images in the data set, then the training set of images contains 36 randomly chosen images from the normal training set and 9 randomly chosen images from the dysplastic set. The remaining 5 images are then used as the test set. This randomised selection was repeated 1000 times and the average classification accuracy was computed in terms of specificity and sensitivity.



(a)



(b)

Figure 4.4: (a) Fourier descriptors for 40 normal Buccal epithelium and 10 dysplastic sections of oral epithelium. (b) Fourier descriptors for 30 normal Tongue epithelium and 10 dysplastic oral epithelium

A true positive is when a dysplastic case is classified as dysplastic. A False negative is where a dysplastic case is classified as normal. True negative may be defined as classifying the normals correctly. False positive is when a normal case is classified as a dysplastic case. Sensitivity and Specificity are defined as shown in equations 4.6 and 4.7 respectively. We move the decision boundary to compute several values of specificity and sensitivity and the results are plotted on a Receiver Operating Characteristics(ROC) curve.

$$Sensitivity = \frac{TruePositive}{(TruePositive + FalseNegative)} \tag{4.6}$$

$$Specificity = \frac{TrueNegative}{(TrueNegative + FalsePositive)} \tag{4.7}$$

Figure 4.6 shows the ROC curve resulting from varying the number of Fourier descriptors chosen to describe the shape of the basement membrane for Buccal 4.6(a) and Tongue Epithelium 4.6(b). Figure 4.5 provides a pictorial representation of how we built the model from raw images to classification.

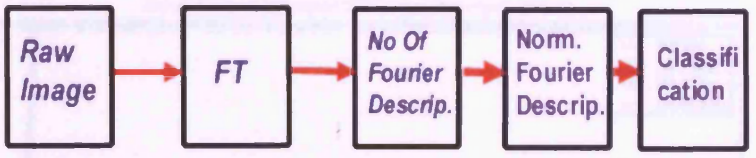
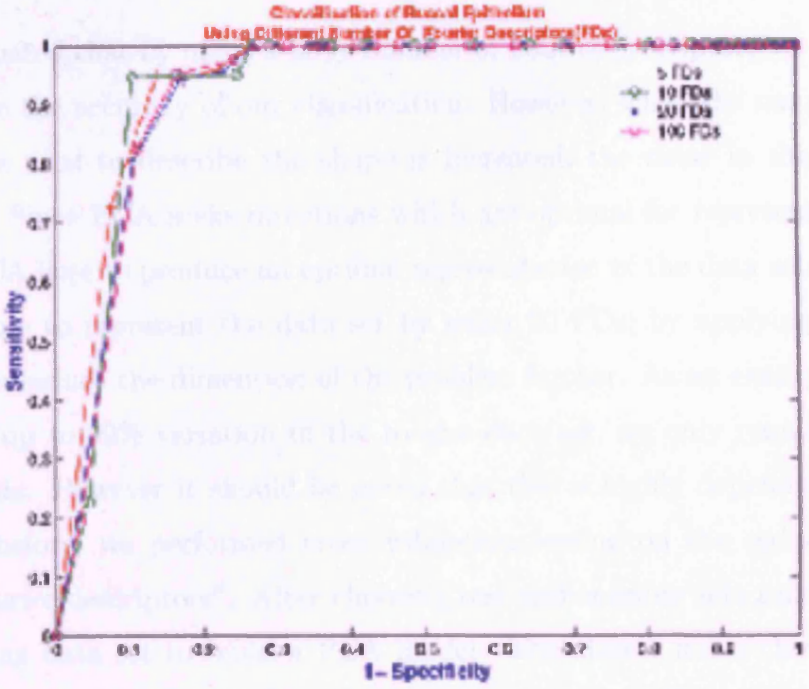
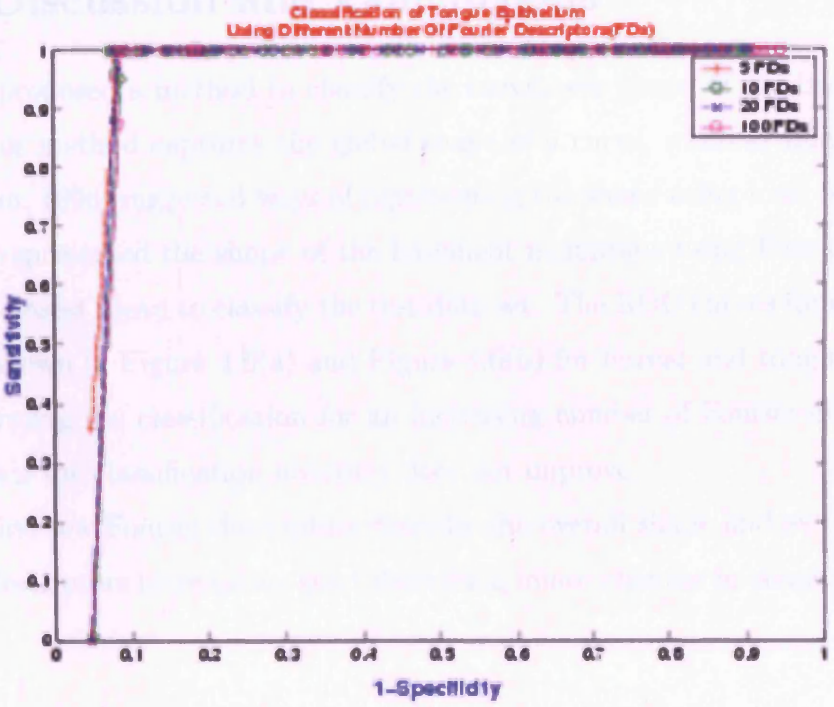


Figure 4.5: A pictorial representation of how the model is built from raw images. We first select our model training set of images, then compute the FDs. Once the FDs are computed, we choose the number of FDs to describe the image, which are then normalised before moving on to use closest mean classification technique to classify unseen images.



(a)



(b)

Figure 4.6: (a) ROC curve for Buccal epithelium using 5,10,20 and 100 Fourier descriptors. (b) ROC curve for Tongue epithelium using 5,10,20 and 100 Fourier descriptors

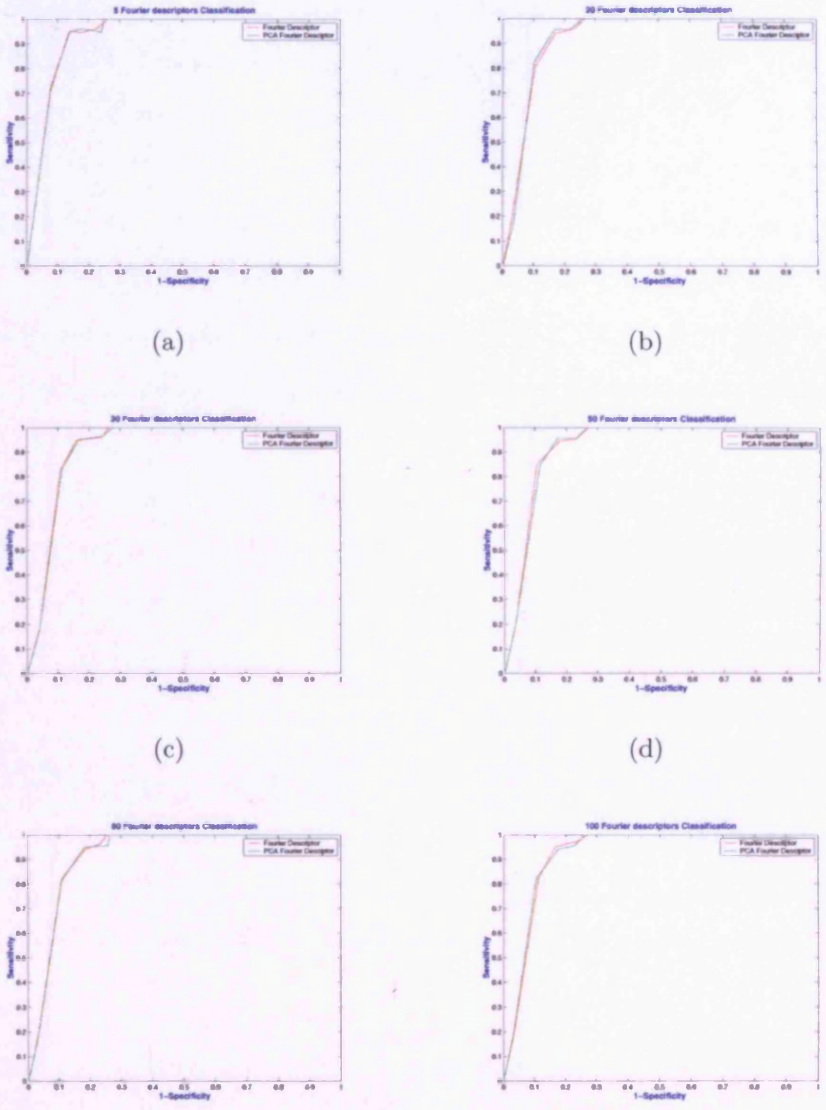
4.4.1 Applying PCA before classification

We anticipated that by using a large number of Fourier descriptors we would be able to improve the accuracy of our classification. However, when the number of Fourier descriptors used to describe the shape is increased, the noise in the data set also increases. Since PCA seeks directions which are optimal for representation of data, we use PCA here to produce an optimal representation of the data set. For example, if we choose to represent the data set by using 20 FDs, by applying the PCA, we are able to reduce the dimension of the problem further. As an example, in order to represent up to 99% variation in the tongue data set, we only required 6 principal components. However it should be noted that this is highly dependent on the data set. As before, we performed cross validation testing on the reduced number of "PCA Fourier descriptors". After choosing test and training sets randomly, we used the training data set to build a PCA model. The closest mean classifier was then used to classify the test samples. The results are shown in Figure 4.7 and Figure 4.8 for buccal and tongue epithelium respectively.

4.5 Discussion and Conclusions

We have proposed a method to classify the curvilinear shape of the basement membrane. Our method captures the global shape of a curve, whereas Landini [Landini and Rippin, 1996] suggested ways of representing the shape using local features of the line. We represented the shape of the basement membrane using Fourier descriptors and used closest mean to classify the test data set. The ROC curves for our classification are shown in Figure 4.6(a) and Figure 4.6(b) for buccal and tongue epithelium. By performing the classification for an increasing number of Fourier descriptors, we have shown the classification accuracy does not improve.

The first few Fourier descriptors describe the overall shape and as the number of Fourier descriptors increase we start describing minor changes in shape, which would

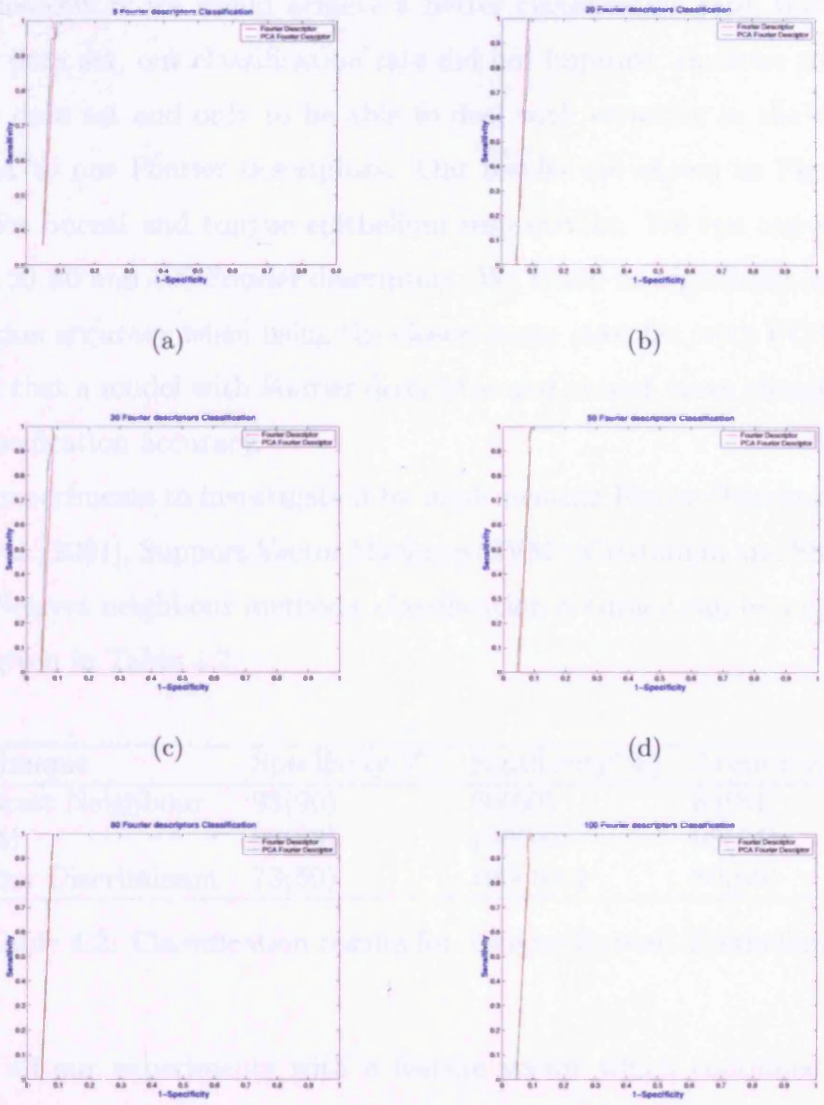


(a) (b)

(c) (d)

(e) (f)

Figure 4.7: Comparing classification accuracy between Fourier Descriptors (FDs) and PCA FDs for buccal epithelium. As can be seen, the classification accuracy does not improve when PCA is applied. (a) 5 FDs and 5 PCA FDs. (b) 20 FDs and 20 PCA FDs. (c) 30 FDs and 30 PCA FDs. (d) 50 FDs and 50 PCA FDs. (e) 80 FDs and 80 PCA FDs. (f) 100 FDs and 100 PCA FDs.



(a) (b)

(c) (d)

(e) (f)

Figure 4.8: Comparing classification accuracy between Fourier Descriptors (FDs) and PCA FDs for tongue epithelium. As can be seen, the classification accuracy does not improve when PCA is applied. (a) 5 FDs and 5 PCA FDs. (b) 20 FDs and 20 PCA FDs. (c) 30 FDs and 30 PCA FDs. (d) 50 FDs and 50 PCA FDs. (e) 80 FDs and 80 PCA FDs. (f) 100 FDs and 100 PCA FDs.

also include the noise in the data set. We anticipated that by increasing the number of Fourier descriptors we would achieve a better classification rate, but due to the noise in the data set, our classification rate did not improve. In order to reduce the noise in the data set and only to be able to deal with variation in the data set, we applied PCA to our Fourier descriptors. Our results are shown in Figure 4.7 and Figure 4.8 for buccal and tongue epithelium respectively. We ran our experiments for $5, 20, 30, 50, 80$ and 100 Fourier descriptors. We found no significant improvement in classification accuracy when using the closest mean classifier with PCA. Therefore we conclude that a model with Fourier descriptor and closest mean classifier provides the best classification accuracy.

We ran experiments to investigate if by implementing Fisher Discriminant Analysis [Duda et al., 2001], Support Vector Machines(SVM) [Cristianini and Shawe-Taylor, 2000], and Nearest neighbour methods, classification accuracy can be improved. Our results are given in Table 4.2.

