THE UNIVERSITY OF HULL

Digital Image Processing Techniques and their Application to the Automation of Palynology

being a Thesis submitted for the Degree of

Doctor of Philosophy

in the University of Hull

by

Walter John Treloar, B.Sc(Hons)

August, 1992

Acknowledgements

Thanks are due to Prof John Flenley for encouraging me to attempt this project in the first place. As an image processing novice I found the advice and help from both Prof Gaynor Taylor and Dr Helen Shen invaluable in explaining many new concepts during the course of this research. For their help in obtaining pollen samples, I would like to thank both Dr Sarah Metcalfe and Annette Parkes. Also to Tony Sinclair for introducing me to the wonders of the SEM!

Many thanks are also due to my proof-reading friends Elaine and Andy who spent much time going over this thesis for me. Mention must also be made to all my other friends particularly Tracy, Sue, Liz, Debbie and Richard who kept me going in times of doubt. Finally, I acknowledge the lads from the Blob Premier Division Card School for making the coffee breaks of the last couple of years so entertaining.

Contents

Chapter 1

Introduction

1.1 Introduction

There are several examples of studies that combine different types of cellular analysis with digital image processing techniques. Rutovitz et al[39], describe a system used to analyse genetic material within cells and Psenner[36] has a semi-automatic system that analyses the sizes of planktonic bacteria. Watanabe and the CYBEST Group[46] use image processing techniques to check cervical smears for cancerous cells. Similar techniques are also used to automate blood cell counts. However research combining palynology with digital image processing is extremely rare to say the least.

This chapter provides a basic introduction to palynology, then outlines the particular problems that this thesis attempts to tackle and how they may be solved using image processing techniques.

1.2 What is Palynology?

Palynology is quite simply the study of all aspects of pollen and spores, such as pollen formation, morphology, dispersal and deposition. Pollen grains are formed in the anthers of flowering plants (gymnosperms or angiosperms). They carry the male sex cell (gamete) to the stigma of the female part of the flower.

Spores are reproductive cells for lower plants such as ferns (pteridophyta) and mosses (bryophyta). For convenience future references to pollen will encompass both pollen and spores.

Pollen are extremely small with the majority falling within the range 25μm to 35μm. Few pollen exceed 100μm in diameter. The outer cell wall or exine is a highly resistant waxy coat of material called sporopollenin. It is the morphological features such as the number, distribution and shape of apertures, the size and shape of the grain, as well as the exine texture that forms the bases of pollen identification.

A major part of palynology is pollen analysis. This is a technique used by palaeontologists, botanists and biogeographers to assist in the reconstruction of past vegetational population structures (assemblages), and thus climatic and environmental conditions. It is based on the identification and counting of various pollen types (or taxa) which have been preserved in deposits such as peat and lake muds. Well defined sequences of pollen taxa throughout a sediment core can be identified. These represent climatic swings from glacial to interglacial periods that, for example, may help to make spatial correlations between climatic events. This is currently very important in modelling climate change for research into the potential threat of the greenhouse effect. Pollen analysis cannot give an absolute age for a deposit, but it can assist in relative dating and for this purpose it is often used by the oil industry.

1.3 Current Problem with Pollen Analysis

The identification of pollen from a sediment core is one of the most time consuming processes in pollen analysis. Samples from different levels of a pollen bearing sediment are treated with various chemical cocktails to remove as much non-pollen material (detritus) as possible. Each sample then requires hours of work by highly skilled personnel to identify and count over 600 pollen grains which in turn may contain over 50 different taxa. Finally, a pollen diagram is produced to show how the concentrations of pollen taxa change throughout the sediment core and therefore time. The automation of any part of this process would allow more time for the evaluation of data rather than on its collection. In particular the automation of the identification and counting of pollen would bring great benefits to research. Results could be obtained more rapidly and with greater objectivity. The pollen would be identified using well defined criteria and would not be susceptible to any human bias. Objectivity may also be increased by using larger data sets per sample than is at feasible at present. Using more samples from a sediment profile would allow a finer resolution analysis of vegetational fluctuations. It is with this aspect of the automation problem that this thesis is concerned.

1.4 Digital Image Processing as a Solution

The technique employed to solve this automation problem is based upon digital image processing. Before any automated identification can be discussed it is important to briefly review the process employed by human researchers to identify pollen. There are many classification schemes that may be followed for pollen identification. Moore and Webb[32], for example, use the type, number and arrangement of apertures over the pollen surface as the initial classification. This is relatively easy with an optical microscope as a pollen grain can be focused through or even rotated to give a 3-D impression. An aperture is a thin or missing part of the exine which is independent of the surface texture of the exine. There are two basic types in pollen; colpi or furrows (illustrated in Plate 12a) and pori or pores. Colporate structures may also form as combinations of these. Spores have neither colpi or pori but do possess a three-branched slit forming a Y-shape structure called a trilete. The subsequent classification levels use a multitude of criteria to separate the pollen into successively smaller sub-divisions. Such criteria include the morphology of the aperture edges which may be thickened to different extents or partially filled. Other surface structures such as thickened bands connecting apertures (arcus), large spines (spinae (see Plates 1, 4, 5 & 10)) or small spines (spinulae (see Plate 3a)) may occur. The above examples are only a small proportion of the different features that may be used to distinguish between pollen taxa.

Any attempt to duplicate such a classification scheme for use with digital image processing techniques would be overly complex. Previous work by Langford[25] ignored all the features mentioned above and classified six pollen taxa using their fine structure (or texture). Examples of pollen texture can be seen in Plates 2b, 3b, 6b, 7b, 8b, 9b, 11b and 12b. Textural features are seldom utilised by human workers for identification purposes as texture is not very distinctive under optical microscopes as they have a narrow depth of field and relatively low magnification. Langford[25] used SEM pollen images in which texture is very well defined due to the greater field depth and magnification. There are many methods that may be used to describe texture in digital image processing, some of which have proved very useful for this task (see chapter 2 for further details). However these may prove to be less effective when used to identify large numbers of pollen taxa. An additional classification level may need to be incorporated.

1.5 Overall Classification Scheme

The overall classification scheme envisaged is summed up in Fig.1.01. The first level is the sample preparation stage where the raw pollen sample is treated in various ways to remove as much detrital material as possible and may also be used to produce a rough sorting of pollen taxa. The refined pollen sample is then mounted on an SEM stub ready for the next stages of the classification scheme where image processing techniques

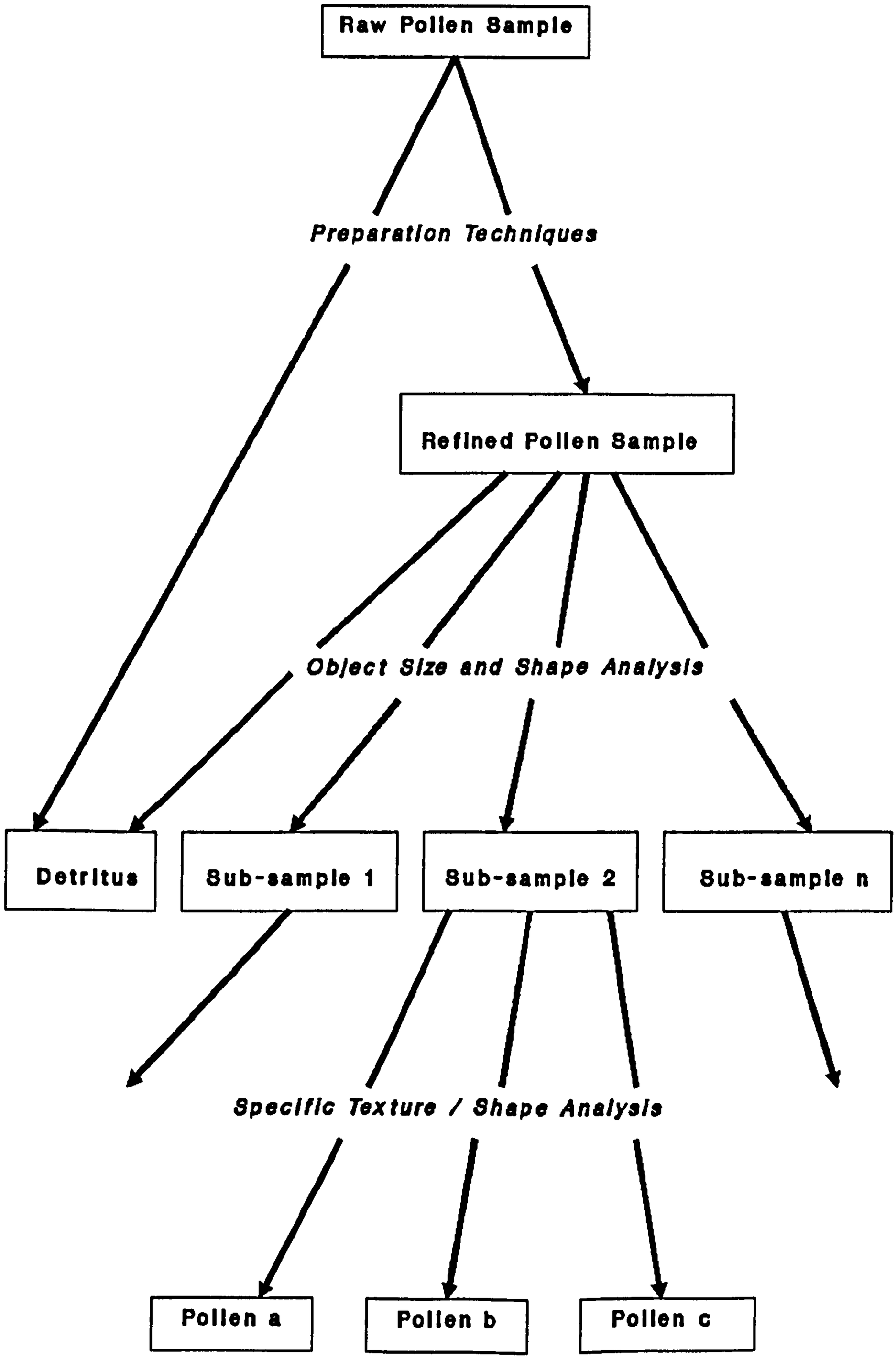


Fig.1.01 Overall Classification Scheme

are utilised. The particles found in the refined sample are split into sub-samples of pollen and detrital material on the grounds of their general size or shape. The exact nature of the size and shape analysis is explored in this thesis. At the final level of the scheme the pollen sub-samples are classified using a set of texture or shape discrimination variables that are optimal for discriminating between that particular sub-sample. Ideally the sub-samples of pollen should consist of as few pollen as possible, thus reducing the number of comparisons required to produce the final classification. Langford[25] has already shown that texture analysis is very useful for the discrimination of small groups of pollen. This thesis also explores shape analysis at this final classification level.

1.6 Outline of Thesis

The main thrust of this thesis is to examine features in the above classification scheme that have as yet been unexplored, namely pollen shape and geometry. The format of the thesis is as follows:

Chapter 2 is an overall review of literature and previous work relating to automation in palynology and image processing. Problems that were left unanswered after earlier research are outlined. Then chapter 3 describes how and why the previous pollen image processing system was changed and describes the

new hardware and software systems used. The chapter concludes by describing the pollen samples used to test the classification scheme on the new system.

Chapter 4 assesses numerous methods that may be employed to detect and locate objects in an image prior to shape and geometric analysis. Chapter 5 explores a number of schemes the may be useful for shape and geometric analysis in the gross classification of pollen. A new method of shape analysis for the lower levels of classification is also explored for spinose (spiny) pollen. The chapter concludes by attempting to classify a set of pollen taxa using the classification scheme outlined in Fig.1.01.

Chapter 6 concludes the thesis by exploring the problem of automatically selecting the best region of a pollen surface from which take a distinctive texture sample for analysis. The potential usefulness of neural networks to the problem of pollen identification is also explored here. Future work and an outline of any further problems to be solved conclude this chapter.

Chapter 2

Research Review

2.1 Literature Review

2.1.1 Automation in Palynology

There is a distinct lack of literature that has been written or work that has been done in the field of automatic identification in pollen analysis. However, there have been developments in the use of computers for other aspects of pollen analysis, for example the analysis of raw pollen taxa frequency data. There are several software packages which now do this and present the information as pollen diagrams eg Squires[41].

Pollen database systems have also been developed, eg Walker et al[44] and Guppy et al[20]. These may be used as tools to aid the identification of unknown pollen taxa. The morphological characteristics of an observed pollen grain are compared with those within the database. The closest match is then given as a suggestion of its identity.

Flenley[11] proposed automation beyond the above mentioned data manipulation techniques to include the data acquisition process of pollen analysis. Both the automatic counting and identification of pollen could be used to clear up many of the problems described in section 1.3. It was envisaged that the once slow and laborious process would become faster and less arduous. Also it could increase the numbers of grains that can currently be practically counted. Thus an increase in the objectivity of results would occur allowing more horizons to

be sampled and increasing the sample size within each horizon. This fine resolution sampling would be very useful in pollen analysis. There is also the possibility of differentiating between taxa that cannot at present be separated by conventional means.

The identification of pollen purely by image based data on computer was attempted by Mirkin and Bagdasaryan[31]. They used the technique of 'template-matching' which is difficult with material such as pollen that has a large amount of variability. Mirkin and Bagdasaryan[31] concluded that the greatest limitation was the small amount of memory available on the computers of that time. However, they did recognize the potential for the computer identification of pollen.

2.1.2 Digital Image Processing

One of the first applications for digital image processing was in the 1920's when photographic images were digitised for transmission along telegraph wires and subsequently reconstructed by specialised printing equipment. It took the arrival of large-scale digital computers and the space program to trigger the development of image processing techniques into what they are today. The Jet Propulsion Laboratory, California, was a

driving force for image processing with its work on image systems for lunar and planetary probes. Since then image processing techniques have undergone vigourous expansion into a multitude of areas. For example, biomedical images such as x-rays are enhanced by computers to make them clearer for human inspection.

The basic hardware components of an image processing system are a video input device, a frame-store (as an add-on within a computer) and a video output monitor.

A video input device (such as a video camera or SEM as in this case) transmits a complete image (frame), line by line, as an analogue video signal where voltage represents brightness.

A frame-store receives this analogue signal from the input device and converts it to a digital format using a high speed analogue to digital converter. Having been formed by discretising the analogue signal to both brightness and spatial co-ordinates the digital image is placed line by line in a rectangular array of memory in the frame-store. It may be considered as a matrix whose column and row indices identify a point in the image and the corresponding matrix element value identifies the grey-level (or brightness) at that point with black represented by the lowest value and white the highest. The elements of such a digital array are called image elements or pixels. The size of the image stored can vary from about 256x256 to 1028x1028 pixels, brightness levels may vary from

64; roughly the number of grey-levels that the human eye can distinguish, to 256 grey-levels. If colour images are used the number of brightness levels can be greatly increased.

The contents of the frame-store are continually converted back into an analogue signal for display on a video monitor. This aids the user by showing any subsequent changes made to the image.