Technique	Specificity(%)	Sensitivity(%)	Accuracy(%)
Nearest Neighbour	93(90)	60(60)	85(84)
SVM	93(83)	100(90)	95(84)
Fisher Discriminant	73(50)	100(100)	80(60)

Table 4.2: Classification results for Tongue(Buccal) Epithelium

We ran all our experiments with a feature vector which contained 20 Fourier descriptors using the 10 fold cross validation technique. Our Fourier descriptors were normalised and we applied PCA techniques to the Fourier descriptors. As can be seen from Table 4.2, combining Fourier descriptors, PCA and SVM yields the most accurate classification results. However more work is needed to analyse what would be the effect of including large or small number of Fourier descriptors in the feature set before stating that best accuracy can be obtained by combining Fourier descriptors, PCA and SVM. We should also do some work to investigate what happens to the

accuracy of the model if the PCA component is removed.

Chapter 5

The Model

We present an individual based model to simulate epithelial cell interactions. The model consists of biological processes such as cell adhesion, differentiation and division. Cell types included are Stem, Transit Amplifying (TA), Intermediate, Mature and Dead. Stem and Transit Amplifying cells are allowed to mitose provided they are on the basement membrane. Cells increase in size before mitosis and subsequently differentiate as suggested by Jensen [Jensen et al., 1999]. Cell sorting arises from the Differential Adhesion Hypothesis(DAH) proposed by Steinberg et al [Steinberg, 1963], and is modelled using the cellular automata approach of Glazier et al [Glazier and Graner, 1993]. We introduce our method of modelling cell behavior but defer simulating biological experiments to a later chapter. We demonstrate in this chapter, that it is possible to generate an ordered structure of epithelial tissue from initial conditions such as a single cell as well as a line of cells.

5.1 Differential Adhesion Hypothesis

Embryonic cells of two types, when randomly mixed and dissociated, can sort themselves to reestablish a homogeneous tissue. Cell sorting has been observed experimentally as the process which drives the ordering process. Steinberg [Steinberg, 1963] postulated the Differential Adhesion Hypothesis (DAH) to explain this process.

The DAH states that cells can explore various configurations and arrive at the global minimum of overall surface energy. Driving forces for the cell sorting arise from differential surface energies between cells and the external medium. Glazier([Graner and Glazier, 1992], [Glazier and Graner, 1993]) presented a 2-dimensional Cellular Automaton (CA) to simulate cell sorting using the DAH. His model is known as the extended large-Q Potts model, which sometimes is referred to as the Glazier-Graner model. The Ising model [Ising, 1925], the precursor to the Glazier-Graner model tries to imitate behaviour in which individual elements (e.g., atoms, animals, protein folds, biological membrane, social behavior, etc.) modify their behaviour so as to conform to that of other individuals in their vicinity. Ising tried to explain certain empirically observed facts about ferromagnetic[1] materials.

In the Ising model, each lattice site can be in one of two different states. In the case of magnetic materials, electrons can be in one of two different states. In physics, this is referred to as the spin associated with an electron which could either be up or down. Potts [Potts, 1952] extended the Ising model by assuming each lattice site can be in one of q different states (when $q = 2$, both models are equivalent). Pott's model was initially used to simulate order-disorder transformations in non-biological patterns. Glazier and Graner focussed on the spatial dynamics of the Potts model with a high number of states. In applying this to biology, a group of the same states(ie lattice sites) refers to an individual biological cell. Savill [Savill, 2003], [Savill and Sherratt, 2003] and Mareé [Mareé, 2000] applied the Glazier-Graner model to modelling biological behaviour. Savill's work focussed on Dictyostelium and stem cells cluster formation. We apply the Glazier-Graner cell sorting algorithm to

[1]The development of extremely strong magnetic properties in certain materials which occurs when magnetic domains (regions at most 1mm in dimension) become aligned in the absence of an applied field, below a temperature known as the Curie temperature. The net magnetization depends on the magnetic history (the hysteresis effect). Above the Curie temperature, these materials become paramagnetic. Iron, nickel, cobalt, and gadolinium are ferromagnetic at room temperature. Ferromagnetism is believed to be caused by magnetic fields generated by the electrons' spins in combination with a mechanism known as exchange coupling, which aligns all the spins in each magnetic domain. This is reproduced from "http://scienceworld.wolfram.com/physics/Ferromagnetism.html"

the simulation of epithelium.

The Glazier and Graner model is usually embedded in a $2D$ lattice, but may be extended to three dimensions. The model describes a collection of N cells, where each cell has a unique identifier, $\rho = 1, 2, \cdots, N$. A cell consists of all lattice sites with identifier ρ. At each time step, a lattice site is chosen at random and the probability of the lattice site being copied into the neighbour (see below) is either 1 if $\Delta H \leq 0$ or $exp(-(\frac{\Delta H}{\kappa t}))$ where ΔH is the change in energy if the copying were to occur, κ is the Boltzmann constant and T is the temperature in the system (see later).

However, when modelling biological cells, the Potts model has to be modified. Glazier and Graner made two extensions to it. Biological cells may be assumed to be of fixed size. By this we mean they don't get infinitely bigger. In order to accommodate this they added to the energy function 5.1 an extra term to describe the area constraint for each individual biological cell. Glazier and Graner took this into account by introducing a second term $\lambda(\nu_\rho - V)^2$ into the energy function shown in Equation 5.1. The adhesive bond strength between biological cells of the same type will be stronger than the adhesive strength between biological cells of different types, according to the DAH. The Γ in equation 5.1 indicates the biological cell type. The surface energy is defined between two lattice sites occupied by two biological cells of different types as $\Gamma_1\Gamma_2$ or same types as $\Gamma_1\Gamma_1$. The Glazier-Graner model is illustrated in Figure 5.1. The Hamiltonián energy function for the Glazier and Graner model may then be given as in Equation 5.1 as suggested by Savill [Savill, 2003] and Mareé [Mareé, 2000]. Savill [Savill, 2003] used the Glazier-Graner model to explain spatial self organisation, cell sorting and slug migration of the cellular slime mold *Dictostelium discoideum.* We use equation 5.1 to simulate cell sorting in our model.

$$H_\rho = \sum_{\text{all } \rho,\rho' \text{neighbours}} \frac{J_{\Gamma_\rho,\Gamma_{\rho'}}}{2} + \sum_{\text{all } \rho,\text{medium}} J_{\Gamma_\rho \Gamma_{\text{medium}}} + \lambda(\nu_\rho - V)^2 \qquad (5.1)$$

where ν_ρ is the area of the biological cell ρ, V is the target area of ρ and λ is the inelasticity or rigidity of the cell membrane. The lower the value of λ, the easier it will

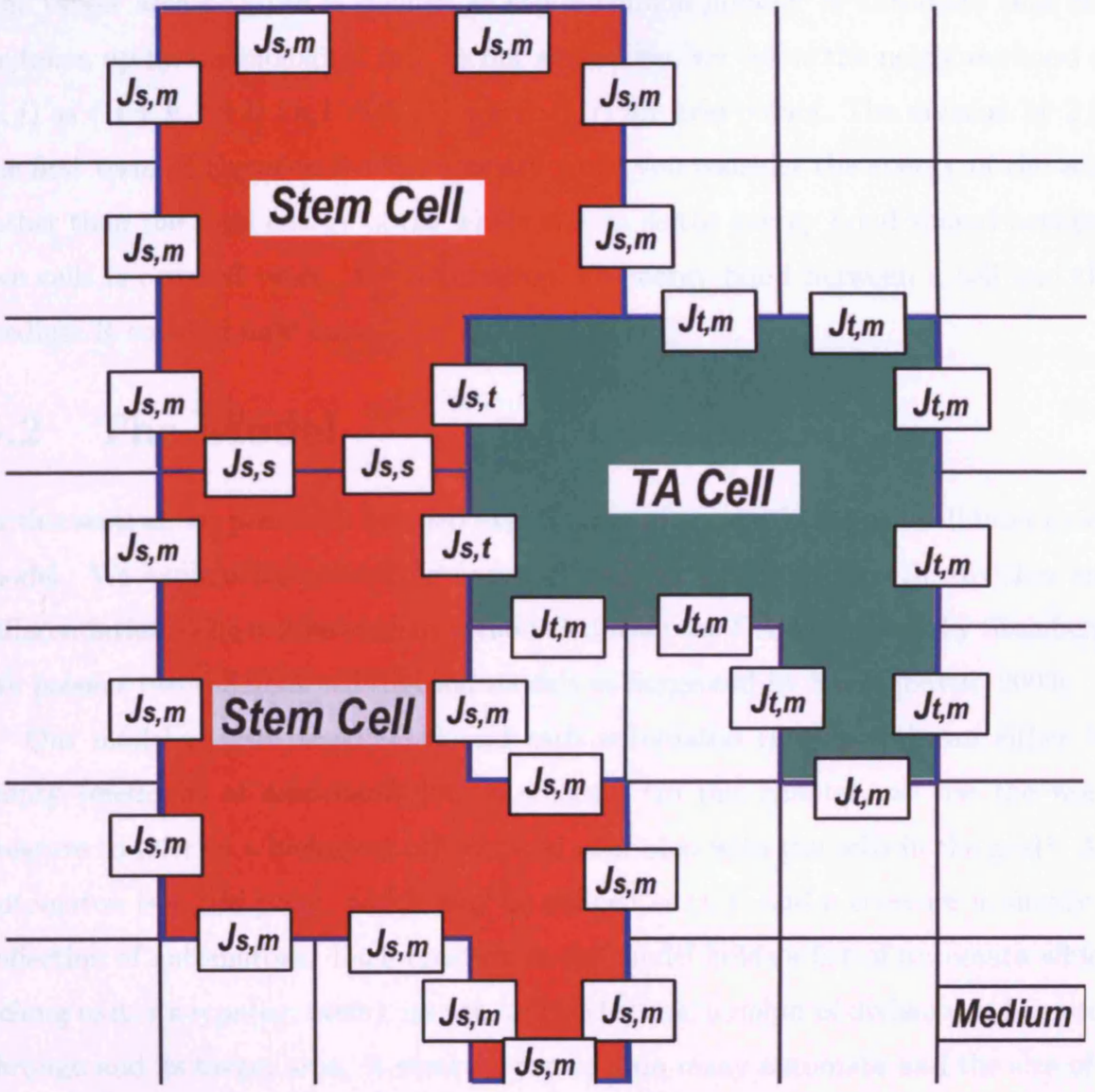


Figure 5.1: A diagrammatic representation of the Glazier and Graner model. We show two different types of biological cells and the medium. All the surface contact energies are also drawn $J_{\rho\rho_1}$. This figure is a modified version of the representation of Glazier-Graner model produced by Mareé [Mareé, 2000]

be to deform the cell membrane. The area of a biological cell refers to the number of lattice sites in the grid which we refer to in this chapter as the number of automata. The target area of a cell is defined as the maximum number of automata that can be taken up by the biological cell. In our simulation, we define the neighbourhood of (i, j) as $\{(i \pm k, j \pm l) \mid k, l \ \epsilon\{0, 1\}\}$ where (i, j) are grid points. The division by 2 in the first term of Equation 5.1 is necessary when you consider the energy of the cells rather than the total energy of the whole system as the energy bond shared between two cells is counted twice. By comparison, an energy bond between a cell and the medium is counted only once.

5.2 The Model

In this section, we present a detailed explanation of our epithelial cell-cell interaction model. We explain our modelling approach for cell adhesion, growth, division and differentiation. The cell movement is modelled from the DAH proposed by Steinberg. We present two different cell division models as suggested by Savill [Savill, 2003].

Our model is grid based (CA) and each automaton (grid point) can either be empty (medium) or associated with a creature (In this chapter, we use the word creature to refer to a biological cell to avoid confusion with the cells in the grid). An automaton is a grid point, which may be defined as (i, j) and a creature is simply a collection of automatons. Each creature in the model holds a list of automata which belong to it, its type(eg: stem), its age, cell cycle time, number of divisions it has gone through and its target area. A creature can contain many automata and the size of a creature is defined by the number of automata associated with it. All automata on a creature membrane have dimensionless free energy bonds associated with them. The energy bonds connect an automaton with its neighbouring 8 automata. The bonds which contribute to the energy are the ones which connect it to another creature with a different state or with the surrounding medium. The strength of these energy bonds will depend on the type of creature each automaton connects to in the grid.

5.2.1 Cell Adhesion

The idea of tissues being held together through cellular adhesion goes back to the work of von Recklinghausen (1862) on squamous epithelia. We build on this idea to construct our epithelial tissue from cells. In an epithelial tissue, cell-cell adhesion is mediated by cadherin and cell-matrix adhesion is mediated by integrins [Davies and Garrod, 1997]. The strength of bonds formed by integrins between cell and extra cellular matrix varies depending on the type of integrins involved [Sheppard, 1996]. For simplicity, in our model, cell-cell and cell-matrix adhesion are not distinguished.

We assume that the bond strengths/adhesive strength are defined on the basis of selective cell adhesion. By selective adhesion we mean that cells of the same type prefer to be together. The adhesive energy between cells of the same type is small in comparison to the adhesive energy between cells of different types. The input values in the model for cell-cell and cell-matrix adhesive strength values are arbitrary and chosen on the basis of selective adhesion (discussed in a later chapter). For example, cell-cell adhesion strength between stem and dead cell in the model could be large in comparison to the adhesive strength between stem and TA cells. Table 5.1 shows an example input of adhesive strength parameters.

5.2.2 Cell Growth

Creatures grow in a randomly chosen direction and are limited to automata on their membrane. Since our model is based on minimising the global energy of a creature, prior to allowing the cell to grow, we compute the two energy configurations. Energy configuration h_1 refers to the current energy state while h_2 refers to the energy state a creature would be in if the creature grew in the randomly chosen direction. We compute energy using Equation 5.1. When computing h_1, the size of two creatures will be the same, but in h_2, one creature grows in size and the other reduces in size by one automaton, in order to accommodate the change. We define $\Delta h = h_2 - h_1$ and the cell is modelled to grow in the randomly chosen direction if the conditions in

Bond Name	Strength	Bond Name	Strength
Stem-Membrane	1	TA-Membrane	4
Inter-Membrane	16	Mature-Membrane	16
Dead-Membrane	16	Stem-Stem	4
TA-TA	4	Inter-Inter	4
Mature-Mature	4	Dead-Dead	4
Stem-TA	4	Stem-Inter	16
Stem-Mature	16	Stem-Dead	16
TA-Inter	6	TA-Mature	16
TA-Dead	16	Inter-Mature	6
Inter-Dead	16	Mature-Dead	6

Table 5.1: Adhesive bond strength between Cell-Cell, Cell-Matrix and Cell-Medium.

Equation 5.2 are satisfied. Figure 5.1 shows a schematic representation of cell growth in our simulation.

$$\begin{aligned} \triangle H \leq 0, \ \mathrm{P} &= 1 \\ \triangle H > 0, \ \mathrm{P} &= \exp - \left(\frac{\triangle H}{\kappa T} \right) \end{aligned} \tag{5.2}$$

where κ is the Boltzmann constant and T is the temperature in the system. The Boltzmann constant relates to the average kinetic energy of a molecule in an ideal gas to its absolute temperature. The Boltzmann constant is given as $1.3810^{-23} JK^{-1}$. The temperature (T), refers to the temperature in the system. When modelling biological cells, we assume that they are in a medium(eg: in some solution). Therefore temperature (T) in our model refers to the temperature of the medium they are in.

5.2.3 Cell Division

Cells undergo a programmed cell cycle before dividing [Murray and Hunt, 1993]. The length of the cycle varies according to cell type. During the cell cycle, cells grow in size and replicate their constituents prior to dividing. We limit cell division to cells

on the basement membrane. If the daughter cells are of the same type as the parent cell, we refer to it as symmetric cell division. If the daughter cells are of different types, this is known as asymmetric division. For example, a stem cell division which produces two stem cells or two TA cells is a symmetric division. A stem cell division which produces one stem cell and one TA cell is known as an asymmetric division.

Dover [Dover and Potten, 1988] found mean stem cell cycle time to be 16 hours in human keratinocytes. Therefore, stem cell cycle time in our simulation is set to 16 hours. There are no biological data at present to suggest a mean cell cycle time for TA cells. However, there are suggestions that TA cell cycle times are shorter in comparison to Stem cells in human keratinocytes. Stem and TA cell cycle time can be set independently in our model. The TA cell cycle is a user set value. The stem and TA cells in the model undergo cell mitosis at the end of their cell cycle. The number of allowed divisions for stem cells is not defined, since they are not modelled to differentiate. However in the case of TA cells, the number of times TA cells divide prior to differentiating is set initially by the user. Watt [Watt, 2001] suggested TA cells undergo 3 cell divisions prior to differentiation. The direction of cell division is chosen in a random fashion relative to the center of mass of the creature. All grid points one side of the division line are assigned to the parent cell and the grid points on the other side are assigned to the daughter cell.

5.2.4 Cell Differentiation

TA cells differentiate after dividing a number of times predetermined by a parameter set by the user or if they are not in contact with the basement membrane. The TA cells terminally differentiate into intermediate cells. The differentiation process from intermediate to mature and mature to dead cells depends on the cell age since becoming intermediate cells.

5.2.5 Modelling the Basement Membrane(BM)

We model the shape of the basement membrane as a straight line. However in real tissue, the basement membrane exhibits a wave like structure whose shape varies according to the type of epithelium. We assume the BM to be flat in our model until we understand the cause of the "wavy" patterns. We resist the temptation to make assumptions which produce desirable model outputs consistent with real tissue shape but which are not supported by biological principles.

5.2.6 Assumptions in the model

When modelling a complex biological system, it is inevitable that assumptions have to be made. However, these assumptions have to be justifiable under our current understanding of epithelial biology. For example, we have assumed that the direction of cell division is random. It is probably the case that epithelial cells do not divide in a random fashion but respond to chemical stimuli. We were unable to find any conclusive biological evidence to support such a notion. By modelling the direction of cell division as random we make sure there is no bias introduced into the model. We also assume that cell sorting is due to the DAH proposed by Steinberg. Other assumptions in our model include: only cells on the basement membrane are capable of division; the direction of cell growth is random; and, the basement membrane is flat.

5.2.7 Constructing the model

Our software is implemented on the Microsoft Windows platform using the C++ programming language. The model implementation is divided into two parts, the physical and the biological. During the physical iteration we randomly select a grid point and decide if the creature which includes that grid point is allowed to grow(see section 5.2.2). If the creature attempts to grow outside a predefined grid boundary, it is placed on the opposite side as if the cells were on a cylinder. It is essential that

we relate the time step of the Glazier-Graner model to the simulated time of cell behavior such as mitosis. It is important to avoid updating cellular functions while the system is momentarily away from equilibrium. This is more likely to happen when a creature has just divided and the two new cells have an unstable geometry. Savill [Savill and Sherratt, 2003] found a time scale of about 50 iterations for updating cell behavior was effective. Hence in our model, the physical iteration is repeated 50 times before iterating the biological iteration. The biological iteration consists of cell differentiation and cell division.

5.3 Results

We use our model to simulate three different experiments and the results are presented in this section. The experiments include varying differentiation ages for intermediate and mature cells and also for varying the temperature in the medium. We simulate these experiments in order to see if the model produces expected results. For example, we expect the differentiated layers to get thicker as we increase the differentiation age, and creatures to break up as temperature in the medium increases [Graner and Glazier, 1992]. We defer simulating novel biological experiments using this model to a later chapter. All input parameters chosen for the experiments are given in Table 5.2. The colour scheme used in our simulation to denote different type of creatures is shown in Figure 5.2. We use randomly generated shades of one colour to denote different creatures of the same type. For example, we use green to denote a TA creature, and we randomly generate different shades of green to denote different TA creatures. In Figure 5.3 and 5.4, we show typical outputs of the model using parameter values given in Table 5.2 for Experiment 1, where the model was initialised with single and multiple stem cells respectively.

In the temperature experiment, we vary the temperature parameter in our model to find the critical temperature above which cells start to break up. As can be seen in Figure 5.5, creatures start breaking up when T is chosen to be more than 5.

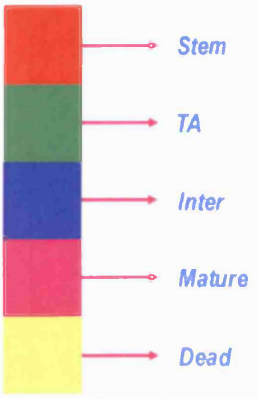


Figure 5.2: Colour key for simulations outputs from our model

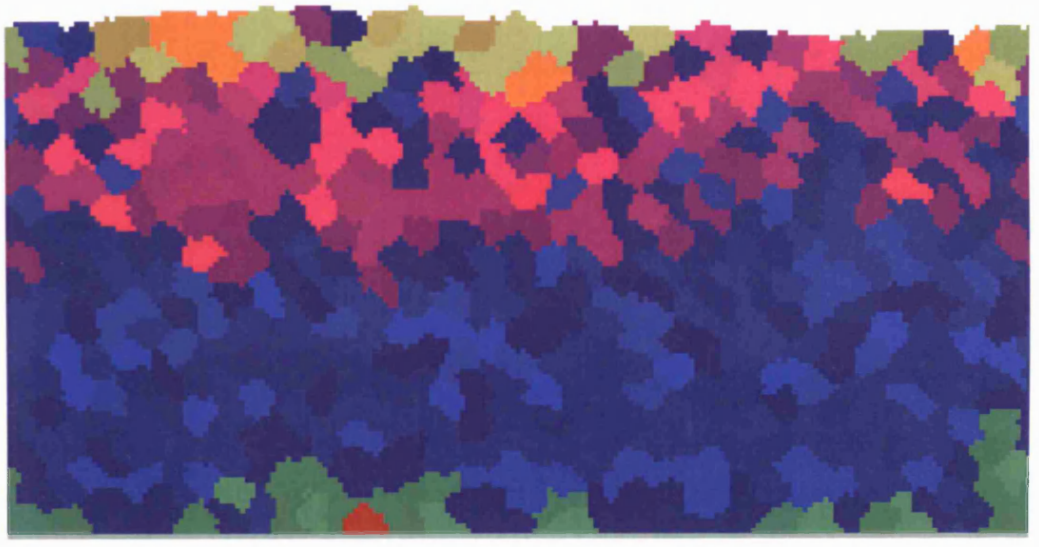


Figure 5.3: A typical output from the model, where input parameters as chosen are given in Experiment 1, Table 5.2. The model is initialised with a single stem cell.

Experiment	Rigidity(λ)	Area	Temp	Inter Age	Mature Age	Stem Cycle Time(h)	TA Cycle Time(h)
Experiment1	3	50	2	50	100	16	8
Experiment1*a*	3	50	**5**	200	300	16	8
Experiment1*b*	3	50	**10**	200	300	16	8
Experiment1*c*	3	50	**15**	200	300	16	8
Experiment1*d*	3	50	**20**	200	300	16	8
Experiment2*a*	3	50	2	**50**	300	16	8
Experiment2*b*	3	50	2	**100**	300	16	8
Experiment2*c*	3	50	2	**150**	300	16	8
Experiment2*d*	3	50	2	**300**	300	16	8
Experiment3*a*	3	50	2	200	**50**	16	8
Experiment3*b*	3	50	2	200	**100**	16	8
Experiment3*c*	3	50	2	200	**200**	16	8
Experiment3*d*	3	50	2	200	**300**	16	8

Table 5.2: Input parameters for experiments carried out using the model. Inter Age column refers to the cell age for differentiating from Inter to Mature. Mature age refers to the cell age for differentiating from Mature to Dead cell. Stem and TA cell cycle time refers to how long the cell cycles before going through mitosis and the time is given in hours.

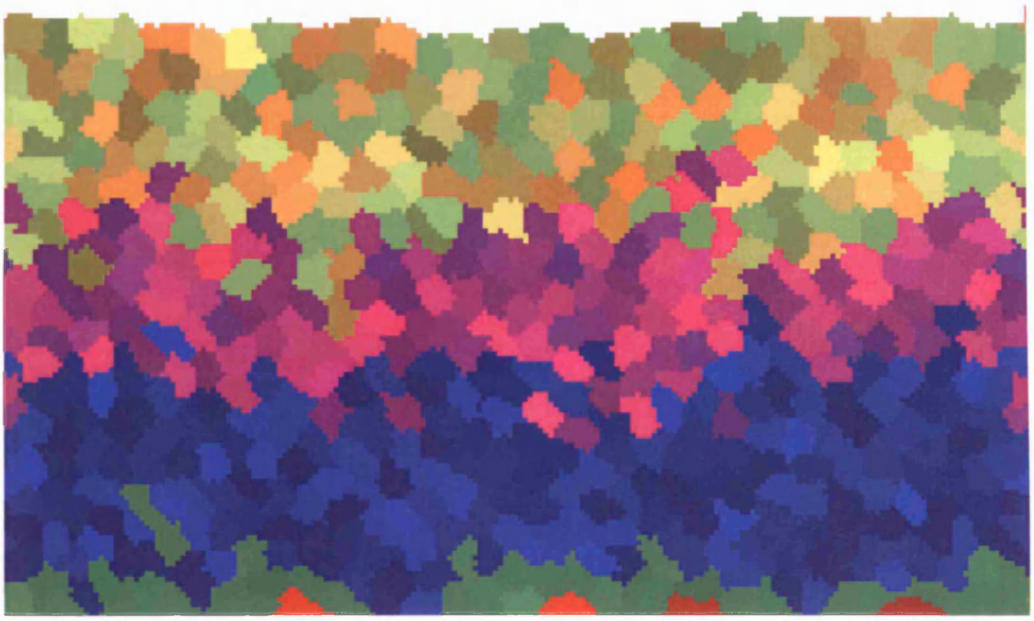


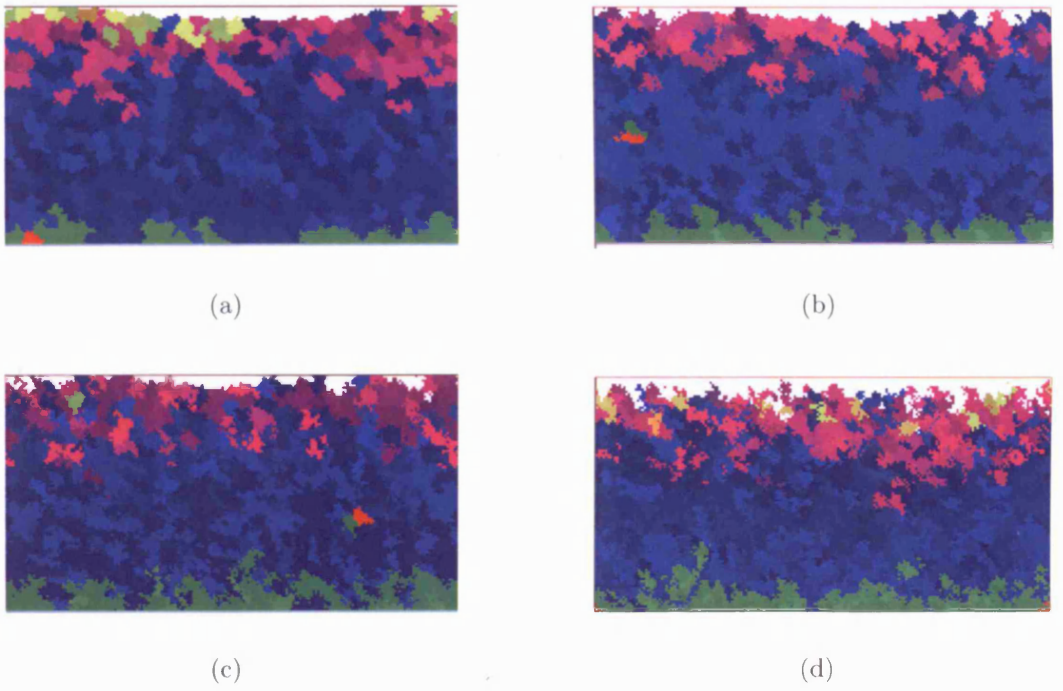
Figure 5.4: A typical output from the model initialised with multiple stem cells. The input parameters chosen are given in Experiment 1, Table 5.2

In the Inter to Mature experiment (I $\Rightarrow$ M), we vary the parameter which controls the aging of inter to mature cell differentiation. As can be seen in Experiment 5.6, the larger the parameter, the thicker an intermediate cell layer becomes.

In the Mature to Dead experiment (M $\Rightarrow$ D), we vary the parameter which controls the differentiation of mature to dead cell. As can be seen in Experiment 5.7, the larger the parameter, the thicker the mature cell layer becomes.