Images within the frame-store can be manipulated and analysed by computer using numerous image processing techniques. The following subsections give an outline of some of these techniques used in this research.

2.1.2.1 Point Operators

Point operators are nonlinear operators that transform input images into output images by mapping input grey-levels to different grey-levels according to certain criteria. They may be used to enhance or restore images and may also be viewed as a way of transforming the histogram of the input image to a new histogram for the output image. Two point operations are discussed here, namely histogram equalisation and thresholding of images.

Histogram equalisation can perform many tasks in image processing but in this study it provides an important first step for texture analysis. It enhances the contrast of an image and produces a standard format image giving a greater ability to compare different images taken under different lighting conditions. It also allows the number of grey-levels within an image to be reduced and thus reduces the computation required on some textural measures such as those proposed by Haralick et al[22] (see below).

The operation is very easily carried out. From the probability histogram (P) of an image a transformation histogram (T) is readily constructed thus:

$$T(k) = \sum_{i=0}^{k} p(i)$$

$0 \leq T(k) \leq 1$ and $0 \leq k \leq L-1$

L = maximum grey-level.

$p(i)$ = probability of grey-level i occurring.

If, for example, a 64 grey-level image is being equalised to a 16 grey-level image then each successive $\frac{1}{16}th$ of the transformation histogram is assigned to successive grey-levels. For instance, a grey-level 0 is assigned to all grey-levels in the transformation histogram which range from 0 to $\frac{1}{16}th$. Likewise grey-level 6 would be applied to all grey-levels with values greater than $\frac{6}{16}th$ and up to $\frac{7}{16}th$.

The second point operator is the thresholding operation. A threshold (θ) is a grey-level that is used to segment an image into two classes, namely object and background. That is:

$$B(i,j) = 1 \qquad if \quad I(i,j) \geq \theta$$

and $$B(i,j) = 0 \qquad if \quad I(i,j) < \theta$$

$B(i,j)$ is the resultant binary image.

$I(i,j)$ is the original image.

The threshold may be set a priori or determined adaptively using the data in $I(i,j)$. There are a wide variety of techniques that may be used to automatically set a threshold. These may be broadly divided into two groups; global and local. Global techniques use a single grey-level value to threshold the entire image and may be sub-divided into either point-dependent or region-dependent techniques. The point-dependent techniques (eg. Otsu[33] and Levine & Nazif[27]) assign a threshold using only the grey-level histogram of the image. A region-dependent technique (eg. Wong & Sahoo[49] and Watanabe et al[46]) however uses a local property within each pixel neighbourhood, the gradient for example. Local techniques split the image into sub-images and assign thresholds for each one.

2.1.2.2 Neighbourhood Operators

Neighbourhood operations are carried out on groups of adjacent pixels. Each pixel in an image is surrounded by eight neighbours. By relating the central pixel to its neighbours the changes in intensity across small areas of the image may be either enhanced or diminished. This operation is carried out by convolving an image with a neighbourhood operator. This is easily demonstrated by the use of an example. The following operator is a Laplacian operator, Prewitt[35].

$$\begin{matrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{matrix}$$

For every 8-neighbourhood in an image this operator represents the following equation:

$$L(i,j) = 8 \times I(i,j) + \sum_{n=1}^{N} -1 \times I(i,j)_n$$

N = number of neighbours.

$L(i,j)$ =new value for central pixel of neighbourhood.

$I(i,j)$ = central pixel of original image neighbourhood.

$I(i,j)_n = n^{th}$ neighbour of central pixel.

This convolution has the effect of enhancing the differences in intensity between adjacent pixels, resulting in edge

enhancement. Examples of other operators are image smoothing (for noise reduction) or image sharpening.

2.1.2.3 Texture Analysis

In the field of texture analysis there is a multitude of techniques that may be used. Previous work by Langford[25] found that two techniques of texture analysis were particularly useful, namely those proposed by Haralick et al[22] and Laws[26] (also Pietikainen et al[34]).

Haralick et al[22] produced a set of second-order statistics to analyse the grey-tone spatial dependence of textures. It has been used in various ways. Don et al[9] used it to measure surface roughness in metals and Weszka et al[47] classified terrain from aerial and satellite images.

The starting point of this type of analysis is the construction of co-occurrence matrices containing information on the spatial organisation of grey-levels of an image. This is best illustrated with the use of an example (see Fig.2.01). A digital image (Fig.2.01(a)) is first equalised (as described above) and then a grey-level distribution matrix constructed (Fig.2.01(b)) using the displacement vector (1,0). This means that each pixel

Digital Image f(x,y)

0	0	1	1
0	0	2	3
0	2	2	3
1	2	3	2

(a)

Grey-level Distribution Matrix

	0	1	2	3
0	2	1	2	0
1	0	1	1	0
2	0	0	1	3
3	0	0	1	0

(b)

Symmetrical Grey-level Distribution Matrix

	0	1	2	3
0	2	0.5	1	0
1	0.5	1	0.5	0
2	1	0.5	1	2
3	0	0	2	0

(c)

Co-occurrence Matrix : P(I,J)

	0	1	2	3
0	.17	.04	.08	0
1	.04	.08	.04	0
2	.08	.04	.08	.17
3	0	0	.17	0

(d)

Fig.2.01 Construction of Co-occurrence Matrix

in the equalised image is compared with the pixel found at a distance of 1 pixel away and at an angle of 0° from the x-axis. The elements within the matrix that are indexed by the grey-levels of the pixel pairs are incremented. The matrix is then made symmetrical (Fig.2.01(c)) about the main diagonal by averaging equivalent elements, eg element (1,2) and (2,1). The co-occurrence matrix (Fig.2.01(d)) is then easily formed by dividing each element by the sum of all elements in the symmetrical grey-level distribution matrix. Each element, $P(i,j)$, is the probability of finding grey-levels i and j separated by a displacement vector of (1,0).

When co-occurrence matrices are used to analyse texture a variety of displacement vectors are used. To reduce the effects of rotation of a texture the displacement vectors over the angles 0°, 45°, 90° and 135° are usually averaged.

Haralick et al[22] proposed 14 statistical measures of distribution within the matrix that relates to textural information. However, Langford[25], Waterhouse[45] and Rutter[40] all used only 6 of the most easily calculated measures in pollen texture analysis. These are as follows:

i) Angular Second Moment or homogeneity of the image is a measure of how even the values are distributed throughout the matrix; the lower the value the greater the homogeneity.

$$ASM = \sum_{i=0}^{L} \sum_{j=0}^{L} P(i,j)^2$$

L = maximum grey-level.

ii) Contrast or local variation (coarseness) of grey-levels is a measure of the moment of inertia about the main diagonal of the matrix. When the probabilities are concentrated about this main diagonal the contrast is low.

$$CON = \sum_{i=0}^{L} \sum_{j=0}^{L} (i-j)^2 \cdot P(i,j)$$

iii) Variance is a measure of the spread of the probabilities in the co-occurrence matrix. The closer the probabilities are to the centre of the matrix the lower the variance.

$$VAR = \sum_{i=0}^{L} \sum_{j=0}^{L} (i-\mu)^2 \cdot P(i,j)$$

$$\mu = \frac{L}{2}$$

iv) Inverse Difference Moment or local homogeneity is given by:

$$IDM = \sum_{i=0}^{L} \sum_{j=0}^{L} \frac{P(i,j)}{1+(i-j)^2}$$

v) Entropy or texture non-uniformity. Gives high values when the elements in the matrix are large and low values when they are unequal.

$$ENT = -\sum_{i=0}^{L} \sum_{j=0}^{L} P(i,j) \cdot \log P(i,j)$$

vi) Correlation returns larger values as elements within the matrix become more similar.

$$COR = \sum_{i=0}^{L} \sum_{j=0}^{L} \frac{(i-\mu_x) \cdot (j-\mu_y) \cdot P(i,j)}{\sigma_x \cdot \sigma_y}$$

$\mu_x(\mu_y), \sigma_x(\sigma_y)$ are the mean and standard deviation of the row (column) sums of the matrix.

Laws[26] developed and investigated a set of texture features based on average values of local properties. A set of standard masks (both 3x3 and 5x5) define grey-level patterns in local pixel neighbourhoods of an image. The features produced represent the degree of affinity between these local pixel neighbourhoods and pre-defined patterns specifying local averaging, edge, spot, ripple, and wave detection. The statistics used for texture discrimination are the sum of the squared or the absolute values of the image pixels after

convolution with a specific mask has been performed. As with the construction of co-occurrence matrices the histogram of the image must first be equalised. If different size images are used the final statistics are normalised to the number of pixels convoluted. Fig.2.02 shows some examples of masks used by Laws.

Pietikainen et al[34] did further analysis of Laws masks. They found that the performance of Laws' features does not depend on the specific numerical values within the masks but upon their general form.

The main weaknesses of the Laws Mask method are:

i) It extracts features closely related to the grey-levels of the image. It may classify similar class textures differently due to the differences of illumination between them.

ii) There is a trade-off between computation time and structural information. To extract more structural information from an image larger masks are required and as these increase, the computation time increases rapidly.

iii) The method does suffer from the directionality problem.

3 x 3 Masks

L3xL3

1	2	1
2	4	2
1	2	1

L3xE3

-1	0	1
-2	0	2
-1	0	1

L3xS3

-1	2	-1
-2	4	-2
-1	2	-1

5 x 5 Masks

L5E5

-1	-2	0	2	1
-4	-8	0	8	4
-6	-12	0	12	6
-4	-8	0	8	4
-1	-2	0	2	1

E5S5

-1	0	2	0	-1
-2	0	4	0	-2
0	0	0	0	0
2	0	-4	0	2
1	0	-2	0	1

L5S5

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

R5R5

1	-4	6	-4	1
-4	16	-24	16	-4
6	-24	36	-24	6
-4	16	-24	16	-4
1	-4	6	-4	1

Fig.2.02 Examples of 5x5 and 3x3 Laws Masks

2.1.2.4 Shape and Geometry

As explained in chapter 1 it is hoped that these features will be useful for the gross classification of pollen. The objective of chapter 5 is to explore this area further so no more time is spent here on it.

2.1.3 Classification Procedures

This sub-section briefly describes how a feature vector consisting of variables that describe a certain characteristic of the pollen taxa (one of the Haralick measures for example) can be used to construct a classifier to separate each possible class of pollen. The first method utilises discriminant functions and has been widely used in previous pollen project research. The second method uses neural networks and has, as yet, been unexplored in this field. A more detailed analysis of neural networks can be found in chapter 6.

2.1.3.1 Discriminant Functions

The classifier employed in all previous research used a Fisher Linear Discriminant Function (see Weszka et al[47]) to separate every possible pair of classes from each other. Consider two classes C_i and C_j each with a set of feature

vectors $[x_{i1}, x_{i2}, \ldots, x_{ik}]$ and $[x_{j1}, x_{j2}, \ldots, x_{jk}]$ where k is the number of variables in the vector. Their optimal linear direction α_{ij}^{k} is given by:

$$\alpha_{ij}^{k} = (\Sigma_i^k + \Sigma_j^k)^{-1} \cdot (\mu_i^k - \mu_j^k)$$

μ_i = sample means of variable k in class i.

Σ_i^k = sample co-variances of variable k in class i.

This is calculated for all the samples in a training set of features and then used to construct a decision boundary, defined by:

$$B_{ij} = \frac{(\mu_i \sigma_j + \mu_j \sigma_i)}{\sigma_i + \sigma_j}$$

μ_i = the sum of the means of $\alpha_{ij}^{k} \cdot x_{ik}$ for each variable k in class i.

σ_i = the sum of the standard deviations of $\alpha_{ij}^{k} \cdot x_{ik}$ for each variable k in class i.

The class to which an unknown sample x is assigned is determined according to the following rules:

If $\sum_t \alpha_{ij}^{t} \cdot x_t > B_{ij}$ then assign to C_i.

If $\sum_t \alpha_{ij}^t \cdot x_t < B_{ij}$ then assign to C_j.

If $\sum_t \alpha_{ij}^t \cdot x_t = B_{ij}$ then indeterminate.

In practice the final case will never occur as these values are real numbers.

As there are more than two possible classes into which an unknown pollen taxa may fall a voting system is used. The feature vector of the unknown taxa is compared with each of the known class pair boundaries. In each case the one to which it is assigned has its vote counter incremented. As soon as all the class pairs have been considered the vote counter with the highest value represents the class into which the unknown is classified. Consider a case where sample x is from class 1 and that there are 4 classes. When x is compared with classes 1-2, 1-3 & 1-4, class 1 should be selected every time. But when it is compared with classes 2-3, 2-4 & 3-4 its selection should be essentially random. Thus the class 1 vote counter will have the largest value and therefore class 1 is selected.

A problem may occur if 2 or more class vote counters are drawn with the maximum value. In this case the above classification procedure is repeated with only drawn class pairs used. If classification is still not possible the system divides the vote between all drawn classes.

A leave-one-out scheme is used to test the success rate of the classifier. This involves removing each sample in turn

from the training set and classifying it using a discriminant function built from the remaining samples. This ensures that the training set is as large as possible and gives a good assessment of the true error rate. It does however increase the computational cost as a new classifier is constructed for each sample.

It is not necessary or practical to use all the variables available in a feature vector. A subset often exists that produces comparable results but with less calculation. Also, increasing the dimensionality of the classifier can lead to diminished classifier performance (Hand[21]). Thus a variable selection procedure is used to search for an optimal subset of variables.

Hotelling's T^2 statistic, a multivariate extension of the standard single variable t-test, is employed as a measure of separability between two classes. The T^2 statistic is useful for this type of search as it is strictly monotonic, ie the T^2 value of a subset of variables is less than or equal to that for the full variable set. It indicates whether any two pollen classes are statistically inseparable using the available data set by producing a measure of the distance between sample means normalised to dispersion within the samples:

$$T^2 = \frac{s_i \cdot s_j}{s_i + s_j}(\mu_i - \mu_j)' \cdot S^{-1} \cdot (\mu_i - \mu_j)$$

s_i = number of samples in class i.

μ_i = mean centroid vector of class i

(sample mean for each variable).

S = assumed common variance-covariance matrix.

In practice however $s_i = s_j = s$ and so the within-class scatter matrix, W, is used in place of S, as it is easier to calculate:

$$T_n^2 = \frac{1}{2} \cdot s \cdot (\mu_i - \mu_j)' \cdot W^{-1} \cdot (\mu_i - \mu_j)$$

where

$$W = \sum_{j=1}^{2} \sum_{i=1}^{n_j} (x_{ij} - \mu_j) \cdot (x_{ij} - \mu_j)'$$

n_j = number of sample points in class j.

x_{ij} = i^{th} point from class j.

μ_j = sample mean for class j.