5.3.1 Determining the homogeneity of the simulation output

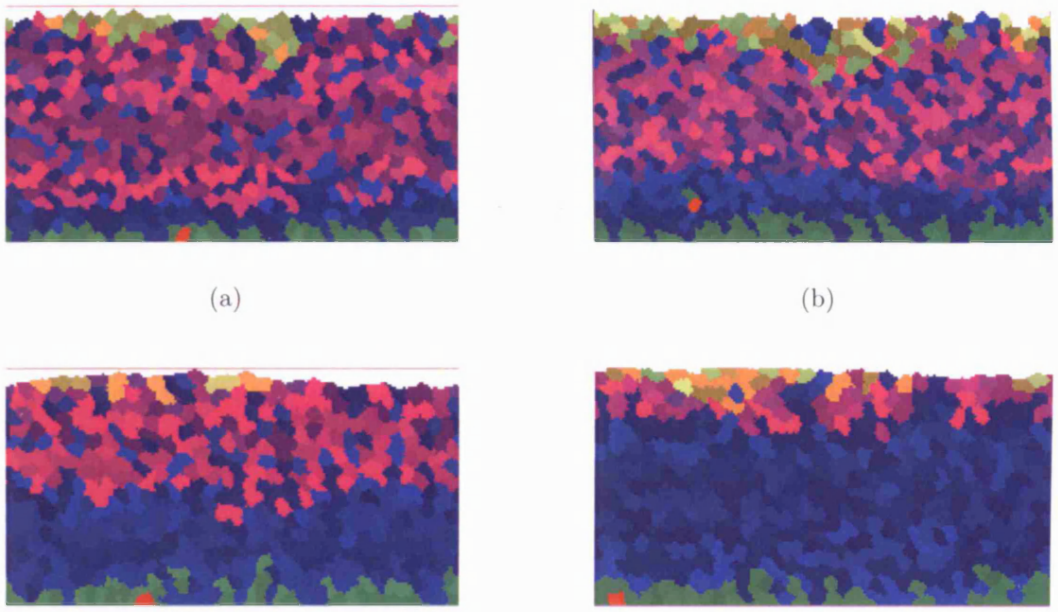
We apply the connected component algorithm suggested earlier in Chapter 3 to the simulation output from our model of epithelial cell interactions. We do this in order to evaluate if our method can be extended to a different system to analyse the homogeneity of the structure. The result is shown in Figure 5.8. As can be seen in Figure 5.8, the algorithm identifies 11 clusters of cells of different types. However, if



(a) (b)

(c) (d)

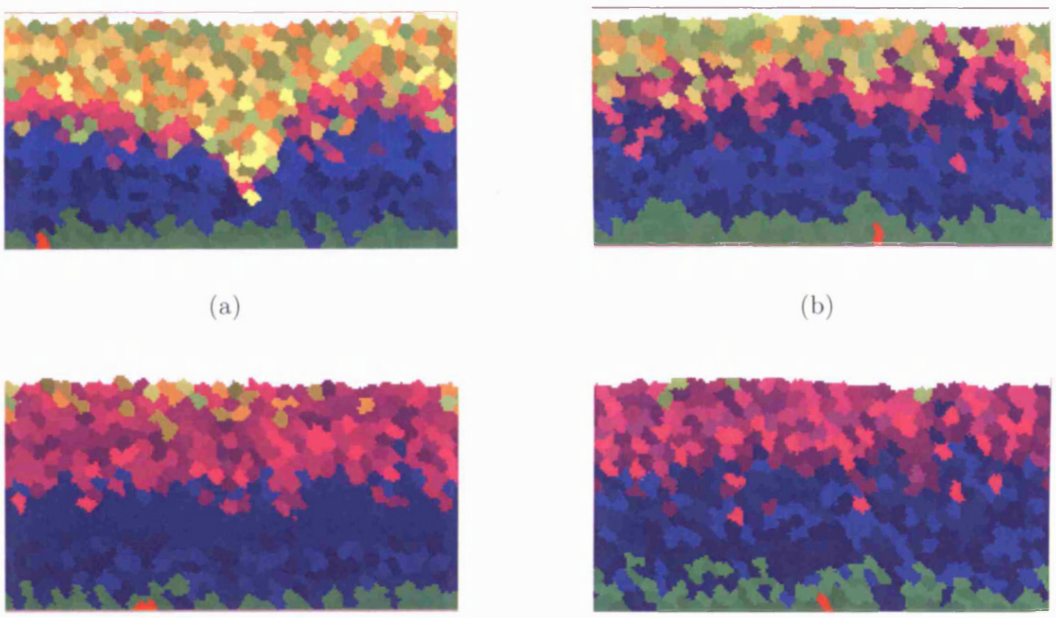
Figure 5.5: In this Experiment we vary the temperature of the model. (a)Experiment $1a$, temperature = 5. (b)Experiment $1b$, temperature = 10. (c)Experiment $1c$, temperature = 15. (d)Experiment $1d$, temperature = 20.



(a) (b)

(c) (d)

Figure 5.6: In this experiment, we alter the differentiation age, which models the cells differentiating from intermediate to mature cells. We expect the intermediate cell layer to get thicker as the differentiation age increases. (a)Experiment 2*a* where I $\Rightarrow$ M differentiation age is set to 50. (b) Experiment 2*b* where I $\Rightarrow$ M age is set to 100. (c) Experiment 2*c* where I $\Rightarrow$ M age is set to 150. (d) Experiment 2*d* where I $\Rightarrow$ M age is set to 300.



(a) (b)

(c) (d)

Figure 5.7: In this experiment, we alter the differentiation age, which models the cells differentiating from mature to dead cells. We expect the mature cell layer to get thicker as the differentiation age increases. (a)Experiment 3*a* where M $\Rightarrow$ D age is set to 50. (b) Experiment 3*b* where M $\Rightarrow$ D age is set to 100. (c) Experiment 3*c* where M $\Rightarrow$ D age is set to 200. (d) Experiment 3*d* where M $\Rightarrow$ D age is set to 300.

we look closely, since the stem cells are separated we require 2 different labels to identify them. Our system consists of 5 different cell types, hence we require a minimum of 5 labels to identify the clusters. However our simulation output requires 15 labels. This is because there is some mixing of cells of different types at the boundary. We mean by boundary, the border between different types of cells.

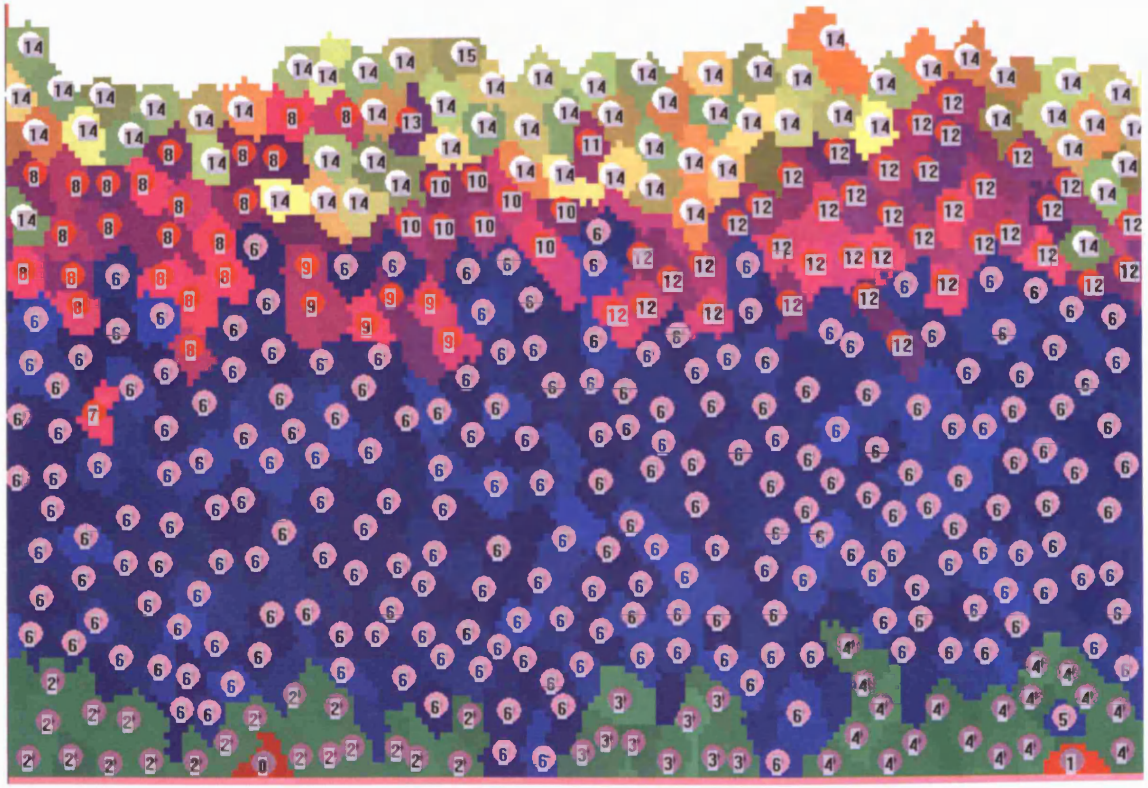


Figure 5.8: Cluster identification suggested in Chapter 3 is applied to an output from our model. As can be seen, the algorithm identifies 11 clusters of cells.

5.4 Discussion and Conclusions

In this chapter, we proposed a cellular automaton approach for modelling cell-cell and cell-matrix interactions in epithelium. Our model simulates biological processes such as division, differentiation and adhesion. We model cell sorting on the basis of the Differential Adhesion Hypothesis(DAH) proposed by Steinberg [Steinberg, 1963] using the Glazier-Graner model. Model simulations were initialised using two different configurations, namely a single cell and a line of cells. Our simulation rules and methods are derived from the biological literature and where possible we cite references to support them.

We believe our model simulates the collective behavior of cells in the epithelium. We do not model cell shape explicitly in our simulation, as we allow cells to remodel their cell membranes and reach a globally minimum energy state. Our model is stochastic in nature. The directions of growth and cell division are chosen at random. It is probable that epithelial cells do not behave in such a random fashion but respond to chemical stimuli. However, we failed to find any conclusive biological evidence for this in the case of epithelial cells. The simple question of how cells know in which direction to divide during development is not yet fully explained although there is evidence from mouse embryo studies which suggest cell polarity may play a role [Magdalena, 2002]. We know through observation that healthy epithelial cells always divide away from the basement membrane. Therefore, we model the direction of cell division as random but assume that cells cannot grow below the basement membrane.

In total our model has 26 input parameters, 20 of which define the bond strengths between creatures. Adhesive bond strength does have some biological basis. We know that adhesion between stem cells and TA cells is stronger than adhesion between TA cells and dead cells. Therefore, we can define the relative size of adhesive bond strength between cell-cell and cell-matrix to reflect this. The other 6 input parameters

(shown in Table 5.2) include cell cycle time, temperature, rigidity, target area and age to differentiation. The cell cycle time is obtained from the published literature. We leave the exploration of different parameters for adhesive bond strengths to a later chapter. Altering the target area allows the creatures to become bigger or smaller.

Higher differentiation age results in thicker intermediate and Mature cell layers, since they will have more "time" to be intermediate or mature cells before differentiating. We have shown in Figure 5.6 and Figure 5.7 several input parameters for time of differentiation of cells. As can be seen in Figure 5.6 and Figure 5.7, bigger values of input parameters results in thicker cell layers. The temperature (T) simulates the temperature of the medium that the cells are in. If T is large then all creatures become disassociated [Glazier and Graner, 1993]. We demonstrate this by running simulations for 4 different temperature values. As can be seen in Figure 5.5, for larger values of temperature, the creatures start to break up. However, in our opinion temperature does not reflect the medium that the cells are in. The temperature parameter plays a role in determining the probability of cell growth. If T becomes large, which in turn increases the probability of growth according to Equation 5.2, this causes the cells to break up in the model.

Our model can simulate the behavior of epithelial cells starting from a single cell as well as from a line of multiple cells. We have suggested a model for epithelial cell-cell interaction. Our rules of interaction are derived from the published epithelial cell biology literature. There are many ways of building on this model, but we believe this as a suitable juncture to end developing the model and observe the kinds of experiments this model can support before extending it further.

Chapter 6

Cell Adhesion and Tissue Formation

Stratified squamous epithelium, may be *non-keratinized* or *keratinized. Non-keratinizing* squamous epithelium covers wet surfaces, for example much of the oral cavity. The *keratinized* layer is formed when squamous epithelium covers dry surfaces such as the epidermal covering of the skin. The structural organisation of both epithelia are similar [Wright and Alison, 1984]. Squamous epithelium is divided into several zones, namely the *basal layer*, *malpighian layer*, *granular layer* and *keratin layer.* The basal layer cells are regarded as forming the proliferative compartment of the epidermis from which cells migrate and differentiate, as they progress to be desquamated at the surface. We focus on how such a structure is formed. We postulate that differential adhesion plays a role in the ordering of cells in the epithelium. We test this using our squamous epithelium model and conclude that differential adhesion aids the development of an ordered structure in squamous epithelium.

6.1 Cell Adhesion

Cell adhesion is necessary for the assembly of individual cells into tissues [Gumbiner, 1996]. A variety of cell adhesion mechanisms are responsible for assembling cells together and for determining the overall architecture of the tissue. Adhesion also aids signalling between cells [Braga and Harwood, 2001]. Although adhesive contacts between cells are stable, they should not be regarded as static, since maintenance of stable connections requires active cellular processes. Cell adhesion molecules are mostly transmembrane molecules, and include members of the integrin, cadherin, immunoglobin and proteoglycan families. While integrins and proteoglycans can execute cell-matrix as well as cell-cell interactions, adhesion via cadherins is mostly intercellular [Vleminckx and Kemler, 1999] since cadherins are usually homophilic[1] and not secreted. Cadherin molecules also play a role in cell differentiation, proliferation, death and polarization [Vleminckx and Kemler, 1999].

Differential cell-cell adhesion has long been suggested as a mechanism for cell sorting in developing tissues [Steinberg, 1963, Steinberg and Takeichi, 1994]. This may be due to cell surface adhesion molecules with different affinities. Evidence that cell sorting is achieved in developing tissues through differential adhesion has been rare. Since single cells attach to each other to form a tissue structure, there have to be processes which cause cells to do so. Basler found evidence of differential adhesion in the development of *Drosophila* wing imaginal disc. The analysis of clones in several compartments of the *Drosophila* wing imaginal disc suggests the presence of adhesion between the cells of different compartments [Dahmann and Basler, 1999].

In this chapter, we aim to understand the driving biological processes behind the organized structure in epithelium. During the development of epithelium, cells divide and attach to other cells or Extra Cellular Matrix(ECM) to form tissues. In the case of epithelium, the resulting tissue is a highly organized structure of cells forming

[1]The attachment of an adhesion molecule in one cell to an identical molecule in an adjacent cell.

layers of different types of cells. We refer to a layer consisting of basal cells as the basal layer, a layer comprising Intermediate cells as the Intermediate layer and so on. There must exist a biological process or biological processes which drives the global emergence of cellular layers resulting from local cell-cell and cell-ECM interactions.

6.2 Methods

Our model is based on the Graner and Glazier framework approach to differential cell adhesion and cell sorting [Graner and Glazier, 1992, Glazier and Graner, 1993]. Each cell in the model is assigned properties such as size, age, type, cell-cycle time, differentiation age and adhesion strength to neighbouring cells. Biological processes like cell division, growth, differentiation, and cell death are also modelled. A detailed description of the model can be found in Chapter 5. The Glazier-Graner model has been extended to study morphogenesis in *Dictyostelium* [Savill and Hogeweg, 1997, Mareé, 2000], avascular tumour growth [Stott et al., 1999], control of stem cell clusters by Notch-mediated lateral induction [Savill and Sherratt, 2003] and migrating fronts of cancer cells [Turner and Sherratt, 2002].

We represent cells on a square lattice. Each cell is assigned a type e.g: stem, TA, inter, mature and dead. Adhesive strengths between cells are modelled as surface energies between neighbouring cell membranes. The cells in the model sort themselves by minimising their surface energy under the constraint of maintaining a target area. Cell growth is modelled by increasing the target area of the cell at every iteration. Cell differentiation is modelled using cell age. A cell differentiates once it is older than a predefined threshold. Cell desquamation occurs when cells reach the upper boundary layer of the tissue.

We set out to test 3 different scenarios using the model for squamous epithelium. In the first scenario, we postulate that for an ordered structure of cells to exist, (where

intermediate cells lie directly above stem and TA cells, mature cells to be above intermediate cells and dead cells above mature cells), adhesive bond strength between neighbouring cells must be small. In other words, the bond strength between intermediate and mature cells must be smaller in comparison to bond strength between stem and mature cells. In our model, only cells in the basal layer are modelled to proliferate. When cells leave the basal layer, they differentiate to intermediate cells. The intermediate cell layer is directly above the basal layer. The intermediate cells differentiate to mature cells. The mature cell layer is above the intermediate layer. The mature cells differentiate to give rise to dead cells which are removed from the model when they reach the top. When we refer to one type of cell lying above another, we refer to our modelling frame work. In the case of biology, above and below don't have any relevance. For example we have a keratinized layer of epithelium on the surface of skin as well as the layer below and the interactions of these cells are no different. In order to model structure with the compartments of layers mentioned above, we set up the inequalities shown in Equation 6.1. We abbreviate stem, TA, intermediate, mature and dead cells with S, T, I, M and D respectively. The differential adhesion between stem and TA cells is denoted by Γ_{S-T}, stem and mature cells as Γ_{S-M} and so on.

$$\begin{aligned}
\Gamma_{S-S}, \Gamma_{T-T}, \Gamma_{S-T} &< \Gamma_{S-M}, \Gamma_{S-I}, \Gamma_{S-D}, \Gamma_{T-M}, \Gamma_{T-D} \\
\Gamma_{I-I}, \Gamma_{T-I} &< \Gamma_{S-I}, \Gamma_{T-D}, \Gamma_{T-M}, \Gamma_{S-M} \\
\Gamma_{M-M}, \Gamma_{I-M} &< \Gamma_{S-M}, \Gamma_{T-M}, \Gamma_{S-D}, \Gamma_{I-D}, \Gamma_{T-D} \\
\Gamma_{D-D}, \Gamma_{M-D} &< \Gamma_{S-D}, \Gamma_{T-D}, \Gamma_{I-D}
\end{aligned} \tag{6.1}$$

Our second scenario is to model the system with no differential adhesion between cells. In other words, we assume that there is no selective adhesion between cells and the surface adhesive energy between all cell types are the same. Inequalities defined to test this scenario are given in Equation 6.2.

$$
\begin{aligned}
\Gamma_{S-S}, \Gamma_{T-T}, \Gamma_{S-T} &= \Gamma_{S-M}, \Gamma_{S-I}, \Gamma_{S-D}, \Gamma_{T-M}, \Gamma_{T-D} \\
\Gamma_{I-I}, \Gamma_{T-I} &= \Gamma_{S-I}, \Gamma_{T-D}, \Gamma_{T-M}, \Gamma_{S-M} \\
\Gamma_{M-M}, \Gamma_{I-M} &= \Gamma_{S-M}, \Gamma_{T-M}, \Gamma_{S-D}, \Gamma_{I-D}, \Gamma_{T-D} \\
\Gamma_{D-D}, \Gamma_{M-D} &= \Gamma_{S-D}, \Gamma_{T-D}, \Gamma_{I-D}
\end{aligned} \tag{6.2}
$$

The final scenario is the reverse of the first in that we assume TA cells can be next to mature cells. When a dysplastic epithelium is observed through the microscope, it is possible to observe mature cells close to the basement membrane. Malignancy is associated with the disassembly of cells that leave their parent tissues and metastasize or invade others. Malignant invasion may result from a decrease in the cohesiveness of the invading cells and/or from an increase in their adhesiveness to extracellular matrix or other components of the invaded tissue [Steinberg and Foty, 1997]. Therefore it is tempting to claim that when we modify the selective adhesion parameters, what we are modelling is dysplasia. However, we believe the processes involved in dysplasia are more complicated. Hence we leave this as future work.

6.3 Results

The input parameters used for the simulations are shown in Table 6.1. Our experiments were initialised with both a single and multiple stem cells. The model output for the 3 scenarios are shown in Figure 6.1 and Figure 6.2 for single and multiple stem cells, where the simulated epithelium exhibits an ordered structure of cells. This is similar in appearance to a histological section of the epithelium analysed using a microscope. We also provide movies constructed from our model outputs for scenario 1, 2 and 3 in the CD attached. The movies are named scenario1.avi, scenario2.avi and scenario3.avi. A description of this movies can be found in Appendix B. Figure 6.1(b), 6.2(b) shows model output obtained for the second scenario where we

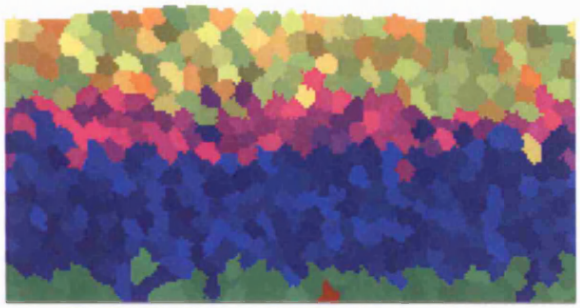
assume that there is no differential adhesion between cells, ie: the surface energy between different types of cells are the same. As can be seen in Figures 6.1(b), and 6.2(b) mature cells are close to the basement membrane which is not the case in normal epithelium and the Stem cells have moved away from the basement membrane. This is not similar to what is observed when a histological section of the epithelium is analysed using a microscope.

Parameter	Value
Rigidity	3
Target Area	45
Temperature	2
Inter Age	200
Mature Age	300
Stem Cycle Time(in Hours)	16
TA Cycle Time(in Hours)	8

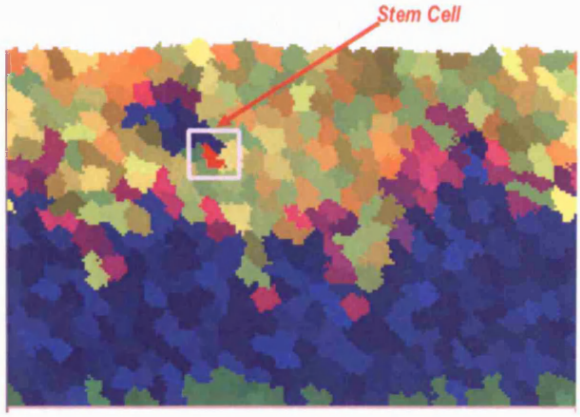
Table 6.1: Input parameters used, when testing the hypothesis using the epithelial model in Chapter 5

In order to test the third scenario we assume the adhesion energy between stem cells and dead cells is smaller in comparison to the adhesive energy between stem and TA cells. We make this assumption to see if this can aid the emergence of ordered structure seen in normal epithelium. The model outputs obtained using scenario 3 are shown in Figures 6.1(c), and 6.2(c).

To further our understanding of the selective adhesion between cells, we plot the age of a cell against the distance from the basement membrane. In a normal epithelium, young proliferating cells can be found close to the basement membrane and as cells age and differentiate they move away from the basement membrane. Figures 6.3(a), and 6.3(b) shows the results obtained, when the model was initialised with single as well as multiple stem cells. As can be seen, in scenario 1, young cells are found close to the membrane. However, in scenario 2 and 3, young cells are not close to the basement membrane.



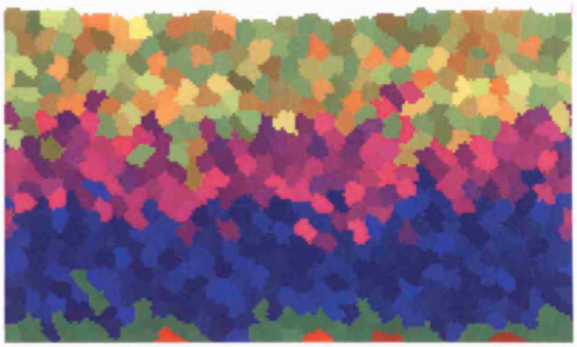
(a)



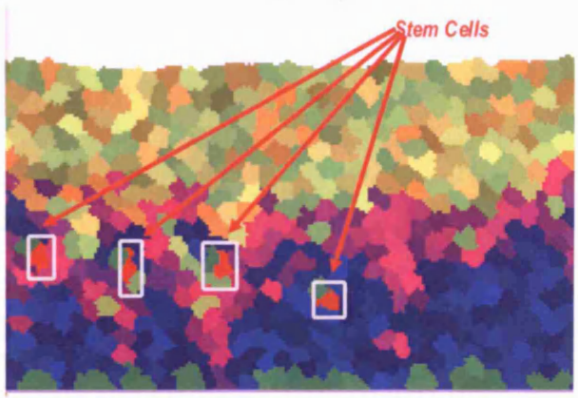
(b)

(c)

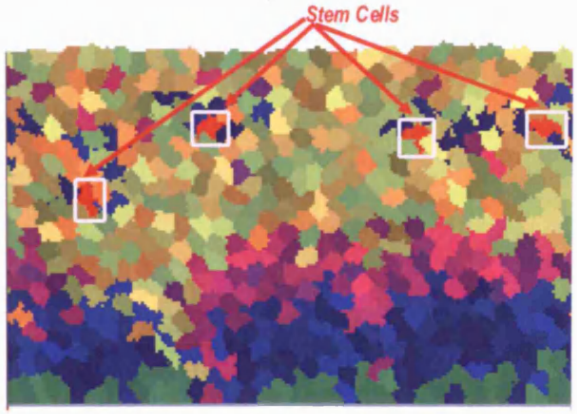
Figure 6.1: Simulations initialised with a single stem cell and the input parameters are given in Table 6.1. (a)As can be seen in Figure 6.1(a), which simulates scenario 1, an ordered structure of epithelium is observed. (b)Figure 6.1(b) shows the epithelial structure for scenario 2. As can be seen, the stem cell is no longer at the basement membrane and has moved up towards the top. Since the stem cell is no longer at the basement membrane, young proliferating cells, which are produced through the stem cell division are now at the top. (c)Scenario 3 result is given in Figure 6.1(c). The stem cell is at the top with dead cells and mature cells are close to the basement membrane.



(a)

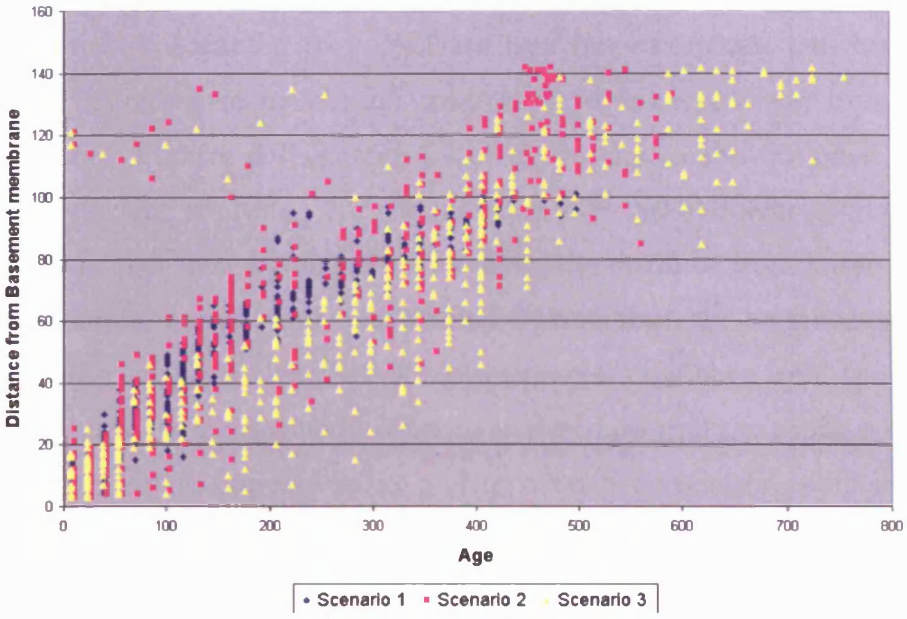


(b)

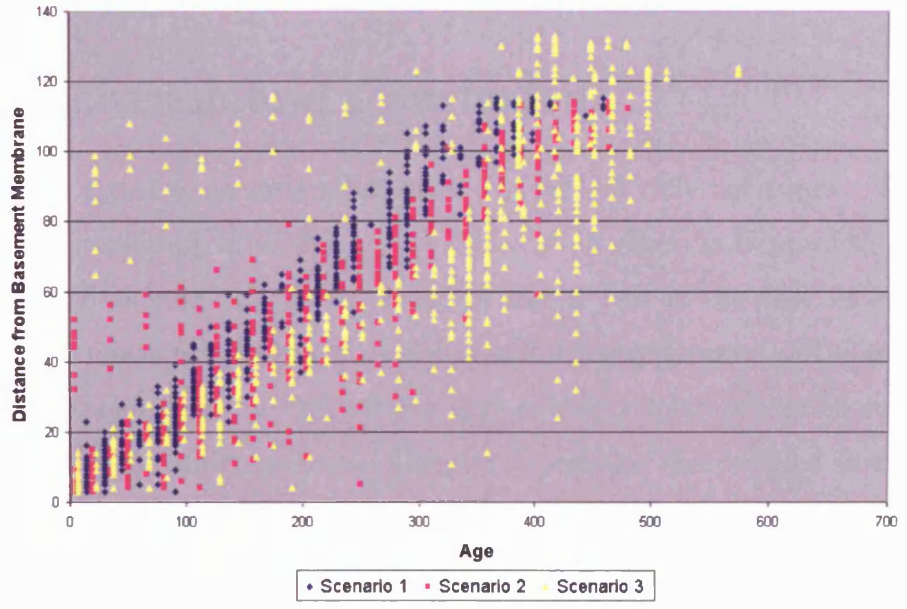


(c)

Figure 6.2: Simulations initialised with multiple stem cells and run using input parameters given in Table 6.1. (a)Figure 6.2(a) shows results obtained for scenario 1. As can be seen, the simulated epithelium exhibits an ordered structure. (b)Results obtained for scenario 2 are shown in Figure 6.2(b). As can be seen, stem cells are no longer at the basement membrane. (c)Scenario 3 is simulated and shown in Figure 6.2(c). As can be seen, the epithelium is no longer ordered.



(a)



(b)

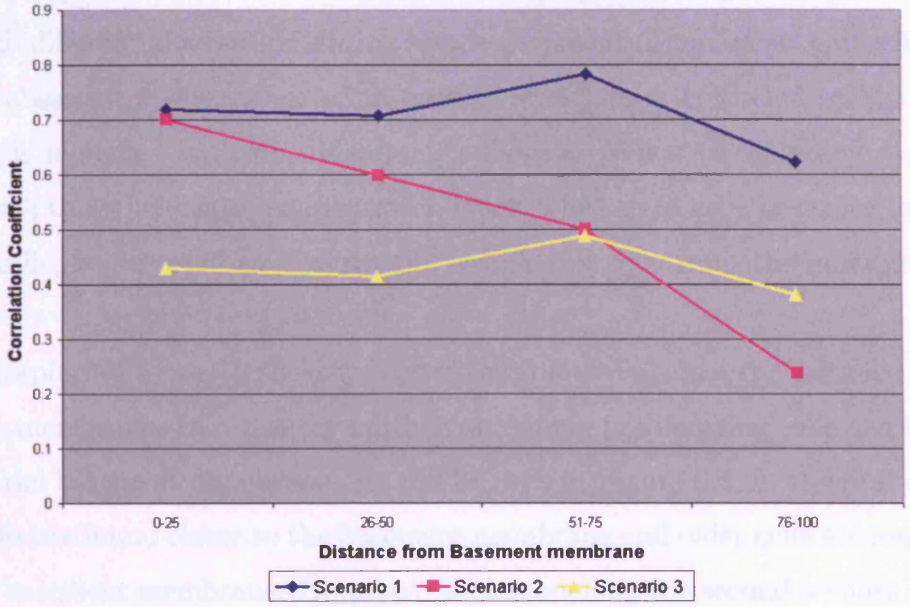
Figure 6.3: We plot the age of a cell against the distance from the basement membrane. Graphs show the results obtained for three different scenarios. (a)Experiment initialised with single stem cell. (b) Experiment initialised with multiple stem cells.