There are a variety of methods that may be employed to search for the optimum variable subset. These may be grouped thus:

i) Exhaustive search methods. These are only applicable when a small number of variables are used and is therefore of little value here.

ii) Accelerated search. These consider all variable sets, but due to the monotonicity of T^2 does not explicitly evaluate all of them.

iii) Suboptimal stepwise methods. These produce the most rapid searches but do not always guarantee finding the best solution.

Due to the many classification combinations and speed of operation required the initial variable selection procedure employed is the suboptimal search using sequential backward elimination (described below). At later stages accelerated searches may be employed to improve the classifier performance.

Sequential backward elimination starts with a complete variable set and removes the variable that reduces the value of T^2 the least. In other words, T^2 is calculated for the complete data set with consecutive variables removed. The variable omitted for the maximum value of T^2 is then permanently removed from the data set. This procedure is repeated until the value of T^2 falls below a predefined separability threshold or stopping point. Both the absolute separability and the McKay[30] variable stopping point have been used in previous pollen classification work. The absolute separability value

has proved to produce the best classification rates and is extensively used in this thesis. The minimum T^2 stopping point used can be expressed as:

$$T^2_{min} = \frac{(2s-2)\cdot N \cdot F}{(2s-1-N)}$$

N = number of variables in the feature vectors

F = 5% level of the F-distribution with N & $(2s-1-N)$ degrees of freedom

$s > N$

A good source of further information on the above functions may be found in Hand[21].

2.1.3.2 Neural Networks

Over the last decade the study of artificial neural networks has come to the fore in computer science (Aleksander and Morton[1]). They attempt to mimic the computational architecture of the human brain (Rumelhart et al[38]), the objective being to incorporate intelligent functions such as learning into computers. Pattern Recognition is one of the primary areas of study within neural network research with classification being one of its fundamental objectives. A neural network, given an input pattern (feature vector) attempts to map it to a specific output category, thus classifying the input pattern.

As this is such a new research tool and therefore in need of much explanation a great deal of chapter 6 has been devoted to it. There the theories behind neural networks are explored along with an analysis of the areas within the current problem that might lend themselves to neural network solutions.

2.2 Review of Previous Pollen Project Research

2.2.1 Preparation Techniques

As well as automating the counting and identification of pollen, it is important to prevent a 'bottle-neck' occurring at the preparation stage. Forster and Flenley[12,13] attempted to automate the preparation of various pollen samples prior to analysis with the development of complex extraction and purification techniques. Their system successfully cleaned out large detrital (non-pollen) material and produced sub-samples of pollen taxa groups. Thus reducing the probability that detrital material would be misclassified as pollen at a later stage. This can be considered to be the initial separation or classification level (Fig.1.01). A more detailed account of the techniques may be found in Forster[12].

2.2.2 Digital Image Processing

The greatest contribution to the automation of pollen identification was that given in a Ph.D thesis by Mitchel Langford[25]. He introduced the discriminant function discussed above and used a branch and bound accelerated search. Using fresh pollen taxa he showed that it was feasible to separate pollen using textural criteria. All six Haralick measures described above were used with displacements of 1, 4 and 8 pixels. Several of the Laws masks were used and edge pair analysis was also tried but this proved to be less useful.

An M.Sc thesis by Helen Waterhouse[45] followed up Langford's work and tried to apply the Haralick measures to six fossil pollen taxa (3 of which were identical to those used by Langford[25].) This provides many problems as type material of fossil pollen is much harder to find than that of fresh pollen. It does however provide a realistic test of the potential of the final system. Unfortunately no separation was possible for any pair of fossil pollen taxa. Closer inspection of her results however shows that they are not as negative as they may at first appear. Very small images were used having dimensions of only 20x20 pixels; this is nowhere near large enough. Even when the images have been equalised to 16 grey-levels there is not enough data to produce a significant co-occurrence matrix for analysis. Langford[25] recommended images with dimensions of at least 64x64 pixels. Also constant displacement vectors of 1, 4 and 8 pixels were used. These may not have been the

best to separate these pollen. Also a constant magnification was used for all six taxa and this would have meant a loss of texture detail on the smaller pollen.

Several final year projects have also contributed to the overall research. Rutter[40] worked in parallel with Waterhouse[45] and obtained the same results. However close analysis of the software uncovered bugs that may well have prevented any separation at all being found. Gundersen[19] produced knowledge-based software that may be used for pollen classification. Beney[3] worked on pollen location.

2.3 Review of Problems Remaining to be Solved

Following the initial work described above there are still numerous problems that need to be solved. These have been raised in both the textural and the automation aspects of the problem. Also the value of other digital image processing techniques have not yet been explored. In the subsequent sections these problems are outlined.

2.3.1 Textural Feature Problems

Langford[25] obtained some very encouraging pollen classification results for six taxa. However the efficiency of the discriminant function declines for a greater number of pollen types as the number of comparisons required increases. More variables will also be required to discriminate between the extra classes. Thus the hierarchical classification scheme proposed in the first chapter uses the textural features at the lower levels where only a few taxa remain to be separated.

There are other drawbacks with pollen texture to be overcome. The main problem is that of automating the choice of the area of the pollen's surface to be sampled. This selection is made difficult by several features of the pollen surface which cause the texture to differ quite markedly. The system would have to deal with pollen grains in all possible orientations. Flattened grains may be viewed edge on and so present only a small area from which to sample texture. Also surface features such as colpi, pori and spines present problems as the texture changes rapidly around them. Similar problems can be found at fold edges as the surface is tilted, also towards the edges of the grain where it curves away giving an increasingly acute view of the pollen texture. Due to the effects of preservation, the problems as described above may be heightened in fossil pollen. It can be increasingly folded, squashed, pitted, holed and torn. Material may also be attached to the grain surface. Such areas need to be recognized and avoided when sampling.

The consistency of the pollen exine structure may not be as fixed as has been assumed. Pollen texture can be observed to vary with changing conditions of temperature and humidity as they mature within the anther. In some species of Fraxinus (Ash) the textural elements may vary in both size and shape over a single grain. Using the average of several discrete areas of the pollen surface at the classification stage may get around this problem. Some variation in textural features within a single pollen taxa may be due to subtle differences in pollen size which may cause the texture to be compressed or stretched. Normalising the displacement vectors used, to the size of the pollen, may account for this problem. The changes in the exine outlined above may be so subtle that texture is not significantly influenced. This is analysed further in chapter 6.

Spores present an additional problem in that they can often lose their distinctive outer layer, leaving a relatively featureless inner layer for identification. It may be necessary therefore to include both possible textures in the classification scheme.

2.3.2 Automation Problems

No work has yet been done on the hardware system. This may only be undertaken once the specific requirements of the software are known. The computer would have to be able to control all aspects of the microscope including stage movement, focusing, magnification and brightness/contrast settings.

2.3.3 Some Unexplored Techniques

There are still some areas of work that may prove to be useful that have, as yet, been unexplored.

If a method to detect the presence of features such as colpi or pori could be developed this would be extremely useful. This could act as an alternative gross classification of pollen taxa. If, for example, a pollen image contained two pori then it could not be classified to a taxa that has one or no pori, but could be to a taxa with two or more. The most obvious problem with such a detector is that it would also have to be able to tell the difference between preservation damage such as creases, folds and pits, and the colpi and pori.

Another area that may prove useful is that of pollen shape and size (Langford[25]). This area is explored extensively in chapter 5 of this thesis.

Chapter 3

Equipment and Samples

3.1 Introduction

In the course of this research the image capture system was changed. This chapter starts with a summary of the previous image analysis system and then describes the reasons for the change. The new system hardware and software are then briefly described. Finally, there is a description of the pollen taxa used and how they were prepared for examination.

3.2 Reasons for Change in System

Langford[25] and Waterhouse[45] both employed the same pollen image capture system. They used the Cambridge Instruments 600 Series SEM to obtain images of pollen. Black and white photographs were taken of images projected onto a fluorescent screen by the SEM. Once these were developed they were digitised by a video camera connected to a Matrox frame-store housed in a DEC PDP 11/23. Image files were then transferred onto a VAX 11/750 for analysis. Both Langford[25] and Waterhouse[45] called for a more direct link between the SEM, frame-store and computer. The method they were using was very time consuming and so limited the number of pollen taxa that could realistically be used. With so many transfers of image data there was a real danger of a loss of image quality. Firstly, the curvature of the fluorescent screen distorted the original images. Secondly, the way the photographs were taken and developed changed the picture quality. Thirdly, lighting of the photographs prior

to capture by the video camera was crucial; reflections from the photographs needed to be avoided. Magnification of the final image may also be inadvertently changed at any of the above steps. The system was also a far departure from an envisaged final integrated system.

So, for the above reasons it was decided to move to a PC-based image capture system. This allowed images to be captured easily and rapidly from the SEM. The system is outlined in the next section.

3.3 System Hardware and Software

The hardware of the system utilised the DT2803 Low Cost Frame Grabber (from Data Translation Ltd[5,6,7]) plugged into a IBM compatible PC. As IBM PC compatibles are so widely used no further description is required here. However a more detailed account of the frame grabber is necessary.

The DT2803 is a single board 256x256x6 bit (64 grey-level) frame grabber for real-time digital image processing. It plugs directly into any IBM or compatible PC expansion slot. A frame is digitised every $\frac{1}{30}th$ second from a video input. The scanning electron microscope (SEM) employed outputs a standard video signal that the DT2803 can digitise. The frame grabber then outputs monochrome images to a video monitor. The main drawback of the frame-store is that it has an aspect ratio of 3:2 causing

a certain amount of image distortion. This problem is dealt with in subsequent chapters. Hardware architecture includes both input and output look-up tables (LUTs) and programme control by a microprocessor.

VideoLab is the component software for the DT2803. It consists of a combined tutorial package (VideoTutor) and comprehensive subroutine library (VideoSub). Any Microsoft compiler may be used with the subroutine library to carry out basic manipulations of the image. The subroutine library includes functions to read and write pixels or regions to the frame-store, construct grey-level histograms of the image and store images to disk.

The Microsoft compiler used for image manipulation software is the Pascal compiler version 4. However, this compiler can be frustrating to use as software compiles and runs relatively slowly. As a result nearly all software for use independently of the frame-store is written with Turbo Pascal version 6.

The SEM used was the Cambridge Instruments 360 Series Stereo Scan Microscope, the newest available in the university. This is vastly superior to the 600 Series used by Langford[25] and Waterhouse[45]. Images are built-up within the SEM's own 1028x1028x8-bit frame-store before being output as a video signal which the DT2803 can capture and manipulate. Controls

on the SEM such as focus, magnification, brightness and contrast are all electronically controlled allowing minute adjustments to be made to any of the parameters.

A wide variety of software was updated and written for use with this new system. Image capture software, for example, was written that offered continuous displays of image histograms to allow the optimum contrast and brightness to be selected. Texture analysis software was upgraded and written for both Laws mask and Haralick measure methods. Edge pair methods were ignored as they were considered to be of little value (Langford[25]). Image manipulation functions such as averaging, median and Sobel operators were written to pass over images. Software relating to the methods discussed in the following chapters were also incorporated into the system.

3.4 Pollen Samples

Fossil pollen has the disadvantage that it can often be damaged or distorted. In some cases it may be difficult for a trained observer to classify it. This may result in the eventual use of a misclassified pollen in the construction of computerised classification rules. Fresh pollen has the advantage that it provides a pure sample and so no time need be wasted identifying each pollen before its image is captured.

Fresh pollen is also easier to prepare for viewing under the SEM. Thus to remove any unnecessary complications fresh pollen taxa were used.

The pollen taxa used originated from Pacific islands such as Fiji, Hawaii and Easter Island. They are commonly found together in fossil cores and so provide a realistic test of the system. Pollen from 12 species was successfully extracted from an initial sample of 25 flowering plants (double that attempted by Langford[25]). This could be built upon to produce a final system classifying over 50 different species. One aim of this investigation is to lay down some basic principles or techniques that may then be applied to larger groups of taxa. The following pollen taxa were used (examples of which may be seen in the plates indicated):

Bidens hendersonensis (Plate 1),

Canthium barbatum (Plates 2(a) and (b)),

Elephantopis mollis (Plates 3(a) and (b)),

Emilia sonchifolia (Plate 4),

Fitchia speciosa (Plate 5),

Macaranga graeffeana (Plates 6(a) and (b)),

Passiflora quadrangalaris (Plates 7(a) and (b)),

Pritchardia minor (Plates 8(a) and (b)),

Pseudoelephantopis spicatur (Plates 9(a) and (b)),

Senecio stokesii (Plate 10),

Spore, Fungal (Plates 11(a) and (b)),

Xylosma suaveolens (Plates 12(a) and (b)).

These taxa were obtained from two sources. The majority were taken from the Herbarium in the Department of Geography at Hull University. The remaining samples were from the collection at Kew Gardens, London.

3.4.1 Preparation

Pre-prepared samples were available suspended in oil, however the oil contaminated the SEM images and proved extremely difficult to remove from the samples. Thus the pollen was extracted directly from dried flowers.

The anthers and as little of the rest of the flower as possible were cut from the dried plants. These were then added to a centrifuge vessel containing hot Potassium hydroxide (KOH) and agitated using a glass rod. After a couple of hours in the hot KOH each sample was centrifuged at 2000rpm for approximately 7 minutes. The pollen then formed a solid cluster at the bottom of the vessel. The KOH was then poured off and replaced with distilled water. The sample was repeatedly spun and drained to remove all the KOH. Once the samples were suspended in almost pure distiled water they were mounted drop by drop on small heated glass coverslips. The water evaporated leaving the pollen behind. Any visibly large particles on the coverslips were removed with a pair of tweezers. Finally, the coverslips were stuck to SEM stubs and splatter-coated with either gold or carbon. They were then ready for analysis under the SEM.

A wide variety of images were captured for analysis. Between 18 and 20 sample images were taken for each taxa in each of the following groups:

i) Images of all pollen at a magnification of 1250x so that comparisons of relative sizes could be made.

ii) Images of each individual pollen at the maximum magnification where it was still possible to see the whole grain. Within each individual taxa the maximum magnification used was kept constant. For examples see all plates indexed by (a).

iii) Texture images of each grain at the minimum magnification that could be used without including any of the pollen edges in the image. For this group the magnification was allowed to vary from grain to grain. For example see all plates indexed by (b).

Chapter 4

Object Detection

4.1 Introduction

Separating objects from the background in SEM images is obviously the first step in the automatic analysis of its shape, size, texture etc. This may be achieved using a number of thresholding techniques. A threshold, k, is a grey-level value that divides an image into two classes (or populations). The background class, C_0, represents pixels with grey-levels $[0,\ldots,k]$ and object class, C_1, denotes pixels with grey-levels $[k+1,\ldots,L]$ where L is the maximum grey-level.

There are a wide variety of techniques that may be used to automatically set a threshold. These may be broadly divided into two groups; global and local. Global techniques use a single grey-level value to threshold the entire image and may be sub-divided into either point-dependent or region-dependent techniques. The point-dependent techniques assign a threshold using only the grey-level histogram of the image. A region-dependent technique however uses a local property within each pixel neighbourhood, the gradient for example. Local techniques split the image into sub-images and assign thresholds for each one.