We compute the correlation coefficient between age and distance from the basement membrane. We wanted to investigate how the correlation between age and distance from the basement membrane varied as the cells moved away from the basement membrane. Figures 6.4(a), and 6.4(b) shows the results obtained using the correlation coefficient analysis. We compute the correlation coefficient by dividing the whole tissue area into 4 compartments using the distance from basement membrane. For example, we select all cells within 50 pixels from the basement membrane and compute the correlation coefficient between a cell's distance from the basement membrane and its age. We do this in order to analyse how well the tissue is structured within compartments as opposed to computing a single correlation coefficient for the whole tissue. We expect to find a higher correlation coefficient when the tissue is structured as in the case of normal epithelium. As can be seen from Figures 6.4(a), and 6.4(b), scenario 1 yields a higher correlation between a cell's age and its distance from the basement membrane in comparison to scenario 2 and 3.

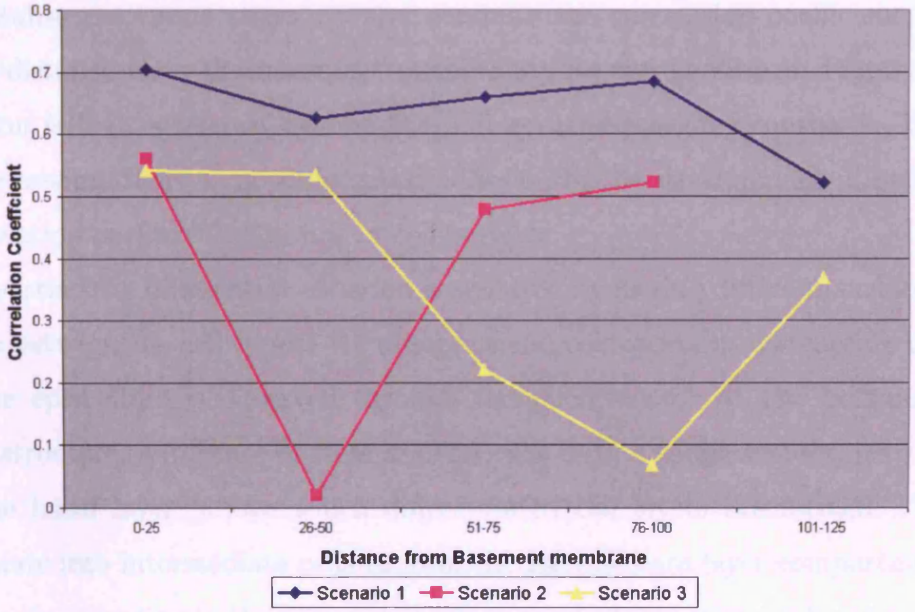
6.4 Discussion and Conclusion

The epithelial tissue is an ordered structure of cells of different types. We suggest that this structure may arise due to the differential adhesion hypothesis suggested by Steinberg [Steinberg, 1963]. In order to test if this is the case in epithelium, we extend the Glazier and Graner model of cell sorting [Graner and Glazier, 1992] by using their algorithm for cell sorting and model cellular processes such as cell division, growth and differentiation. Chapter 5 provides the detailed description of our model of squamous epithelium. The experiments were initialised with single as well as multiple stem cells. The model outputs are shown in Figures 6.1, and 6.2.

As can be seen in Figures 6.1(b), 6.2(b), 6.1(c) and 6.2(c), the stem cells are not attached to the basement membrane. This is due to the strength of the adhesive energy bonds chosen between the stem cells and the basement membrane.



(a)



(b)

Figure 6.4: We plot the correlation between cell age and their distance from the basement membrane. The distance from basement membrane is divided into compartments. The results were obtained using the input parameters shown in Table 6.1. (a)Experiment is initialised with single stem cell. (b) Experiment is initialised with multiple stem cells.

Figures 6.1(b), and 6.2(b) show the simulation output for which we postulate that there is no differential adhesion during the development of squamous epithelium. We failed to observe the emergence of an ordered structure of epithelium when the cell adhesion is modelled with no differential adhesion. When the adhesive bonds are chosen such that the energy configuration is low when stem cells are close to mature or dead cells, we also fail to observe the ordering of cells from the emerging tissue structure.

The graphs in Figure 6.3 show the age of cells plotted against the distance from the basement membrane. In a healthy epithelium, young proliferating cells can be found closer to the basement membrane. As can be seen in Figure 6.3, in the first scenario, young cells are found closer to the basement membrane and older cells are found away from the basement membrane. However, when modelling the second scenario, we find old cells close to the basement membrane. In order to further our understanding of correlation between cell age and distance from the basement membrane, we compartmentalise the tissue structure and compute the correlation coefficient between age and distance from the basement membrane. As can be seen in Figure 6.4, the correlation is high when we assume the differential adhesion hypothesis. However, when we assume there is no differential adhesion in the development of epithelium, the correlation coefficients are low in comparison.

Cell sorting via differential adhesion is achieved by having different surface energy densities between the cell types. We observe a compartmentalized structure of tissue, when the epithelium is observed through the microscope. At the bottom of this ordered structure, a mixture of stem and TA cells form a compartment, here referred to as the basal layer, above which differentiated cell layers are formed. TA cells differentiate into intermediate cells to form the intermediate layer compartment. For this to occur according to the differential adhesion, the bond strength between TA and Inter cells must be smaller in comparison to other differentiated layers. Intermediate cells differentiate to form the mature cell layers and mature cells differentiate to form

the dead cell layers.

We chose input parameters to satisfy Inequalities 6.1 for the strength of adhesion between cells and showed differential adhesion may play a part in the development of epithelium for a set of input parameters. There are many combinations of input parameters which could be used in the simulation to get an ordered structure and the parameter ranges we have chosen are not special, just one possible combination which could satisfy the inequalities 6.1. However, if DAH is indeed the driving force behind the ordered structure in epithelium, and is modelled using Glazier-Graner's cell sorting algorithm, then ordered structures may be shown to arise if the inequalities 6.1 are satisfied. Since there are many parameters in the model, it is difficult to state categorically that no other combination gives rise to ordered structure even when there is no differential adhesion. This statement can only be made once we have explored all possible ranges of input parameters and this may be a criticism of the model.

It is interesting to note that for an ordered structure to stay in dynamic equilibrium, cell division must also play a role. We initialised our model with a developed epithelium and ran the model by turning off the cell division process with adhesion input parameters satisfying scenario 1. Since no division was taking place at the basement membrane, the intermediate cells directly above it were simply aging and not getting pushed up by new cells. As the intermediate cells aged, they became mature and dead cells. Hence we found mature and dead cells close to the basement membrane, even though our input parameters for adhesion satisfied inequalities. Therefore, we conclude that pressure from cell division aids the maintenance of the ordered structure of epithelium. This behaviour was also observed during experiments where cell division was turned off in a biology experiment [2].

We have shown that the cell organisation in the epithelium may be due to DAH.

[2]Private communication with Helen McNeil at Cancer Research UK on 16th June 2004. Fiona Watt's lab, also at Cancer Resrach UK, carried out experiments where they turned off cell division and found the layers to collapse on top of each other

However, the interesting question is whether this statement can be supported by the biology. The biologists are aware that compartment boundaries play an important role in pattern formation. However, it is unclear how these boundaries/compartments are established during the development of epithelium. Biological experiments supporting adhesion aided cell sorting have been completed for the development of the embryonic central nervous system and for wing development [Dahmann and Basler, 1999], in Drosophila.

We have shown that in an *in silico* model, the structural organisation of epithelium may arise from the differential adhesion proposed by Steinberg. Our model does not include cell polarity nor cell signalling. There is evidence to suggest that adhesion molecules (eg: cadherins) not only play an adhesive role, but also a signalling role [Vleminckx and Kemler, 1999]. It would also be of great interest to see whether compartment boundaries arise *in vivo*, because cells from adjoining compartments differ in the activity of a single kind of cell adhesion molecule. The cell adhesion molecules are essential to maintain a stable tissue structure and the expression of cell adhesion molecules are tightly regulated to control cell proliferation, mobility, differentiation and survival, and many of these processes are misregulated in malignant tumors [Thomas and Speight, 2001]. It will be interesting to develop the model further to introduce several cell adhesion molecules into the model, but we leave this for discussion in a later chapter.

Chapter 7

Conclusions and Future Work

We set out to construct an *in silico* model of epithelial cell-cell interactions. Any such model would have to be compared with *in vivo* or *in vitro* experiments to determine its validity. Once validated, the model could be used to test novel biological hypotheses before expensive and time consuming laboratory research was undertaken. In this chapter we report on the conclusions arising from our attempts to construct and evaluate an epithelial cell-cell interaction model.

In Chapter 3, we produced a rational reconstruction of Stekel's model of epithelium [Stekel et al., 1995] in order to familiarise ourselves with agent based modelling techniques. Several aspects of the model were shown to be inadequate or at least without strong evidence in the biological literature. These areas included the direction of cell division, the determination of neighbour cells, the modelling of compressive forces and cell shape. In particular, Stekel's suggested connective tissue attractive and repulsive forces appear to have no biological basis. Such assumptions can be made to produce desirable graphical output from the model, but do not enhance our current understanding of biological processes. We present a method to analyse spatial clusters and the method is applied to Stekel's model output. The method uses techniques from computational geometry and image processing.

Chapter 4, considered ways of using histological sections of oral epithelium to validate *in silico* model outputs. We proposed a method which combines Fourier Descriptors, Principal Component Analysis (PCA) and classification techniques to distinguish labelled classes of tissue images using the curvilinear shape of features found in the image set. We demonstrated this combination of techniques on histological images of normal and dysplastic tissue samples taken from two sites of the oral mucosa, the tongue and inner cheek (buccal surface). The Fourier descriptor shape model for normal and dysplastic epithelium may be compared with *in silico* outputs and used to evaluate the model hypothesis with increased confidence.

In Chapter 5, we proposed a model to simulate epithelial cell-cell interactions. We used the Glazier-Graner algorithm for cell sorting and added biological processes such as cell division, differentiation and adhesion. Our model simulates the behaviour of epithelial tissue starting from a single cell as well as a line of cells. Our rules of cell interaction are derived from the published epithelial cell biology literature. There are many ways of building on this model, but we believe this is a suitable juncture to end developing the model and observe the kinds of experiments the model can support before extending it further.

We evaluated our model further in Chapter 6, by testing it against an existing hypothesis from biology. It has long been postulated that differential adhesion is the driving process behind cell sorting and the formation of compartments during embryonic development [McNeill, 2000, Dahmann and Basler, 1999]. We test this using our model and conclude that by modelling differential adhesion and allowing cells to proliferate in the model it is possible to show the emergence of compartments in epithelium.

7.1 Future Work

We explore several avenues of future work in this section. Obvious extensions of the model include detailed simulation of biological processes such as cell signalling, differentiation, apoptosis and extension to three dimensions. We attempted to extend Stekel's model to three dimensions. Movies obtained from our 3-D simulation can be found in the attached CD rom and a description of them is provided in Appendix B. We provide 2 movies, namely stekel3D and rotate3D. The stekel3D provides a simulation movie for Stekel's model in 3-D. Rotate3D shows the same simulation by rotating the viewing angle. It should be noted that our reconstruction in 3-D did not contain all elements of the Stekel's model. This was performed as an exploratory exercise to extend the model into 3-D.

It is important to keep in mind that cellular processes such as differentiation and signalling are active research areas in cell biology. Hence we have to strike a balance between modelling a biological process and including sufficient, well justified biological information into the model to simulate it. For example, the cell cycle may be simulated by implementing a detailed model like Novak's [Novak et al., 1998]. We discuss some biological processes here which we have omitted so far, but which could improve our model further. We also discuss some possible extensions to our image analysis work.

7.1.1 Cell Polarity

The plasma membrane of epithelial cells is typically divided into two domains: an apical surface and a basolateral surface. This polarized organisation is the basis for the function of these cells in the transport of ions across the epithelium. The epithelial cells rely on two fundamentally different mechanisms to ensure the localisation of plasma membrane proteins: the selective targeting of a protein to a specific cell-surface domain and the selective stabilisation of a protein at a specific cell-surface

domain [Yeaman et al., 1999]. Epithelial cell-cell adhesion is mediated principally by E-cadherin [Kemler, 1992] and it accumulates membrane proteins at sites of cell-cell contacts in polarized epithelial cells [McNeill et al., 1990]. Cell adhesion to the extra-cellular matrix is mediated by the integrin superfamily of adhesion receptors [Hynes, 1992] and these interactions generate differences in protein distributions which define the apical-basal axis of polarity.

We assume, in our model, that cell adhesion is uniform on cell membranes and we do not differentiate between apical and basolateral adhesive surfaces. The model may be improved by defining different adhesive strengths along the cell membrane as is the case between apical and basolateral cell membrane in epithelial cells. The apical and basolateral membranes may be computed using cell polarity, and the adhesive strength of the bonds can then be defined as a distribution of the polarity vector. It is interesting to note that by introducing the polarity vector of a cell when initialising the model, we no longer need to define the basement membrane. The *in silico* cell using its polarity can define its own basement membrane. By this we mean a cell can compute the automata directly below it using the direction of its polarity vector.

7.1.2 Cell Cycle

Every living organism is composed of one or more cells, and new cells can only arise by the division of pre-existing cells. Instructions governing cell division are encoded in the sequence of DNA in the chromosomes. Broadly speaking, rapidly dividing human cells have a cell cycle life time of 24 hours, although this may vary according to cell type. Cell cycle can be separated into two fundamental parts, namely Inter-phase and Mitosis.

During cell division, the cell must replicate all of its components and allocate them to two identical daughter cells. Since DNA stores the genetic information of the cell, it has particular importance among the cellular constituents. Hence, during

cell division DNA must be accurately replicated and chromosomes must be precisely segregated. The DNA replication phase of the cell cycle is called the **S Phase** and the phase of chromosomes segregation is called **Mitosis or M phase**. The gaps between DNA replication and Mitosis are denoted by G_1 and G_2 respectively.

Not all cells proliferate continuously. Cells which cease to proliferate, but retain the capacity to re-initiate progress through the cell cycle at a later time, are known to be in a **Quiescent state**. Quiescence can be induced in many types of cells by manipulation of their environment. This state represents an important level of biological control, since it indicates that cell multiplication may be held back, but can be re-initiated rapidly when required. For example in tissue generation following injury. Quiescent cells remain in G_1 phase of the cell cycle. Moreover, when cells exit quiescence, they remain in the G_1 phase of the cell cycle for some time, before they enter into the S Phase.

By following a population of cells over time, it can be seen that not only does it take longer for the population to double in size, but also that a significant fraction of cells fail to complete the cell cycle. These cells are known as **senescent**. In senescence, cells enter G_1 phase, but they never leave it to enter the S phase again. As time goes on, the amount of senescence in a cell population increases until there comes a point at which a plateau phase occurs and further increase in cell population is halted.