Brightness and contrast settings on the Cambridge 360 SEM can be changed very easily. This allows the difference between the bright objects and dark backgrounds to be highly exaggerated. All images were taken with brightness and contrast settings at 44% and 17% respectively to take advantage of this facility.

4.2 Otsu Method

The first thresholding measure to be tested is that proposed by Otsu[33]. In his brief study of thresholding methods Langford[25] found this to be the most useful and suggested that an image should be convoluted with a large (9x9 or 11x11) smoothing filter before applying this technique. As it is a global point-dependent thresholding method an optimal threshold is set using the normalised grey-level histogram or probability distribution, p_i, of an image with no other a priori knowledge. The method employs discriminant analysis to evaluate the "goodness" of threshold and thus determines an optimal threshold. A measure of separability, $\sigma_B^2(k)$, is evaluated for each potential threshold of classes C_0 and C_1 thus:

$$\sigma_B^2(k) = \frac{[\mu_T \cdot \omega(k) - \mu(k)]^2}{\omega(k) \cdot [1 - \omega(k)]}$$

where

μ_T = total mean level of the original picture

$$\mu_T = \sum_{i=0}^{L} i p_i$$

$\omega(k)$= zeroth-order cumulative moment of the histogram up to the k^{th} level

$$\omega(k) = \sum_{i=0}^{k} p_i$$

$\mu(k)$= first-order cumulative moment of the histogram up to the k^{th} level

$$\mu(k) = \sum_{i=0}^{k} i p_i$$

The optimal threshold, k^*, is selected by maximising σ_B^2 thus:

$$\sigma_B^2(k^*) = \max_{0 \le k \le L} \sigma_B^2(k)$$

Otsu however assumed that the measure of class separability, $\sigma_B^2(k)$, was always unimodal. This is however not always the case. Kittler and Illingworth[24] showed that for certain ratios of object-to-background pixel populations, $\sigma_B^2(k)$, was multimodal and therefore the best threshold may not be set. A modification procedure was proposed to be used whenever the separability function was found to be multimodal.

If the separability function is bimodal, for example, there will be two candidate thresholds at k^1 and k^2. For each of these candidate thresholds the means of the two classes that they separate are calculated. The grey-level histogram value

at the candidate threshold point k' is compared with the values at the means μ_{i1} and μ_{i2} of the two populations formed by dividing the histogram at that threshold. Thus the candidate threshold that satisfies the following is accepted:

$$p_{k'} < p_{\mu i1}$$

and $$p_{k'} < p_{\mu i2}$$

$i = [1, 2]$

$p_{\mu ij}$ = the mean of the j^{th} population isolated by i^{th} candidate threshold

If it is not possible to reject a threshold at this point it is assumed that both thresholds are equally valid and that the image can not be binarised without losing meaningful information.

4.3 Combination of Methods

Here a threshold is set using two different methods. These are then combined to give a single optimum threshold that maximises the two methods. Wong and Sahoo[49] combined a uniformity and shape measure to produce a threshold that maximised both the uniformity and shape information in the resultant binary image.

4.3.1 Uniformity Method

Uniformity is a global point-dependent threshold measure, k^{*1}. It was introduced by Levine and Nazif[27] to assess the performance of a segmentation method. For every possible threshold value, k, the uniformity measure, $U(k)$, is calculated thus:

$$U(k) = 1 - \frac{\sigma_0^2 + \sigma_1^2}{F_1}$$

F_1 = positive normalisation factor

σ_i^2 = variance of class i given by:

$$\sigma_i^2 = \sum_{k \epsilon C_i} (k - \mu_i)^2 p_i(k)$$

C_i = pixel class i where $[i = 0, 1]$

μ_i = mean grey-level of pixels in C_i

The threshold, k^{*1}, is assigned to the maximum uniformity value.

4.3.2 Shape Method

Shape, the second threshold measure, k^{*2}, is a global region-dependent method as it includes information on the grey-levels that make up the shape features of an object (eg its edges). After all, it is the occurrence of edges that

separates the object from the background. The shape measure, $S(k)$, is computed for every possible threshold value using the formula:

$$S(k) = \frac{\sum_{h(x,y)} sgn(h(x,y)-\mu_N)\cdot\Delta(x,y)\cdot sgn(h(x,y)-k)}{F_2}$$

$h(x,y)$ = grey-level at location x,y

μ_N = mean grey-level of 8-neighbourhood of $h(x,y)$

F_2 = positive normalisation factor

$$sgn(x) = \begin{cases} +1, & if \quad x \geq 0 \\ -1, & if \quad x < 0 \end{cases}$$

$\Delta(x,y)$ = generalised gradient value, calculated thus:

$$\Delta(x,y) = \left[\sum_{i=1}^{4} D_i^2 + \sqrt{2} D_1 (D_3 + D_4) - \sqrt{2} D_2 (D_3 - D_4) \right]^{1/2}$$

$$D_1 = h(x+1,y) - h(x-1,y)$$

$$D_2 = h(x,y-1) - h(x,y+1)$$

$$D_3 = h(x+1,y+1) - h(x-1,y-1)$$

$$D_4 = h(x+1,y-1) - h(x-1,y+1)$$

As with the uniformity measure the optimum threshold, k^{*2}, is set at the maximum shape value.

4.3.3 Combining the Threshold Functions

The two threshold functions returned, $S(k)$ and $U(k)$ are easily normalised to produced probability density functions. The maximum value of the sum of both these functions is taken to be the threshold value, k^*. This value thresholds a binary image thereby maximising both the uniformity and the shape information of the image.

4.4 CYBEST Method

This global region-dependent method was proposed by Watanabe and the CYBEST Group[46] as a thresholding solution for cervical cancer cell pre-screening. The method is based on the differential histogram of an image. The first step is to calculate the grey-level difference, s_{ij}, of every pixel x_{ij} and its 8-neighbours x_{ij}^n (where $n = 1..8$) thus:

$$s_{ij} = \sum_{n \in N} d_{ij}^n$$

$$d_{ij}^n = x_{ij} - x_{ij}^n \quad \text{if} \quad (x_{ij} - x_{ij}^n) > 0, \quad d_{ij}^n = 0, \quad \text{otherwise.}$$

The differential histogram, $s(k)$, may then be calculated by summing s_{ij} over the whole image and then averaging by the number of pixels in the image with grey-level k. A threshold, k^{*1}, may now be set at the maximum value of $s(k)$ which represents

the grey-level with the greatest average differential. A second threshold, k^{*2}, is set by repeating the above method but by using an inverse form of d_{ij}^n, thus:

$$d_{ij}^n = x_{ij}^n - x_{ij} \quad if \quad (x_{ij}^n - x_{ij}) < 0, \quad d_{ij}^n = 0, \quad otherwise.$$

The final threshold, k^*, is set by combining k^{*1} and k^{*2} thus:

$$k^* = \frac{(k^{*1} + k^{*2})}{2}$$

4.5 Local Threshold Techniques

As well as the global techniques described above a couple of simple local techniques were tried. These attempt to produce binary images by setting individual thresholds for each 8-neighbourhood of pixels.

4.5.1 Gradient Direction

The first method uses the pair of Sobel gradient operators shown on the following page to calculate the directionality of the gradient at each pixel.

$$G_x = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \qquad G_y = \begin{matrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix}$$

G_x represents the change in gradient in the x direction and G_y the change in the y direction. If both G_x and G_y are equal to 0 (or less than a threshold), then there is no directionality at the target pixel. It can be assumed that if a pixel has no directionality then it must be a part of the image background, whereas a pixel with directionality is highly likely to be part of an object.

4.5.2 Edge Magnitude

This second method convolutes an image with eight edge detectors (see below) and sets each target pixel to the maximum value they return.

$$\begin{matrix} 1 & 0 & -1 \\ 1 & 0 & -1, \\ 1 & 0 & -1 \end{matrix} \qquad \begin{matrix} 0 & -1 & -1 \\ 1 & 0 & -1, \\ 1 & 1 & 0 \end{matrix} \quad \textit{etc...,}$$

After the convolution a preset threshold value is used to highlight the edges and blank the rest of the image.

4.6 Set A Priori

The final method tests the viability of having a single threshold value to binarise every image. The advantage of this is that it would make the process of constructing a binary image almost instantaneous. The proposition is made more realistic when the nature of the images is considered ie the high contrast between object and background. The best thresholds from a small number of images are averaged and then used to threshold all subsequent images.

4.7 Results, Discussions and Conclusions

The relative performances of the global thresholding methods can be seen in Fig.4.01.

The method proposed by Otsu[33] performed extremely well. There was no need for prior convolution with a large smoothing operator as Langford suggested. The images Langford[25] used had brightness and contrast levels which maintained texture quality. With no prior smoothing there was a danger that some edge elements would be missed but this problem was removed effectively in this study by the exaggerated brightness and contrast levels which reduced the texture influences. The smoothing convolution also tends to reduce the prominence of some edge features such as spines. The speed of this technique

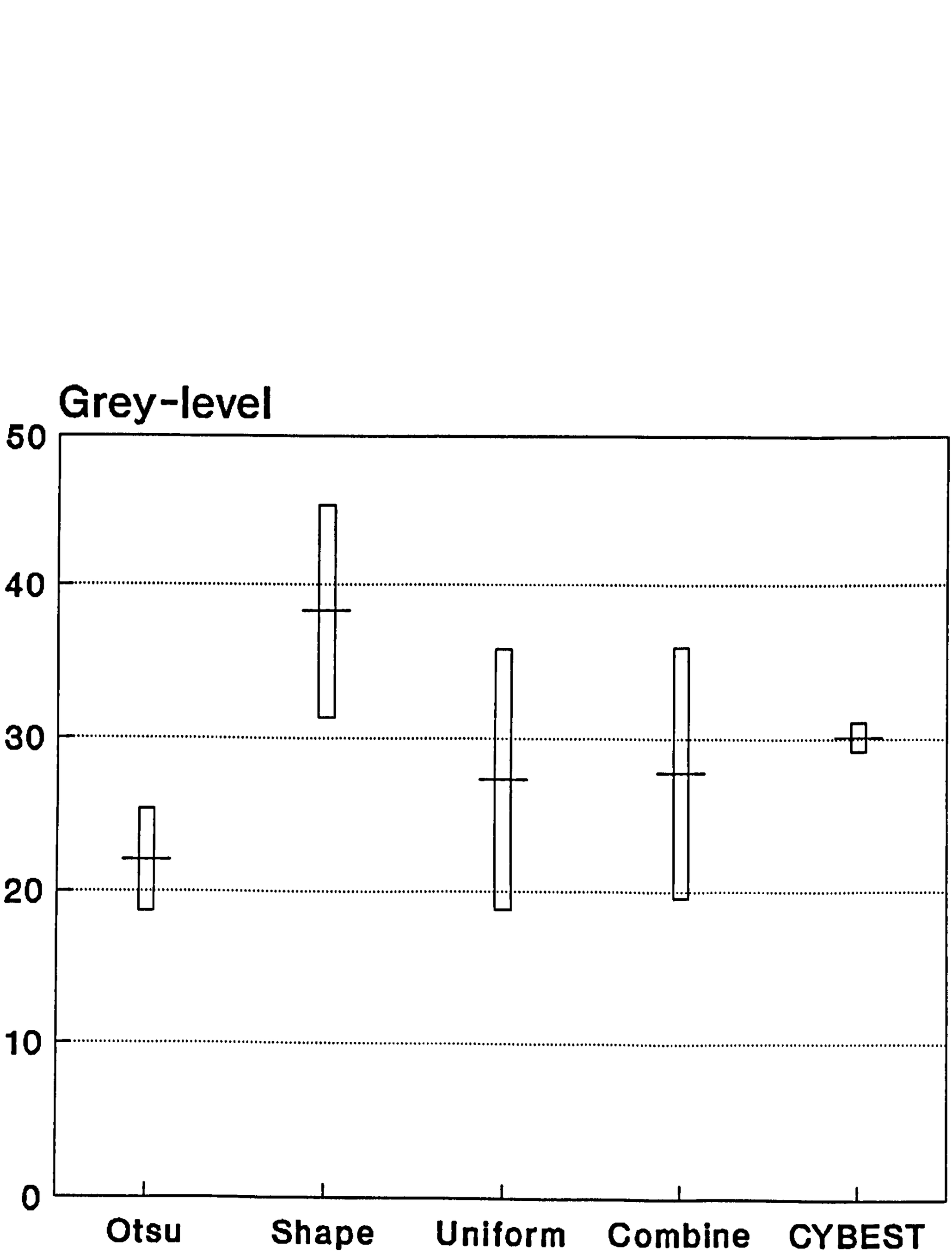


Fig 4.01 : Global Threshold Methods

was greatly enhanced as no time was wasted convoluting the image with a large operator. The method was in fact the fastest of all those used.

In the vast majority of cases it set good thresholds, with intact pollen edges and little noise from the background pixels. However, on a couple of occasions the Otsu method set a threshold that was a couple of grey-levels too high resulting in pollen with parts of their edges missing in the binary image. This happens where there are shadows on an object. The electrons in the SEM are emitted at a 7° angle from the detector, so shadows can occur due to the uneven illumination. While this is not a common problem it is annoying. A simple solution is to reduce the Otsu threshold set by a couple of grey-levels for every image. This would have little effect on the good quality images but would eliminate the shadow problem on the rest.

Not once in over 300 images was the Otsu threshold function multimodal. As a result no comment can be made on the modification procedure proposed by Kittler and Illingworth[24].

Both independently and together the Shape and Uniformity methods proved to be completely inadequate as the thresholds set were much too high. The uneven illumination had a marked effect on the images produced. Areas of the pollen with only slight electron shadows were lost completely. As well as setting inadequate thresholds these methods also took much longer to compute than any of the other methods explored.

The CYBEST thresholding method, while relatively fast did tend to set thresholds too high, resulting in the shadow problems discussed above.

Images produced using both local techniques were very encouraging. Both methods, while not as fast as the Otsu method, did work quite rapidly. Their great strength was that they always picked out the complete pollen edges as they were less susceptible to local changes in grey-level than the global techniques. However, they were more susceptible to noise in the background which did tend to clutter the images somewhat. With the correct edge and directionality threshold this problem is reduced. We are now back to a thresholding problem. What level of edge or directionality do we use as a cut-off? To set this automatically would add more time to the production of a binary image and to have a set threshold for both these techniques would make them inflexible. While these are promising techniques more work needs to be done to set meaningful thresholds.

The thresholds set a priori had limited success. Shadows, as ever, were a problem and the method's inflexibility in responding to them was its greatest weakness.

The Otsu method was selected as the best all-round thresholding technique and was subsequently used in the production of binary images to be analysed in the next chapter.