The total number of cells which can be created from a founder cell is termed the proliferative potential. The proliferative potential is highly variable between cell types. Hence it is possible to define a population of cells by its proliferative potential. The proliferative potential is a programmed feature of the cell, rather than a random loss of the capacity to divide. The cells in G_0 are activated to proliferate by growth factors, such as the Epidermal growth factor(EGF) and Insulin like growth factor(IGF), advancing to the G_1 after several hours of stimulation. If the growth factors are removed early in G_1, cells revert back to the quiescent state. Growth

factors act by binding extracellularly to their specific trans-membrane receptor proteins. Growth factors are a highly heterogeneous group of molecules,whose ability in specific circumstances is to regulate cell proliferation. It is a fundamental property of all growth factors that they act only upon a characteristic and defined set of target cell types.

There is no universal growth factor with identical action on all cells. Each growth factor has a designated biological specificity. Another important aspect of growth factors is that they act locally within tissue rather than systematically. Growth factors not only exhibit positive actions, such as inducing an event which would not occur otherwise, but also act as negative regulators. For example, the system that controls the multiplication of cell populations has both positive and negative regulators. The family of growth factors, TGF_β, amongst many biological activities, are examples of negative regulators of cell proliferation.

Our model may be extended to include a more detailed model of the cell cycle. Cells are currently modelled to divide after completing the appropriate cell cycle time. The model may be extended to include check points for cells going through their phases prior to dividing. The growth factors may be introduced into the model by incorporating a time dependent diffusion equation. This equation can be solved using the Finite Difference Method(FDM). The source for the growth factors may be placed along the basement membrane and different types of cells can be modelled to respond to different types of growth factors.

In the case of oral cancer, patients treated with chemotherapy or radiotherapy are often plagued with mucositis[1]. Severity depends on the quality of dental hygiene, the treatment schedule, the irradiated area and the amount of radiation given, as well as the age of the patient. Late effects can be characterized by thinning of the mucosa (the soft tissues of the oral cavity), and mucosal ulceration and necrosis. The radiation removes healthy cells as well as tumour cells and patients may experience discomfort

[1]Inflammation of the lining of the mouth

due to several layers of the epithelium being removed. However, growth factors may be introduced locally to improve the speed of the natural healing processes in epithelium. It may be possible to study the effects of locally introducing the growth factors to promote healthy cells to divide more frequently. It may be possible to devise better treatment protocols of induced growth using this extended model [Shochat et al., 1999]

7.1.3 Growth Control in Epithelium

The epithelium undergoes self renewal and the whole process is regulated to be in a dynamic equilibrium. Growth control in epithelia requires the cell population to be maintained at an appropriate level. The hypotheses may be based on a functional demand or that the population must be in some way genetically predetermined. However, any global theory describing the dynamic equilibrium state must be able to deal with local physiological adaptation to an induced environmental change and wound healing. One of the questions which needs to be answered is how the cell population "senses" its cell number, how it realises if a proliferative response is indicated, and how the population recognises that the proliferative rate should be curtailed.

The sensor for the population may be modelled at the cellular level and this would imply that the normal population size is genetically preprogrammed. This hypothesis would only work, if the sensor is placed in stem cells and TA cells, which gives rise to cell division. In the case of tissue removal, the stem cells somehow sense the deficit (this may be through signalling) and increase the proliferation until the predetermined population size is reached. However, it may be difficult to account for hyperplasia or increased food intake in which the actual size of the population increases. Alternatively, the population size may be determined by the functional demands made upon it [Goss, 1964]. Normal functions would impose a certain cell number; and if the functional demands are increased either by removal of cell tissue or

due to physiological adaptation as with increased food intake, increased growth due to functional demands would ensure that the demand has been met. Wright [Wright and Alison, 1984] provides a detailed explanation of growth control mechanisms in epithelium.

Our model may be extended to include the growth control mechanism of epithelium to keep the model in a dynamic equilibrium. When our model is initialised with a stem cell or a line of cells, we need to implement the first hypothesis using a "sense" mechanism to reach a dynamic equilibrium. However, if we wish to simulate oral carcinoma or wound healing, we will have to implement the functional demands hypothesis. It is our opinion that development and maintenance of the epithelium are two different biological processes and should be modelled separately. Hence, a "sense" mechanism could be implemented during the developmental phase and the functional demands hypothesis might be modelled to maintain the epithelium in a dynamic equilibrium.

7.1.4 Image Analysis

In this section, we discuss the possible extensions of our analysis of images of histological sections of the oral epithelium. The basement membrane in the images was extracted manually. Our attempts at extracting the basement membrane from the images using edge detection algorithms did not yield satisfactory results especially where there are sharp edges in the basement membrane, a common feature of dysplasia. This problem arises since prior to applying Canny [Canny, 1986] edge detection techniques we needed to blur the image with a Gaussian filter to smooth the noise. When this procedure is carried out, the sharp edges are smoothed and do not get picked up by the edge detection algorithm. Techniques from Active Shape Models (ASM) cannot be used, because it is not possible to place landmark points on the image since it is a continuous line. This is in contrast to a face, for example, where

landmark points can easily be identified as unique anatomical features on the nose, eyes and mouth. However, we believe snake algorithms [Trucco and Verri, 1998] may be applied to detect edges automatically.

We employed supervised learning techniques for classification. It may be possible to employ unsupervised learning techniques to classify images. We showed by using SVM techniques one is likely to get better accuracy, but there is nothing to suggest that our classification accuracy cannot be improved further by optimising other classification techniques to our data set. This may be an area for further work. We have not explored the possibility of using this to aid clinical diagnosis. We would require a larger data set and our classification model would require some attention since oral dysplasia is usually classified as normal, mild and severe. Hence the classification is now a three class problem as opposed to two. We believe our method could be extended to aid clinical diagnosis. However, more work needs to be done to test this.

Appendix A

Principal Component Analysis

Principal component analysis(PCA) is a statistical analysis method. PCA enables us to reduce the variation in a data set to a relatively small number of uncorrelated parameters. Our feature vector, x is constructed using the Fourier descriptors of a given image. For example, if the basement membrane shape is characterised by k Fourier descriptors then the feature vector $\mathbf{x}$ may be defined as follows

$$\mathbf{x} = (x_1, x_2, \ldots\ldots\ldots, x_k)^T \tag{A.1}$$

Given Equation A.1 and n examples in the data set, the mean vector $\bar{\mathbf{x}}$ may then be given as

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \tag{A.2}$$

where $\mathbf{x}_i$ is the feature vector of the i^{th} example in the data set.

We then compute the deviations $(\mathbf{x}_i - \bar{\mathbf{x}})$ from the mean and construct the scatter matrix (also known as the covariance matrix) $\mathbf{S}$, which is given by

$$\mathbf{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \tag{A.3}$$

and compute the eigenvalues of the matrix S.

The scatter matrix is positive semi-definite and hence each eigenvalue λ_i is non-negative. The eigenvalues are ordered and numbered as $\lambda_1, \lambda_2, \ldots\ldots, \lambda_k$ such that $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \lambda_k \geq 0$ where k is the size of the feature vector. If the corresponding eigenvectors are $\mathbf{p}_1, \mathbf{p}_2, \ldots\ldots, \mathbf{p}_k$ then

$$\mathbf{S}\mathbf{p}_j = \lambda_j \mathbf{p}_j, \quad \lambda_j \geq \lambda_{j+1} \quad and \quad \mathbf{p}_j^T \mathbf{p}_j = 1 \quad for\, j = 1, 2, \ldots\ldots, k \tag{A.4}$$

The eigenvectors form an n-dimensional spanning set of the feature space. Hence every point in this n-dimensional space can be computed from a sum of the mean and a linear combination of the eigenvector. The trace of the covariance matrix is equal to the sum of the squared differences, which represents the total variation of all the elements of the feature vector $\mathbf{x}$, over all the samples in the data set and is equal to the sum of all the eigenvalues. Eigenvalues describe the variance of each component of the feature vector. The eigenvector corresponding to the largest eigenvalue, describes the most significant mode of variation in the variables used to derive the covariance matrix. Discarding small eigenvalues and their corresponding eigenvectors, noise in the data set can be minimised. Most of the variation can usually be explained by a relatively small number of modes, t. One method for calculating t is to choose the smallest number of modes such that sum of their variances explains a sufficiently large proportion of λ_T, the total variance thus

$$\lambda_T = \sum_{j=1}^{k} \lambda_j \tag{A.5}$$

and t is chosen as the minimum value of g such that Equation A.6,

$$\left[\frac{\sum_{j=1}^{g} \lambda_j}{\lambda_T} \right] \geq f \tag{A.6}$$

where f is a predetermined valve in the range $[0, 1]$ that determines the accuracy of the statistical model. f is usually chosen to be between 0.95 and 0.99.

Any feature vector in the data set can be approximated using the mean and a weighted sum of the deviations obtained from the first t modes:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \tag{A.7}$$

where

$$\mathbf{P} = (\mathbf{p_1}, \mathbf{p_2}, \ldots, \ldots, \mathbf{p_t}) \tag{A.8}$$

is the matrix of the first t eigenvectors and

$$\mathbf{b} = (b_1, b_2, \ldots, \ldots, b_t)^T \tag{A.9}$$

is a vector of weights. It should be noted that by definition in Equation A.8, $\mathbf{P}$ is an $n \times t$ matrix and in general not a square matrix, hence $\mathbf{P}^{-1}$ is not defined. However, since $\mathbf{p}_k^T \mathbf{p}_k = 1$, $\mathbf{P}^T$ can be used instead as shown in Equation A.10

$$\mathbf{p}^T = \begin{pmatrix} \mathbf{p}_1^T \\ \mathbf{p}_2^T \\ \vdots \\ \mathbf{p}_n^T \end{pmatrix} \tag{A.10}$$

in place of $\mathbf{P}^{-1}$ to project out the weights required to parameterise a new feature vector via Equation A.11

$$\mathbf{b} = \mathbf{P}^T (\mathbf{x} - \bar{\mathbf{x}}) \tag{A.11}$$

Equation A.11 allows us to generate new examples of the data by varying the parameters $b_1, b_2, \ldots, \ldots, b_k$, within suitable limits so that new examples will be similar to the training set data. The limits for the b_k are derived by examining the distributions of the parameter values required to generate the training set. Since the variances of b_k over the training set can be shown to be λ_k, suitable limits are given by

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \tag{A.12}$$

since most of the population (99.73%) lies within three standard deviations of the mean.

Appendix B

Movie Descriptions

In this section, we provide short descriptions of the simulation movies supplied with this thesis. All movies can be found on the attached CD and are of *avi* format. The movies can be viewed using suitable software like **Real One player or Windows Media Player**.

- **Stekel.avi** This movie is acquired from our implementation of Stekels' [Stekel et al., 1995] model for squamous epithelium. As can be seen, the model is initialised with two stem cells. As cells divide and differentiate, the rete pegs are formed. The results produced from our implementation in the form of model outputs are similar to the published work from Stekel [Stekel et al., 1995].

- **Scenario1.avi** We demonstrate our model for epithelial cell interactions in this movie as explained in Chapter 6. The model was initialised with two stem cells. As can be seen from the movie, stem and basal cells grow and divide. As cells differentiate, they move away from the basement membrane and begin to form layers of different types of cells.

- **Scenario2.avi** Here we show the results obtained for the scenario2 explained in Chapter 6. In this scenario, we assume that there is no differential adhesion. As

can be seen, as stem cells divide they move away from the basement membrane. This is because there is no difference in the adhesive strengths between stem cells and other cells in the model.

- **Scenario3.avi** This movie demonstrates the scenario 3 explained in chapter 6. We assume here that stem cells prefer to be close to dead cells than the basement membrane and so on. As can be seen, stem cells move off the basement membrane. We fail to observe any cell sorting in this scenario.

- **Stekel3D.avi and rotate3D.avi** We demonstrate the results of some preliminary work to extend Stekels' model to 3D. The extended model does not include all biological details from the Stekels model. As can be seen from the simulation, the stem cells divide and produce TA cells, which differentiate when they leave the basement membrane. Since we did not include the connective tissue attractive and repulsive force in the extended model as suggested by Stekel [Stekel et al., 1995], stem cells can be seen to leave the basement membrane. The rotate3D.avi shows the simulation from a different angle.

Bibliography

[Alonso and Fuchs, 2003] Alonso, L. and Fuchs, E. (2003). Stem cells of the skin epithelium. *Proceedings of the National Academy of Sciences*, 100(suupl 1):11835.

[Aranjo, 1999] Aranjo, C. (1999). Detection and categorisation of epithelial dysplasia. Master's thesis, University College London, Dept. Of Computer Science, UCL, Gower Street, London.

[Ball, 2000] Ball, P. (2000). Biophysics: Science in motion. *Nature*, 406:244–45.

[Braga and Harwood, 2001] Braga, V. and Harwood, J. (2001). Super glue. *Nature Cell Biology.*

[Canny, 1986] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, 8(6).

[Chellappa and Bagdazian, 1984] Chellappa, R. and Bagdazian, R. (1984). Fourier coding of image boundaries. *IEEE Trans. Pattern Anal. Machine Intell.*, 6:102–05.

[Clem et al., 1992] Clem, C., Boysen, M., and Rigaut, J. (1992). Towards 3-d modelling of epithelia by computer simulation. *Analytical Cellular Pathology*, 4:287–302.

[Clem et al., 1997] Clem, C., König, D., and Rigaut, J. (1997). A three-dimensional dynamic simulation model of epithelial tissue renewal. *Analytical and Quantitative Cytology and Histology*, 19(2):174–183.

[Clem and Rigaut, 1995] Clem, C. and Rigaut, J. (1995). Computer simulation modelling and visualization of 3d architecture of biological tissues. simulation of the evolution of normal, metaplastic and dysplastic states of the nasal epithelium. *Acta Biotheoretica*, 43:425–442.

[Cristianini and Shawe-Taylor, 2000] Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines.* Cambridge University Press.

[Curtis, 1960] Curtis, A. (1960). Cell contacts: Some physical considerations. *American Naturalist*, 94:37–56.

[Curtis, 1961] Curtis, A. (1961). Timing mechanisms in the specific adhesion of cells. *Exp.Cell.Res.Suppl.*, 8:107–122.

[Dahmann and Basler, 1999] Dahmann, C. and Basler, K. (1999). Compartment boundaries: at the edge of development. *Trends in Genetics*, 15(8):320–6.

[Davies, 1990] Davies, E. (1990). *Machine Vision: Theory, Algorithms and Practicalities.* Academic Press.

[Davies and Garrod, 1997] Davies, J. and Garrod, D. (1997). Molecular aspects of the epithelial phenotype. *Bioessays*, 19(8):699–704.

[Dover and Potten, 1988] Dover, R. and Potten, C. (1988). Hetrogeneity and cell cycle analysis from time-lapse studies of human keratinocytes in vitro. *Journal of Cell Science*, 89.

[Drasdo, 2000] Drasdo, D. (2000). Buckling instabilities in one layered growing tissues. *Physical Review Letters*, 84:4424–27.

[Drasdo et al., 1995] Drasdo, D., Kree, R., and McCaskill, J. (1995). Monte carlo approach to tissue-cell populations. *Physical Review E*, 52(6):6635–55.