Chapter 5

Shape and Geometric Analysis

5.1 Introduction

Binary images constructed using the thresholding techniques described in the previous chapter are now analysed to extract information about the shape and location of objects within each image. The method chosen to store both the shape and location of each object is that proposed by Freeman[14]. As well as an edge encoding system Freeman[14-18] produced a wide variety of shape analysis functions to describe the form of an object. This chapter describes the initial separation of pollen from detrital material and then examines a variety of shape analysis functions.

5.2 Separating Detrital Material and Pollen

No matter how many filtering procedures a pollen sample goes through prior to analysis it will always contain a certain amount of detrital material. Inorganic solids are easily removed chemically and so present little problem. However, it is the organic plant material that is the problem as attempts to classify it as pollen are almost certainly doomed to failure.

Solutions to this problem may be derived from both the preparation and image processing techniques. Research by Forster and Flenley[12,13] on preparation techniques has produced some very helpful results. They have produced a system where detrital material is largely removed from samples leaving

only smaller material behind. This makes the subsequent task of separating pollen and detrital material from each other using image processing techniques much easier. Unfortunately, the equipment used by Forster and Flenley[12,13] was unavailable and therefore large detrital material has been removed manually from some images leaving only the smaller particles. As the detrital material is in the form of relatively small particles it may be detected on the basis of size. As we shall see in the next section, any object that has a shorter perimeter than a set threshold is considered to be detritus and is ignored.

A chemical analyser attached to an SEM may also prove useful as a solution to this problem. By analysing x-ray absorption of target areas of a sample an evaluation of its chemical constituents may be made. This would detect non-organic material with little difficulty, however, and more importantly, its ability to separate sporopollenin (a specific pollen macro-molecule) from lignin or cellulose (other plant macro-molecules) has not been tested. It may also be possible to classify detritus on the grounds of texture although these methods may add an unnecessary amount of complexity to the problem.

A secondary problem is one of pollen overlap. If overlapping pollen are not separated there is a great chance that the classification scheme will fail as pollen size, shape and texture data could be confused. Preparation techniques offer a solution to this problem by increasing the dilution of the original

sample and thereby reducing the probability of overlap. Also if the samples are kept moving as they are mounted there is limited time for them to become aggregated.

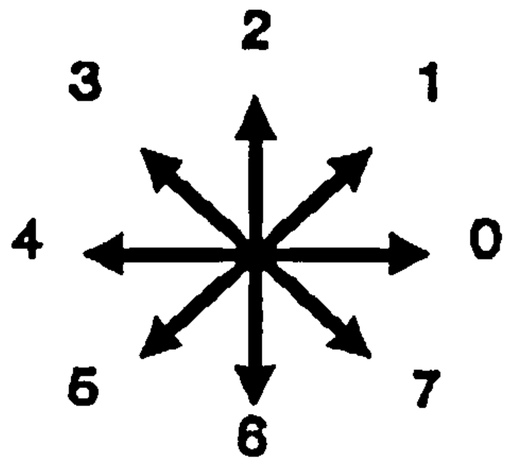
5.3 Edge Encoding using Freeman Chain Code

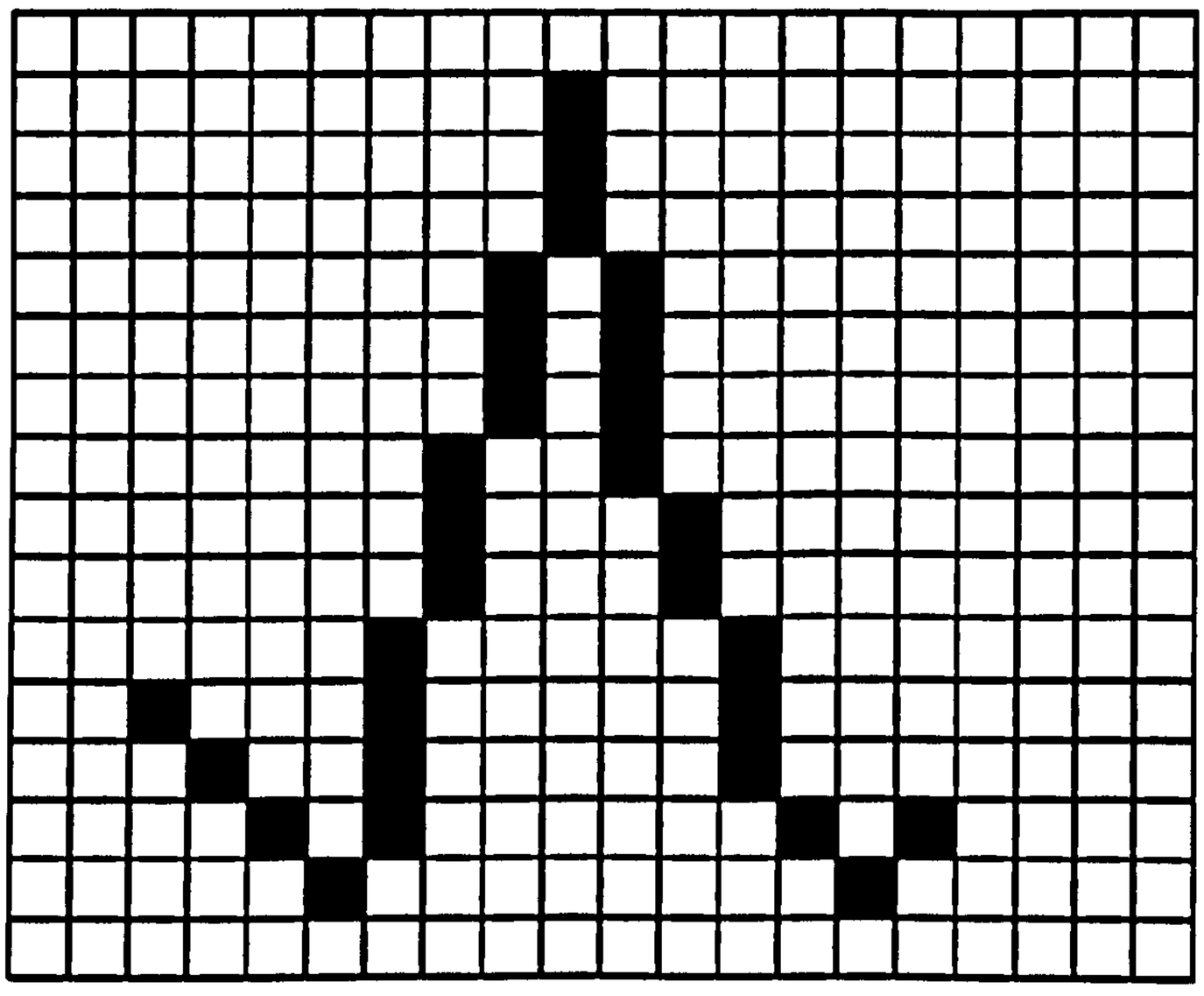
Rather than storing the (x,y) co-ordinates of every pixel on the pollen binary image edge, the direction in which each successive edge pixel lies is recorded. As each pixel has 8 possible neighbours the position of the next may be stored as a digit from 0 to 7 where 0 represents an adjacent pixel along the positive x-axis and the digits 1 through 7 represent consecutive 45° turns in an anti-clockwise direction. Fig.5.01 illustrates the coding scheme applied to a short edge segment. The starting point of the code is termed the initium and the end the terminus. Each digit refers to one directed line segment known as a link that represents a unique x and y component, for example a code of value 1 describes a change in co-ordinates of (1,1). A chain is a line structure made up of a number of links and can be represented thus:

$$A = \alpha_1 \alpha_2 .... \alpha_n$$

A = chain of length n.

α_x = link number x.





A = 7771222122122122667666767667771

D1 = 00210011011011040110011110102

D4 = 3330001101104334000010110 3

Fig.5.01 Chain Code Generation

Employing the methods described in the previous chapter binary images of pollen were made. A border one pixel thick was 'wiped' black. This prevented the subsequent object search from fruitlessly trying to trace a boundary that was not contained wholly within the image. It would be quite simple to include a portion of code that detected an object that was not entirely contained within the image under investigation and caused the next image captured to incorporate it.

The process of extracting the objects and their chain codes from the binary images is describe in the steps outlined below:

1. Start the scan from the top-left of the image (the origin).

2. At each successive point check for an object ie a pixel with the value of the maximum grey level.

3. Once an object pixel is located its co-ordinates are stored.

4. The rest of the object is traced and stored thus:

   a) Object maximum-minimum location array is set to out of range values. The array stores the co-ordinates that enclose the object and is used to delete it at a later stage.

   b) The Freeman location counter is set to 4 (ie it points back to the last non-object pixel).

c) From the current pixel a search begins for the next object pixel at the pixel pointed to by the Freeman location counter.

d) If no edge is found the Freeman location counter is incremented by a modulo-8 addition of 1. This produces an anti-clockwise search of pixels around the current pixel. A check is also made that the object is not a single point - this would result in an infinite loop as the search cycles about it. The procedure jumps to 5 if a single piont is found.

e) If an edge is found the value of the Freeman location counter is stored.

f) The maximum-minimum location array is updated.

g) The current pixel is updated to that pointed to from the last current pixel location.

h) The Freeman location counter is set to point to the last pixel searched before the current pixel was found.

i) Steps (c) to (h) are repeated until the original co-ordinates are reached again.

5. Then using the maximum-minimum location array the object is deleted thus preventing further chain codes being produced

for the same object.

6. If the edge just traced is longer than a previously set threshold length (eg a chain more than 100 links long) then it is stored to disk or is used in further shape analysis procedures. Objects with shorter chains are ignored as detritus.

7. The scan continues from the original co-ordinates of the object just located as described from 2 to 6. This continues until the bottom-right of the image is reached.

5.4 Image Distortion

It is important when looking at shape that the image being used has a 1:1 relationship with the 'real' world. In other words the aspect ratio of the image is 1:1. The frame-store used in this study however has a 3:2 aspect ratio. This resulted in the images used being distorted in such a way that they were stretched along the vertical axis. By deleting every third horizontal line from the image and then compressing the remaining lines together this distortion was rapidly removed. For any future work a frame-store with a 1:1 aspect ratio is recommended.

5.5 Basic Information Extracted from Code

As well as the basic chain coding scheme, Freeman[14-18] produced numerous algorithms that extracted shape information from the chain code. A selection of the most useful of these is described below.

5.5.1 Area Enclosed by Chain

The following formula may be used to calculate the area enclosed by a chain.

$$Area = \sum_{i=1}^{n} a_{ix}\left(y_{i-1} + \frac{1}{2}a_{iy}\right)$$

y_i = the y ordinate of link i.

$y_i = y_{i-1} + a_{iy}$

a_{ix} = the x component of the link a_i

(ie if $a_i = 0$ then $a_{ix} = 1$ & $a_{iy} = 0$).

As closed chains are used the ordinate of the initium y_0 can be arbitrarily selected. This formula calculates the area encircled in a clockwise direction. The result has a negative value for anti-clockwise encirclement. The direction of enclosure is easily removed by using the absolute value of the result.

5.5.2 Length of Object Perimeter

The length of an object's perimeter can be calculated very rapidly using the following equation:

$$L = n_e + [n_o \cdot \sqrt{2}]$$

n_o = number of odd-valued chain code links.

n_e = number of even-valued chain code links.

5.5.3 Moments of Inertia

The first moment of inertia of an area enclosed by a chain code along a given axis (in this case the x-axis) is calculated thus:

$$M_1^x = \sum_{i=1}^{n} \frac{1}{2} a_{ix} \left[y_{i-1}^2 + a_{iy} \left(y_{i-1} + \frac{1}{3} a_{iy} \right) \right]$$

a_{ix} = the x component of the link a_i

(ie if $a_i = 7$ then $a_{ix} = 1$ & $a_{iy} = -1$).

y_i = the y ordinate of link i.

5.5.4 Centroid of the Chain

The centroid of a chain, or its centre of gravity, may easily be calculated given the first moments of inertia about both major axes and also the area enclosed by the chain. The location of the centroid with respect to a specified axis is given by the ratio of the moment about this axis to the enclosed area. Thus the centroid co-ordinates $(\bar{x},\bar{y})$are:

$$\bar{x} = \frac{M_x^1}{Area}$$

$$\bar{y} = \frac{M_y^1}{Area}$$

The absolute values of both the moment and area values will need to be used to give valid co-ordinates within the frame-store co-ordinate system.

5.5.5 Maximum & Minimum Dimensions

The maximum and minimum diameters of an object can be easily calculated from a chain code. The distance from the initium to the link at the half-way point along the code is calculated using Pythagoras' theorem. Subsequent calculations are made using pairs of links around the chain . The maximum and minimum distances encountered are stored for future analysis. This method assumes a certain amount of symmetry in the object being

measured. Fresh pollen is generally symmetrical and is therefore suitable for measurements using this method. Fossil pollen may be more asymmetric due to preservation damage and this method may require further refinement. Regions of the code where the maximum and minimum dimensions are found may require closer scrutiny to check that no other combinations of link pairs produce larger or smaller diameters.

5.5.6 Combination of Shape Measures

The basic geometric measures described above may be combined to produce further features. Ratios of measures such as object perimeter length to area or maximum to minimum diameter may be used. A measure of the compactness of the object may be calculated using the following equation:

$$Compactness = \frac{Perimeter^2}{Area}$$

5.6 Moment Invariants

These may be used to extract features from images and are invariant under image rotation and reflection. Since the introduction of moments by Hu[23] they have had varying degrees of success in applications such as aircraft identification (Dudani et al[10]), scene matching (Maitra[29]) and character

recognition (Casey[4]). The moments are calculated using the second- and third-order central moments (μ_{pq}) of an image ($I(i,j)$) given by:

$$\mu_{pq} = \sum_i \sum_j (i-\bar{x})^p (j-\bar{y})^q I(i,j)$$

$\bar{x}$ and $\bar{y}$ are the co-ordinates of the centre of gravity calculated above (see section 5.5.4). The co-ordinate system of the image is shifted so that the origin coincides with $(\bar{x},\bar{y})$ thus:

$$\mu_{pq} = \sum_i \sum_j i^p \cdot j^q \cdot I(i,j)$$

The following equations represent the seven low-order invariant moments most commonly used:

$$\phi_1 = \mu_{20} + \mu_{02}$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$$

$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2]$$
$$+ (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[(3\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

$$\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$$

$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2]$$
$$-(\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[(3\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

It must be noted that moments $\phi_1, \ldots, \phi_6$ are all invariant under rotation and reflection. However, ϕ_7 is sign sensitive to reflection, but its magnitude is unchanged and so the absolute value of this moment is used.

The moments calculated above can have a very large dynamic range causing the classification system described in chapter 2 to be unusable as the within-class scatter matrix cannot be inverted. Tien[42] suggested that it is more practical to use the logarithm of the moment magnitudes. Invariant moments (ψ_i) are therefore calculated thus:

$$\psi_i = \log|\phi_i|, \qquad i = 1, 2, \ldots, 7$$

By substituting normalised central moments (η_{pq}) for the central moments (μ_{pq}) in the invariant moment equations Hu[23] produced moments that are invariant under scale changes. Normalised central moments are calculated thus:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$$

$$\gamma = \frac{p+q}{2} + 1$$

μ_{00} = image area.

Maitra[29] proposed a further adaptation to the moment calculations to produce six functions invariant under contrast changes. The functions (β_i) are given by:

$$\beta_1 = \frac{\sqrt{\phi_2}}{\phi_1}$$

$$\beta_2 = \frac{\phi_3 \cdot \mu_{00}}{\phi_2 \cdot \phi_1}$$

$$\beta_3 = \frac{\phi_4}{\phi_3}$$

$$\beta_4 = \frac{\sqrt{\phi_5}}{\phi_4}$$

$$\beta_5 = \frac{\phi_6}{\phi_4 \cdot \phi_1}$$

$$\beta_6 = \frac{\phi_7}{\phi_5}$$

The different moment invariant functions may be applied to object boundaries, binary images or even the raw images of objects. Moments calculated from the object boundary contain some information not carried by moments taken from either raw or binary images (Dudani et al[10]). Information about the high-frequency part of an image is contained disproportionately in the moments derived from the object boundary. In other words, minute details of the pollen shape are best characterised by moments calculated from the pollen boundary ie directly from its chain code. Coarse structural features on the other hand are best described by moments calculated from binary or raw images. These moments also have the advantage that they are less susceptible to noise. A combination of moments from both boundary and binary or raw images will maximise the shape information extracted from an object.

5.7 Shape from Boundary Tracing

Boundary tracing techniques employ the centre of gravity (see section 5.5.4) as a standard reference point within an object to produce a "centroidal profile". A plot is made of the distances from the centre of gravity of an object to each boundary point in sequence. The maximum distance is chosen as the starting point of the trace as well as being a normalising factor to reduce all distances to the range 0 to 1. The distance

1 is the maximum distance and 0 at the centre of gravity. If a fixed number of edge points are sampled the resultant plot is scale and rotation invariant.

This study uses two different methods to define which edge points to sample. The first samples points at equal distances around the object perimeter using the "true" length (see section 5.5.2) of the edge between each successive point. Alternatively, points that subtend equal angles to the object centroid are used. If, for example, 100 points are sampled then points at angles 0°, 3.6°, 7.2°,... etc are employed. Angles may be easily calculated using the cosine rule:

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc}$$

A = angle subtended to centroid from edge points.

a = distance from point 1 to point 2.

b = distance from centroid to point 1

c = distance from centroid to point 2

The main difference between these two methods is the distribution of their sample points around the object perimeter. The first method has an even distribution of points around the perimeter, whereas the equal angle method has higher sample densities on parts of the edge closest to the centroid.

Due to the nature of this method the classification techniques used for the earlier methods cannot be used. An unknown shape is classified by matching its profile with a set of standard shape profiles. The closest match should occur with the most similar standard shape profile. Moore & Webb[32] present a set of standard pollen shapes often found in nature. The set of seven standard pollen shapes that most closely resemble the pollen classes used are shown in Fig.5.02. These standard shapes produced the profiles in Fig.5.03 (equal distance method) and Fig 5.04 (equal angle method). The closeness of the match between an unknown pollen shape and a standard is determined by the square of variance σ_{xy} between them:

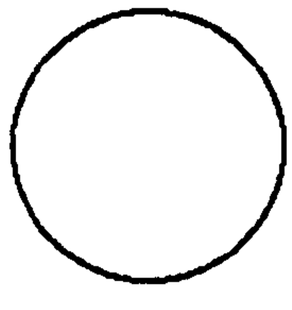
$$\sigma_{xy} = \sum_{i=1}^{N} (x_i - y_i)^2$$

N = number of sample vectors.

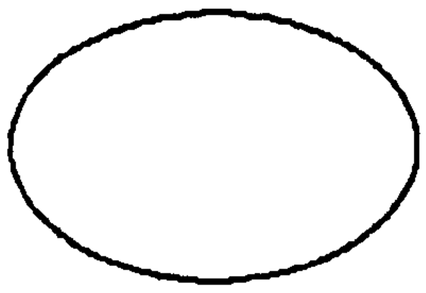
$x_i = i^{th}$ element of unknown pollen shape.

$y_i = i^{th}$ element of standard pollen shape.

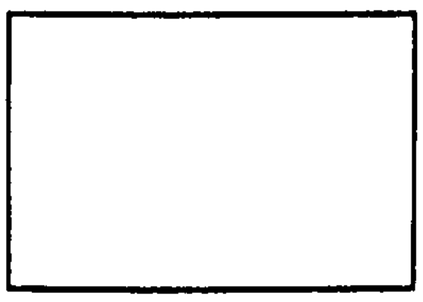
The standard shape that produces the lowest σ_{xy} when compared to an unknown is the shape into which the unknown is classed. Using profiles ordered with their maximum distances first, may not always produce the best matches, therefore σ_{xy} is calculated with the unknown profile ordered to each of the top ten longest distances in turn.



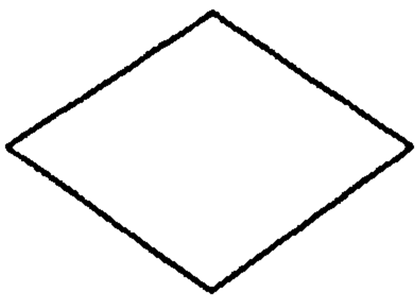
Circular



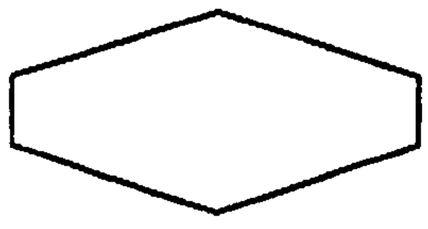
Elliptic



Rectangular

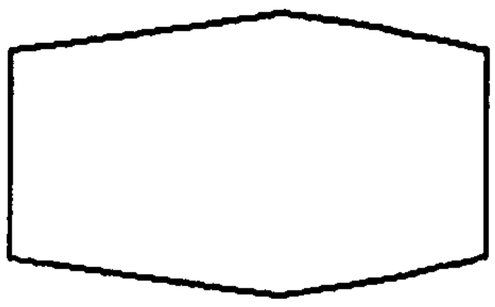


Rhombic acute (diamond4)



Rhombic acute truncate (diamond5)

Elliptic acuminate acute (lens)



Elliptic truncate (pillbox)

Fig.5.02 Standard Pollen Shapes

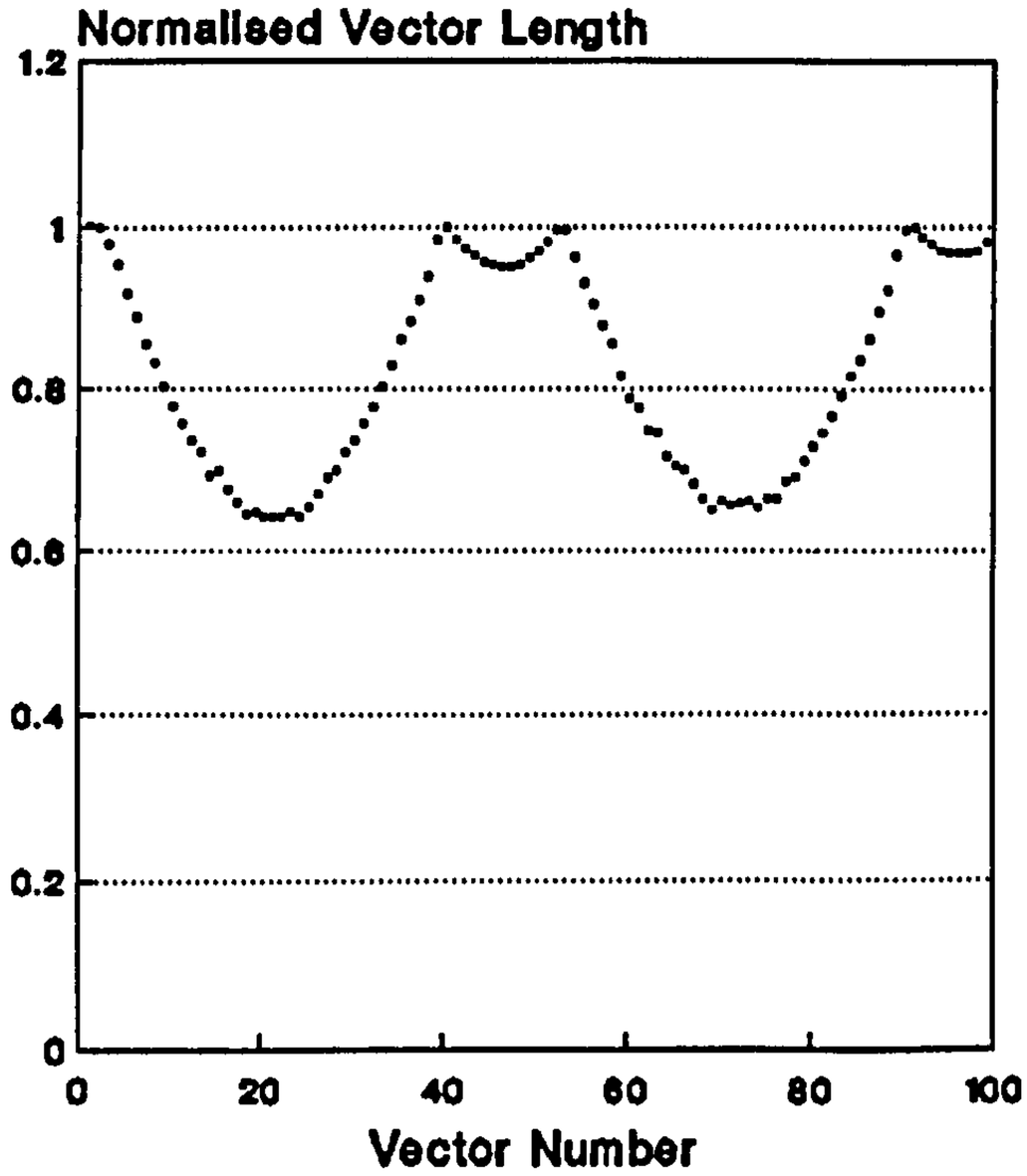


Fig.5.03(g) Elliptic truncate (Pillbox)