[Drasdo and Loeffler, 2001] Drasdo, D. and Loeffler, M. (2001). Individual based models to growth and folding in one layered-tissues: Intestinal crypts and blastulation. *Nonl.Analysis*, 47:245–56.

[Dubertret and Rivier, 1997] Dubertret, B. and Rivier, N. (1997). The renewal of the epidermis: a topological mechanism. *Biophysics Journal*, 73(38-44).

[Duda et al., 2001] Duda, R., Hart, P., and Stork, D. (2001). *Pattern Classification.* Wiley Interscience.

[Fuchs and Cleveland, 1998] Fuchs, E. and Cleveland, D. (1998). A structural scaffolding of intermediate filaments in health and disease. *Science*, 279:514–19.

[Fuchs and Raghavan, 2002] Fuchs, E. and Raghavan, S. (2002). Getting under the skin of epidermal morphogenesis. *Nature Reviews: Genetics*, 3(3):199–209.

[Glazier and Graner, 1993] Glazier, J. and Graner, F. (1993). Simulation of the differential driven rearrangement of biological cells. *Phyical Review E*, 47:2128–2154.

[Goss, 1964] Goss, R. (1964). *Adaptive Growth.* Logos, Academic Press, New York.

[Graner, 1993] Graner, F. (1993). Cell surface adhesion drive cell-rearrangement? part1: Biological cell-sorting. *J. Theor. Biol.*, 164:455–476.

[Graner and Glazier, 1992] Graner, F. and Glazier, J. A. (1992). Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical Review Letters*, 69(13):2013–2016.

[Graner and Sawada, 1993] Graner, F. and Sawada, Y. (1993). Cell surface adhesion drive cell-rearrangement? part2: A geometric model. *J. Theor. Biol.*, 164:477–506.

[Granlund, 1972] Granlund, G. (1972). Fourier pre-processing for hand print character recognition. *IEEE Transactions on Computers*, C-21:195–201.

[Green and Gaudry, 2000] Green, K. and Gaudry, C. (2000). Are desosomes more than tethers for intermediate filaments? *Nat. Rev. Mol. Cell. Biol.*, 1:208–216.

[Gumbiner, 1996] Gumbiner, B. (1996). Cell adhesion: The molecular basis of tissue architecture and morphogenesis. *Cell*, 84:345–57.

[Gumbiner et al., 1988] Gumbiner, B., Stevenson, B., and Grimaldi, A. (1988). The role of the cell adheion molecule uvomorulin in the formation and maintenance of the epithelial junction complex. *Journal of Cell Biology*, 107:1575–87.

[Hillier, 1996] Hillier, B. (1996). *Space is the Machine.* Cambridge University Press, Cambridge.

[Hillier and Hanson, 1984] Hillier, B. and Hanson, J. (1984). *The Social Logic of Space.* Cambridge University Press, Cambridge.

[Honda et al., 1996] Honda, H., Tanemura, H., and Imayama, S. (1996). Spontaneous architectural organisation of mammalian epidermis from random cell packing. *Journal of Investigative Dermatology*, 106(312-315).

[Hynes, 1992] Hynes, R. (1992). Integrins:versatality, modulation and signalling in cell adheison. *Cell*, 69:11–25.

[Ising, 1925] Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschr. f. Physik*, 31:235–258.

[Izaguirre et al., 2004] Izaguirre, J., Chaturvedi, R., Huang, C., Cickovski, T., Coffland, J., Thomas, G., Forgacs, G., Alber, M., Hentschel, G., Newman, S., and Glazier, J. (2004). Compucell, a multi-model framework for simulation of morphogenesis. *Bioinformatics*, 20(7):1129–37.

[Jain, 1989] Jain, A. (1989). *Fundamentals Of Digital Image Processing.* Prentice-Hall.

[Jensen et al., 1999] Jensen, B., Lowell, S., and Watt, F. (1999). The spatial relationship between stem cells and their progeny in the basal layer of human epidermis: a new view based on whole-mount labelling and lineage analysis. *Development*, 126:2409–2418.

[Jones and Watt, 1993] Jones, P. and Watt, F. (1993). Separation of human epidermal stem cell from transit amplifying cells on the basis of differences in integrin function and expression. *Cell*, 73:713–724.

[Jones and Watt, 1995] Jones, P.H.and Harper, S. and Watt, F. (1995). Stem cell patterning and fate in human epidermis. *Cell*, 80:83–93.

[Kemler, 1992] Kemler, R. (1992). Classical cadherins. *Semin. Cell Biology*, 3:149–55.

[Kimmel, 2000] Kimmel, R. (2000). Two lineage boundaries coordinate vertebrate apical ectodermal ridge formation. *Genes Development*, 14:1377–89.

[Kohavi, 1995] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence.*

[Landini and Rippin, 1996] Landini, G. and Rippin, J. W. (1996). How important is tumour shape? quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis. *Journal Of Oral Pathology*, 179:210–217.

[Lavker and Sun, 1983] Lavker, R. and Sun, T. (1983). Epidermal stem cells. *Dermatology*, 81:121s–127s.

[Leicester et al., 1998] Leicester, S., Finney, J., and Bywater, R. (1998). Description of molecular surface shape using fourier descriptors. *Journal of Molecular Graphics*, 6.

[Lestrel, 1997] Lestrel, P. (1997). *Fourier Descriptors and Their Applications in Biology.* Cambridge University Press.

[Loeffler et al., 1986] Loeffler, M., Stein, M., Wichmann, R., Potten, H., Kraur, C., and Chwalinski, S. (1986). Intestincal cell proliferation. a comprehensive model of steady-state proliferation in the crypt. *Cell Tissue Kinetics*, 19:627–45.

[Mackenzie, 1997] Mackenzie, I. (1997). Retroviral transduction of murine epidermal stem cells demonstrates clonal units of epidermal structure. *Journal of Investigative Dermatology*, 109:377.

[Magdalena, 2002] Magdalena, Z. (2002). Patterning of the embryo: the first spatial decisions in the life of a mouse. *Development*, 129:815–29.

[Mareé, 2000] Mareé, A. F. M. (2000). *From Pattern Formation to Morphogenesis.* PhD thesis, Theoretical Biology/Bioinformatics Group, University of Utrecht.

[McNeill, 2000] McNeill, H. (2000). Sticking together and sorting things out: Adhesion as a force in development. *Nature Reviews:Genetics*, 1:100–108.

[McNeill et al., 1990] McNeill, H., Ozawa, H., Kemler, R., and Nelson, W. (1990). Novel function of the cell adhesion molecule uvomorulin as an inducer of cell surface polarity. *Cell*, 62:309–316.

[Meineke et al., 2001] Meineke, F., Potten, C., and Loeffler, M. (2001). Cell migration and organisation in the intestinal crypt using a lattice free model. *Cell Proliferation*, 34(4):253–66.

[Mombach, 1999] Mombach, J. (1999). Simulations of embryonic cell self-organisation: a study of aggregates with different concentrations of cell types. *Physical Review E.*

[Mombach et al., 1993] Mombach, J., DeAlmedia, C., and Iglesias, R. (1993). Mitosis and growth in biological tissues. *Physical Review E.*, 48:598–602.

[Mombach and Glazier, 1996] Mombach, J. and Glazier, J. (1996). Single cell motion in aggegrates of embryonic cells. *Physical Review Letters*, 76:3032–35.

[Mombach et al., 1995] Mombach, J., Glazier, J., Glazier, R., and Zajac, M. (1995). Quantitative comparison between differential adhesion models and cell sorting in the prescence and abscence of fluctuations. *Physical Review Letters*, 75:2244–47.

[Murray and Hunt, 1993] Murray, A. and Hunt, T. (1993). *The Cell Cycle.* Oxford University Press.

[Nixon and Aguado, 2002] Nixon, M. and Aguado, A. (2002). *Feature Extraction and Image Processing.* Newnes.

[Nose et al., 1988] Nose, A., Nagafuchi, A., and Takeichi, M. (1988). Expressed recombinant cadherins mediate cell sorting in model systems. *Cell*, 54:993–1001.

[Novak et al., 1998] Novak, B., Csikasz-Nagy, A., Gyorffy, B., Nasmyth, K., and Tyson, J. (1998). Model scenarios for evolution of the eukaryotic cell cycle. *Phil.Trans.R.Soc.Lond B*, 353:2063–76.

[Okebe et al., 1992] Okebe, A., Boots, B., and Sugihara, K. (1992). *Spatial Tessellations, Concepts and Applications Of Voronoi Diagrams.* Probability And Mathematical Statistics. John Wiley and Sons, John Wiley and Sons Ltd, Baffins Lane, Chichester, West Suusex, PO19 1UD, England.

[Pelfer and Polakis, 2000] Pelfer, M. and Polakis, P. (2000). Wnt signalling in oncogenesis and embryogenesis - a look outside the nucleus. *Science*, 287:1606–09.

[Perez-Moreno et al., 2003] Perez-Moreno, M., Jamora, C., and Fuchs, E. (2003). Sticky business: Orchestrating cellular signals at adherens junctions. *Cell*, 112:535–48.

[Persoon and Fu, 1977] Persoon, E. and Fu, K. (1977). Shape descrimination using fourier descriptors. *IEEE Trans.Syst.Man and Cybern.*, 7:170–79.

[Persoon and Fu, 1986] Persoon, E. and Fu, K. (1986). Shape descrimination using fourier descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:388–397.

[Potten and Loeffler, 1990] Potten, C. and Loeffler, M. (1990). Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. lessons for and from the crypt. *Development*, 110:1001–20.

[Potts, 1952] Potts, R. (1952). Some generalized order-disorder transformations. *Proc. Camb. Phil. Soc.*, 48:106–109.

[Reynolds, 1987] Reynolds, C. W. (1987). Flocks, herds, and schools: A distributed behavioral model, in computer graphics. *SIGGRAPH '87 Conference Proceedings*, 21(4):25–34.

[Runswick et al., 2001] Runswick, S., O'Hare, M., Jones, L., Streuli, C., and Garrod, D. (2001). Desmosomal adhesion regulates epithelial morphogenesis and cell positioning. *Nature Cell Biology*, 3:823–830.

[Savill, 2003] Savill, N. (2003). Mathematical models of hierarchically structured cell populations under equilibrium with application to the epidermis. *Cell Proliferation*, 36(1-26).

[Savill and Hogeweg, 1997] Savill, N. and Hogeweg, P. (1997). Modelling morphogenesis: From singlce cells to crawling cells. *Journal of Theoretical Biology*, 184:229–235.

[Savill and Sherratt, 2003] Savill, N. and Sherratt, J. (2003). Control of epidermal stem cell clusters by notch mediated lateral-induction. *Developmental Biology*, 258(141-153).

[Schaller et al., 2003] Schaller, G., Beyer, T., and Meyer-Hermann, M. (2003). Weighted dynamic delaunay triangulations in tumour-growth modelling. *International Conference, Linking mathematical and biological models in cancer research.*

[Sheppard, 1996] Sheppard, D. (1996). Epithelial integrins. *Bioessays*, 18(8):655–660.

[Shochat et al., 1999] Shochat, E., Hart, D., and Agur, Z. (1999). Using computer simulations for evaluating the efficacy of breast cancer chemotherapy protocols. *Journal of Mathematical Models and Methods in Applied Sciences*, 9(4):599–615.

[Speight et al., 1996] Speight, P., Farthing, P., and Bouquot, J. (1996). The pathology of oral cancer. *Current Diagnostic Pathology*, 3:165–76.

[Steinberg, 1963] Steinberg, M. (1963). Reconstruction of tissues by dissociated cells. *Science*, 141:401–408.

[Steinberg and Foty, 1997] Steinberg, M. and Foty, R. (1997). Intercellular adhesions as determinants of tissue assembly and malignant invasion. *Journal of Cellular Pathology*, 173:135–39.

[Steinberg and Takeichi, 1994] Steinberg, M. and Takeichi, M. (1994). Experimental specification of cell sorting, tissue spreading and specific spatial patterning by quantitative differences in cadherun expression. *Pro.Natl.Acad.Sci. U.S.A.*, 91:206–209.

[Steinberg, 1996] Steinberg, S. (1996). Adhesion in development: A historial overview. *Developmental Biology*, 180:377–88.

[Stekel et al., 1995] Stekel, D., Rashbass, J., and Williams, D. (1995). A computer graphic simulation of squamous epithelium. *J.Theor.Biol.*, 175:283–293.

[Stott et al., 1999] Stott, E., Britton, N., Glazier, J., and Zajac, M. (1999). Stochastic simulation of benign avascular tumour gowth using the potts model. *Mathematical and Computer Modelling*, 30:183–98.

[Thomas and Speight, 2001] Thomas, G. and Speight, P. (2001). Cell adhesion molecules and oral cancer. *Critical Reviews in Oral Biology and Medicine*, 12:479–498.

[Trucco and Verri, 1998] Trucco, E. and Verri, A. (1998). *Introductory Techniques For 3-D Computer Vision.* Prentice Hall.

[Tsukita et al., 2001] Tsukita, S., Furuse, M., and Itoh, M. (2001). Multifactional strands in tight junctions. *Nat.Rev.Mol.Cell.Biol.*, 2:285–93.

[Turner and Sherratt, 2002] Turner, S. and Sherratt, J. (2002). Intercellular adhesion and cancer invasion: A discrete simulation using the extended potts model. *Journal of Theoretical Biology*, 216:85–100.

[Vasioukhin et al., 2000] Vasioukhin, V., Bauer, C., Yin, M., and Fuchs, E. (2000). E directed actin polymerizations is the driving force for epithelial cell-cell adhesion. *Cell*, 100:209–219.

[Vleminckx and Kemler, 1999] Vleminckx, K. and Kemler, R. (1999). Cadherins and tissue formation: integrating adhesion and signalling. *BioEssays*, 21:211–20.

[Watt, 1998] Watt, F. (1998). Epidermal stem cells, patterning and the control of stem cell fate. *Phil.Trans.R.Soc.Lond.B.*, 353:831–37.

[Watt, 2001] Watt, F. (2001). Stem cell fate and patterning in mammalian epidermis. *Current Opinion in Genetics and Development*, 11(410-417).

[Wilson, 1907] Wilson, H. (1907). On some phenomena of coalescence and regeneration in sponges. *Journal of Experimental Zoology*, 5:245–58.

[Wolfram, 1983] Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of Modern Physics*, 55:601–44.

[Wolfram, 1994] Wolfram, S. (1994). *Cellular Automata and Complexity: Collected Papers.* Westview Press.

[Wolfram, 2002] Wolfram, S. (2002). *A New Kind of Science.* Wolfram Media Incorporated.

[Wright and Alison, 1984] Wright, N. and Alison, M. (1984). *The biology of epithelial cell populations*, volume 1. Clarendon Press, Oxford.

[Yeaman et al., 1999] Yeaman, C., Grindstaff, K., and Nelson, W. (1999). New perspectives on mechanisms involved in generating epithelial cell poarity. *Physiological Reviews*, 79:73–98.

[Zhang and Lu, 2002] Zhang, D. and Lu, G. (2002). A comparative study on shape retrieval using fourier descriptors with different shape signatures. *In Proc. Of The Fifth Asain Conference on Computer Vision, ACCV*, pages 646–651.