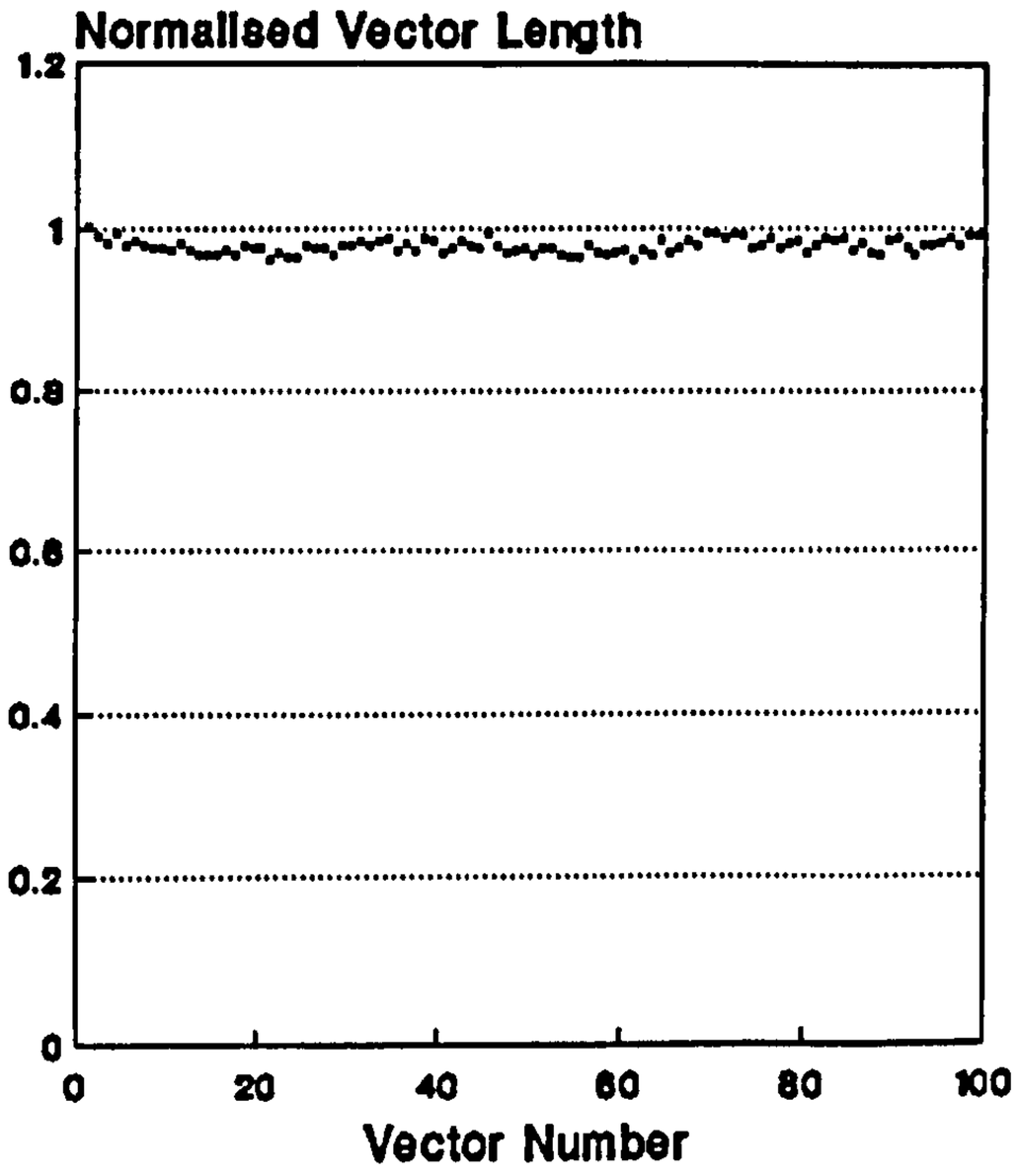


Fig.5.03(a) Circle

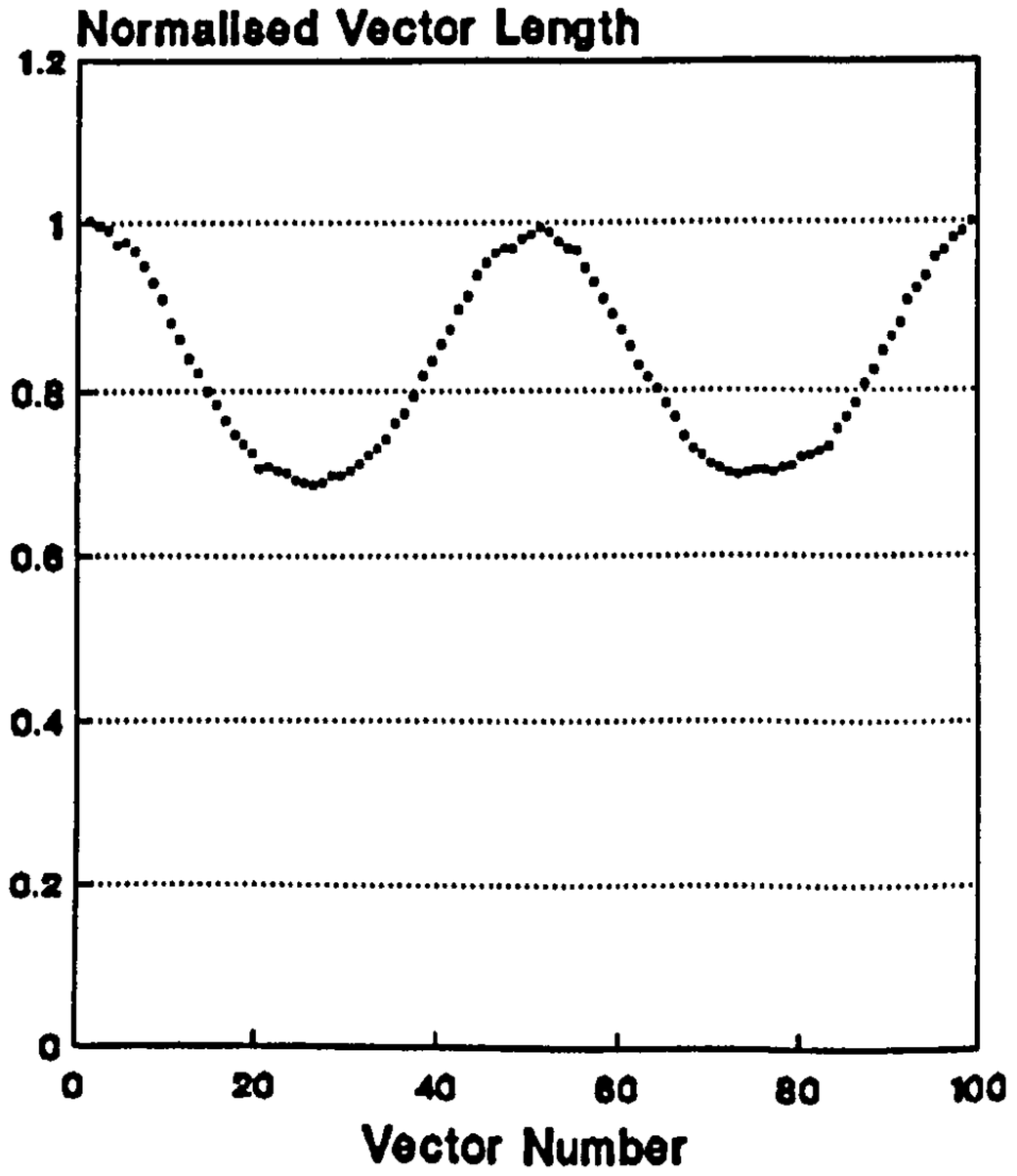


Fig.5.03(b) Ellipse

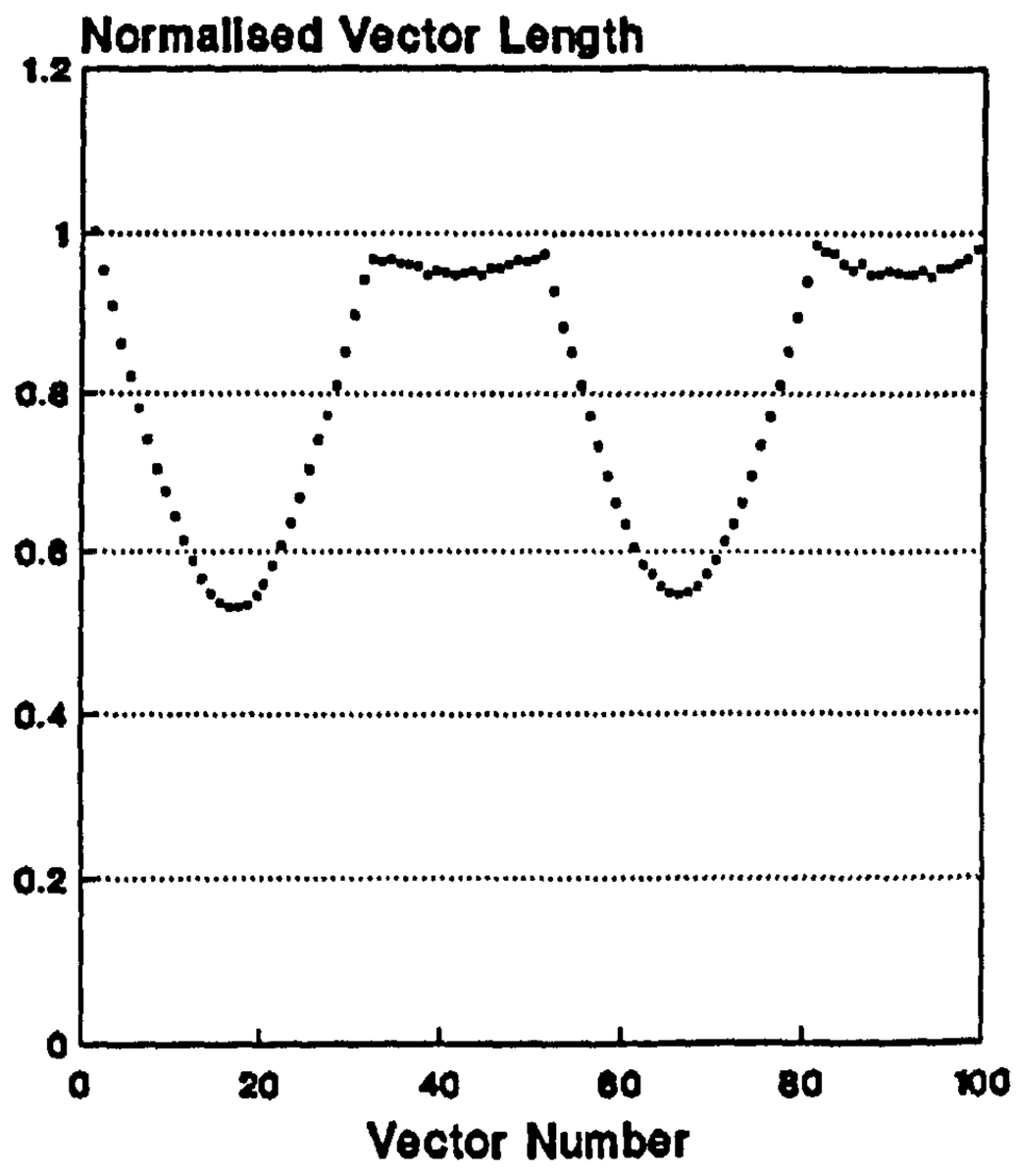


Fig.5.03(c) Rectangle

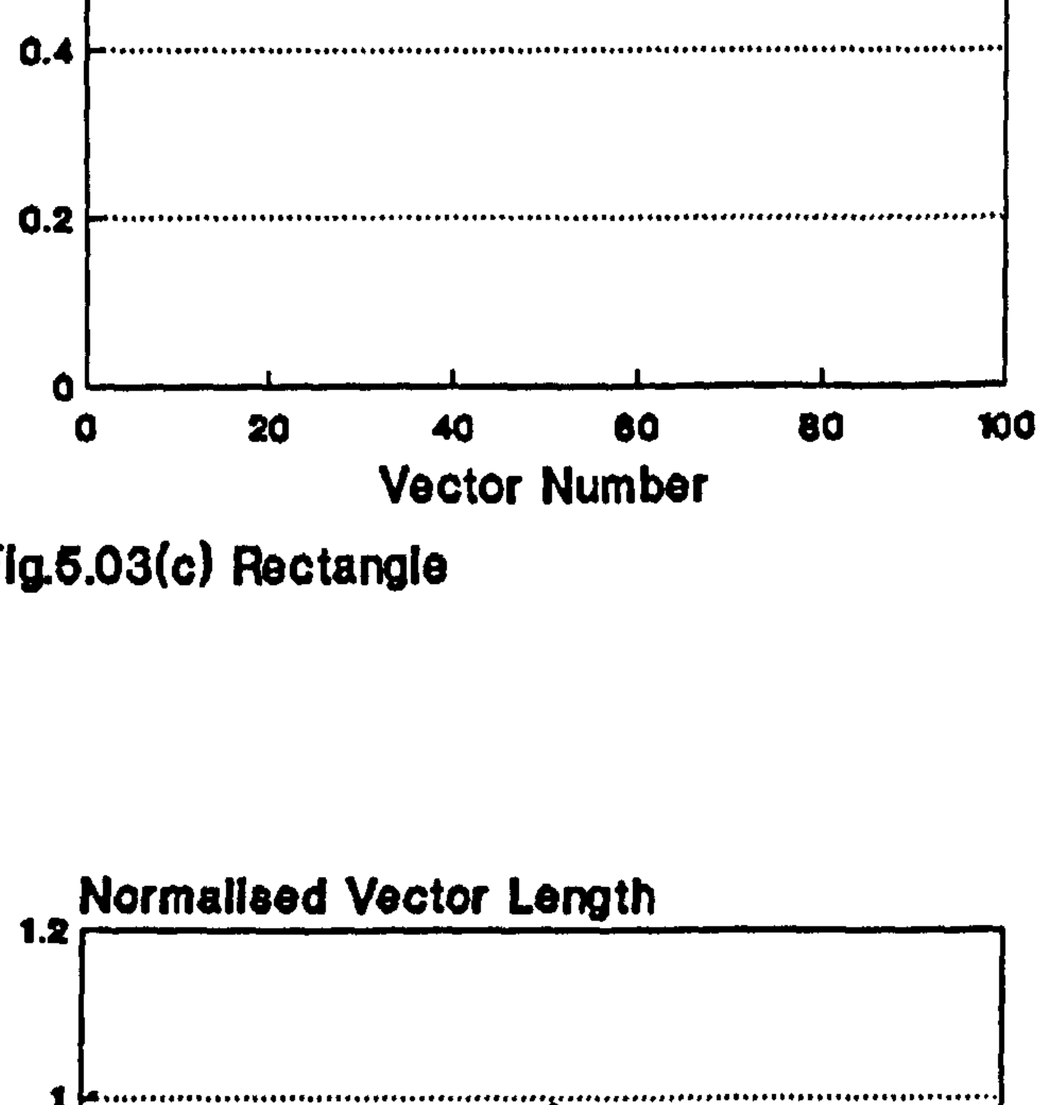


Fig.5.03(d) Rhombic acute (Diamond4)

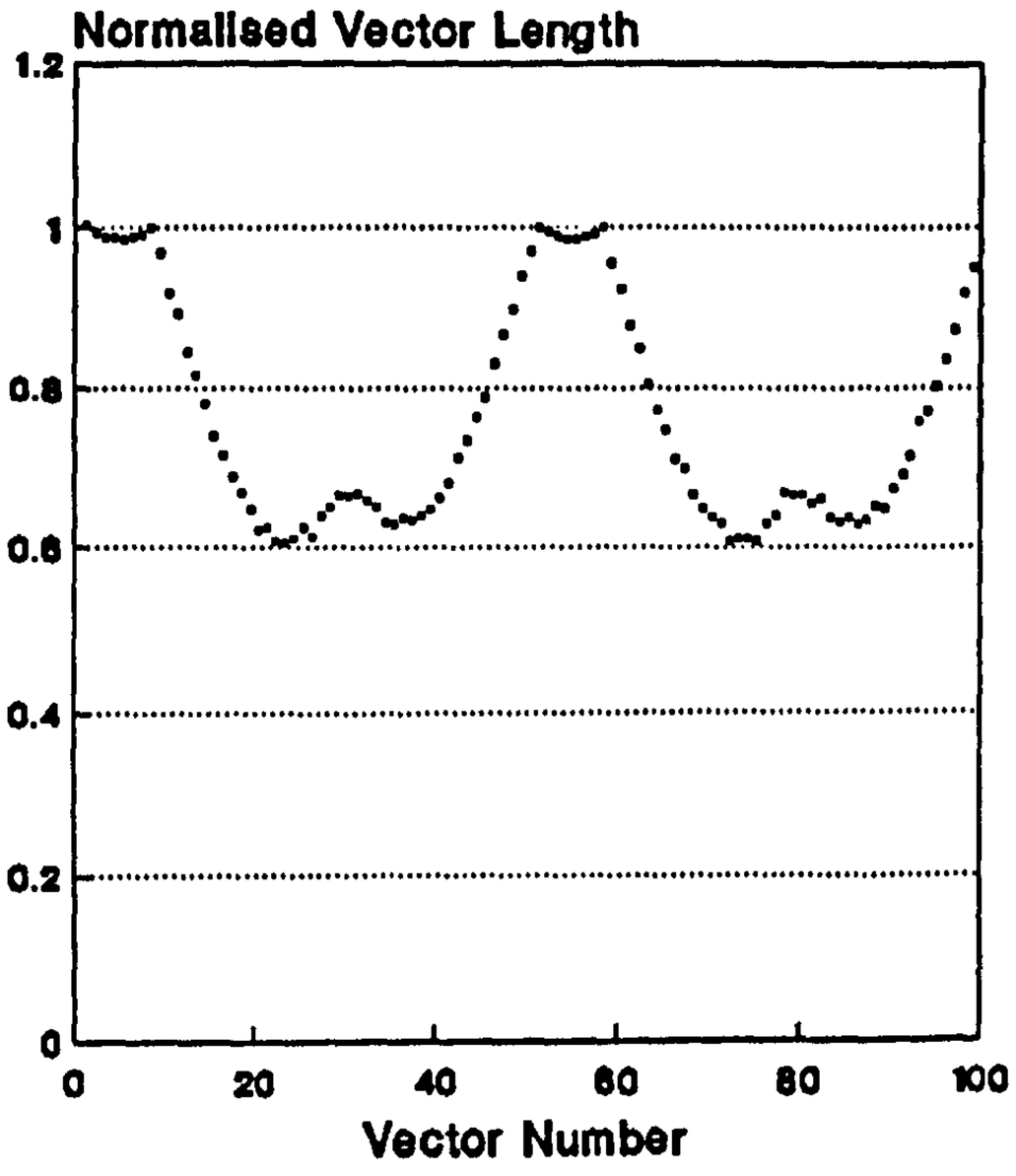


Fig.5.03(e) Rhombic acute truncate

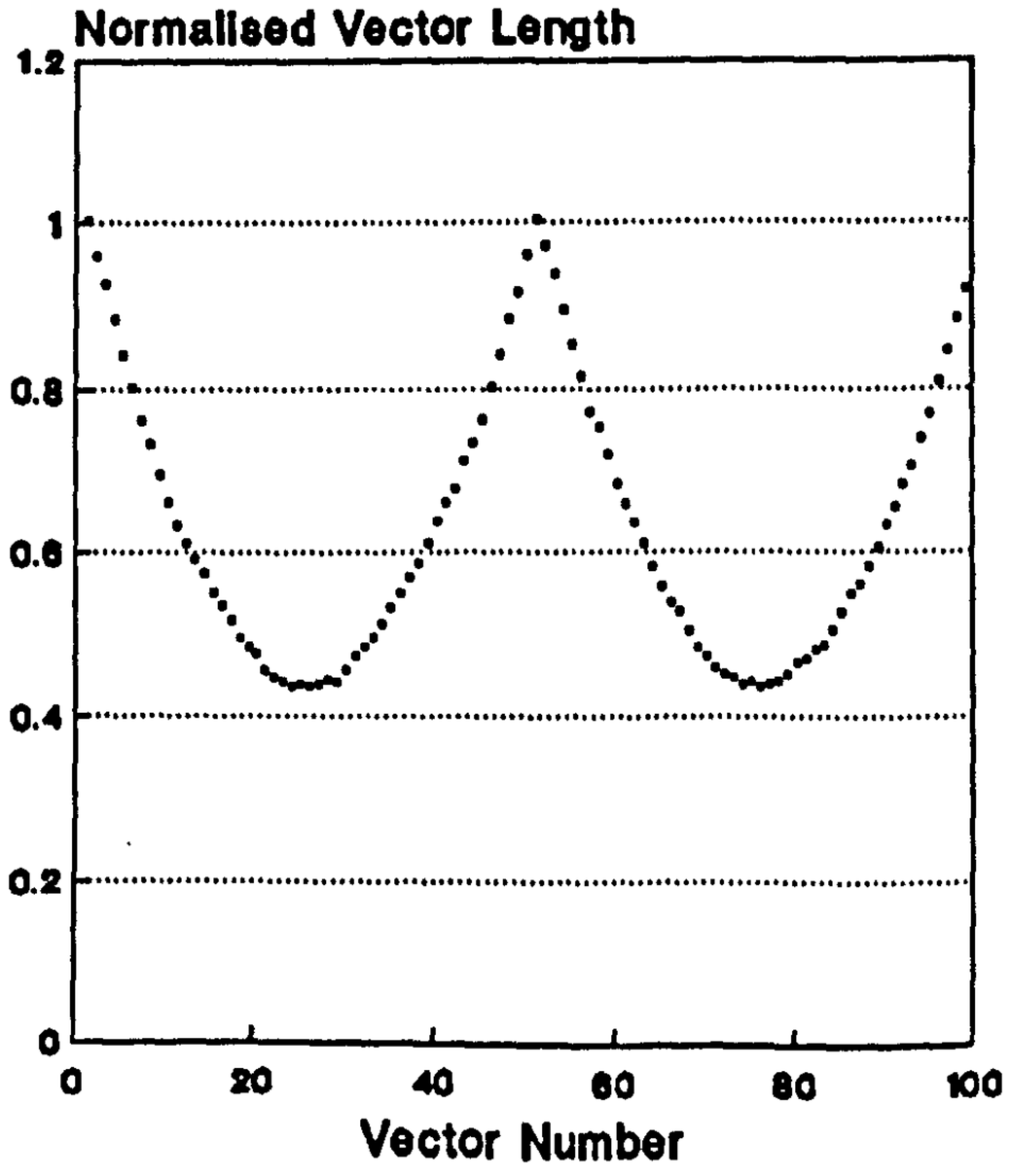


Fig.5.03(f) Elliptic acuminate acute

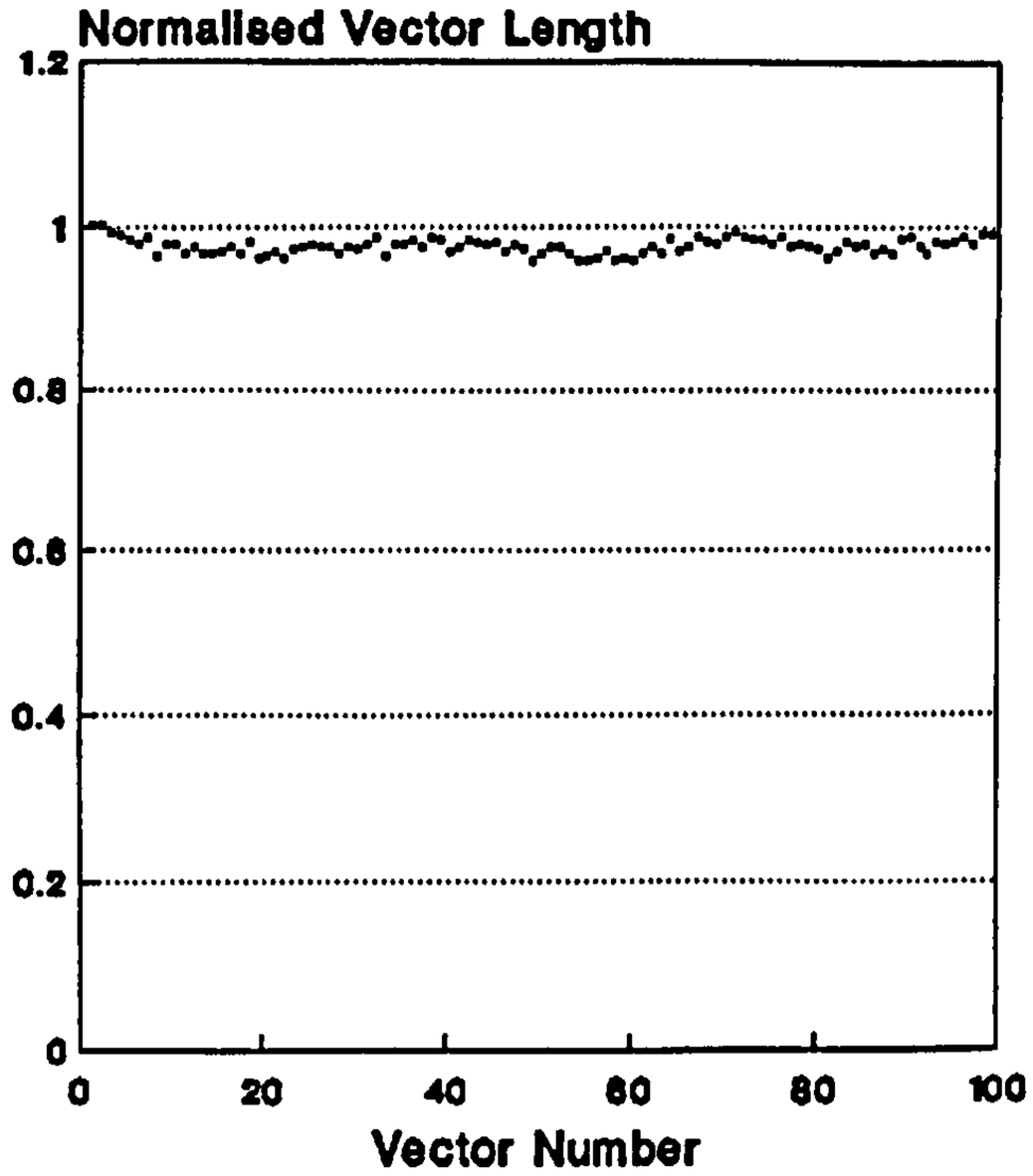


Fig.5.04(a) Circle

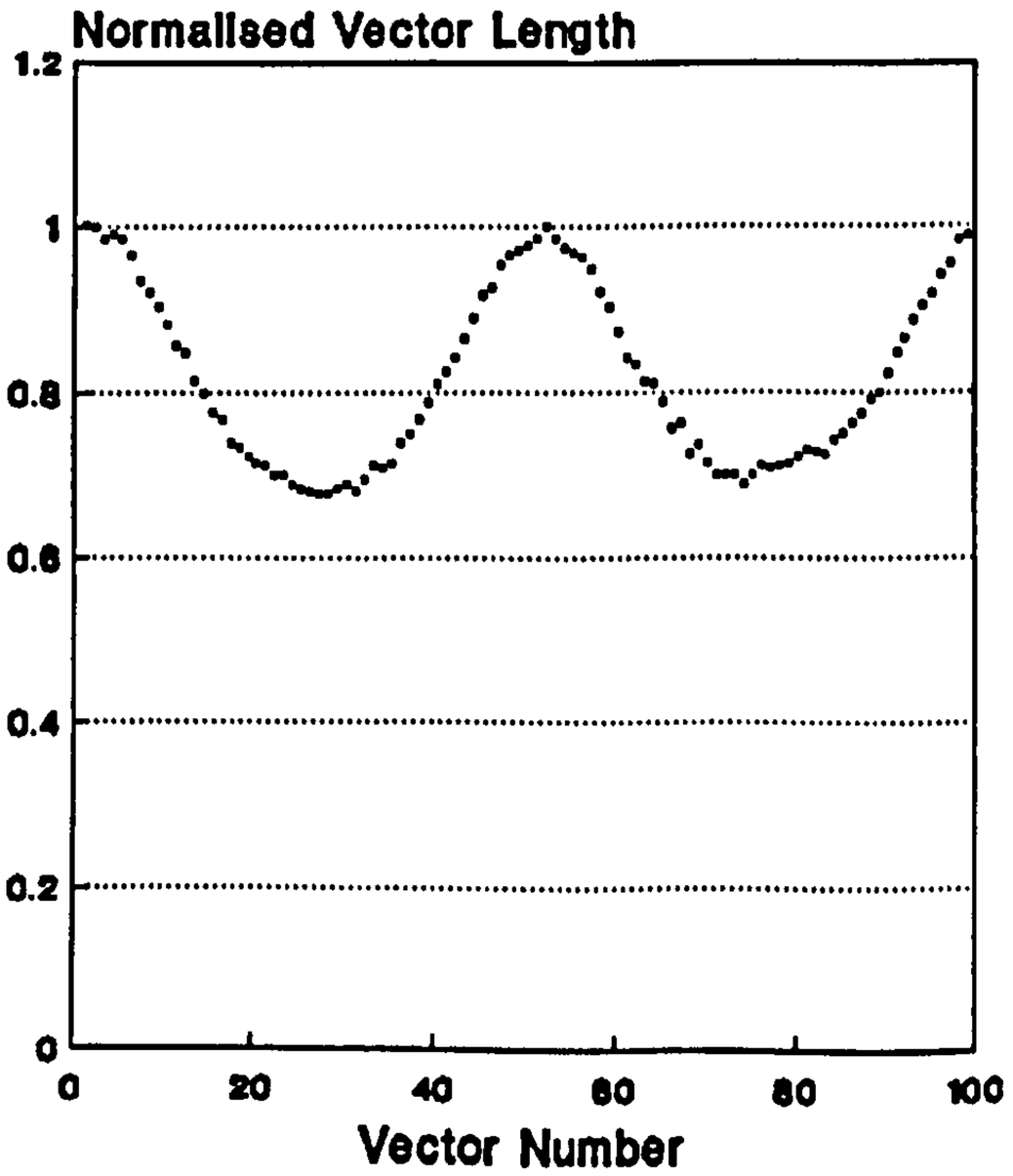


Fig.5.04(b) Ellipse

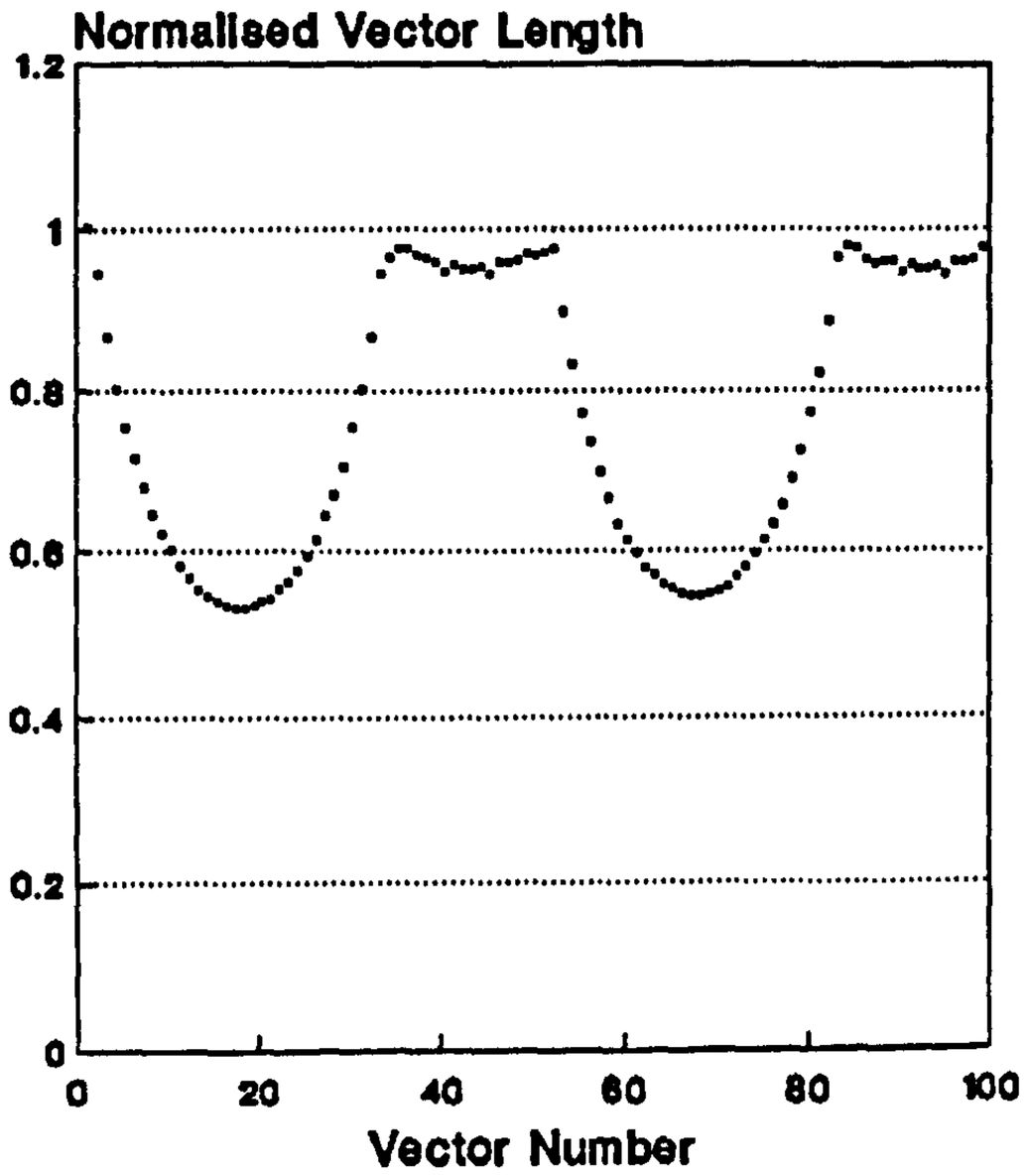


Fig.5.04(c) Rectangle

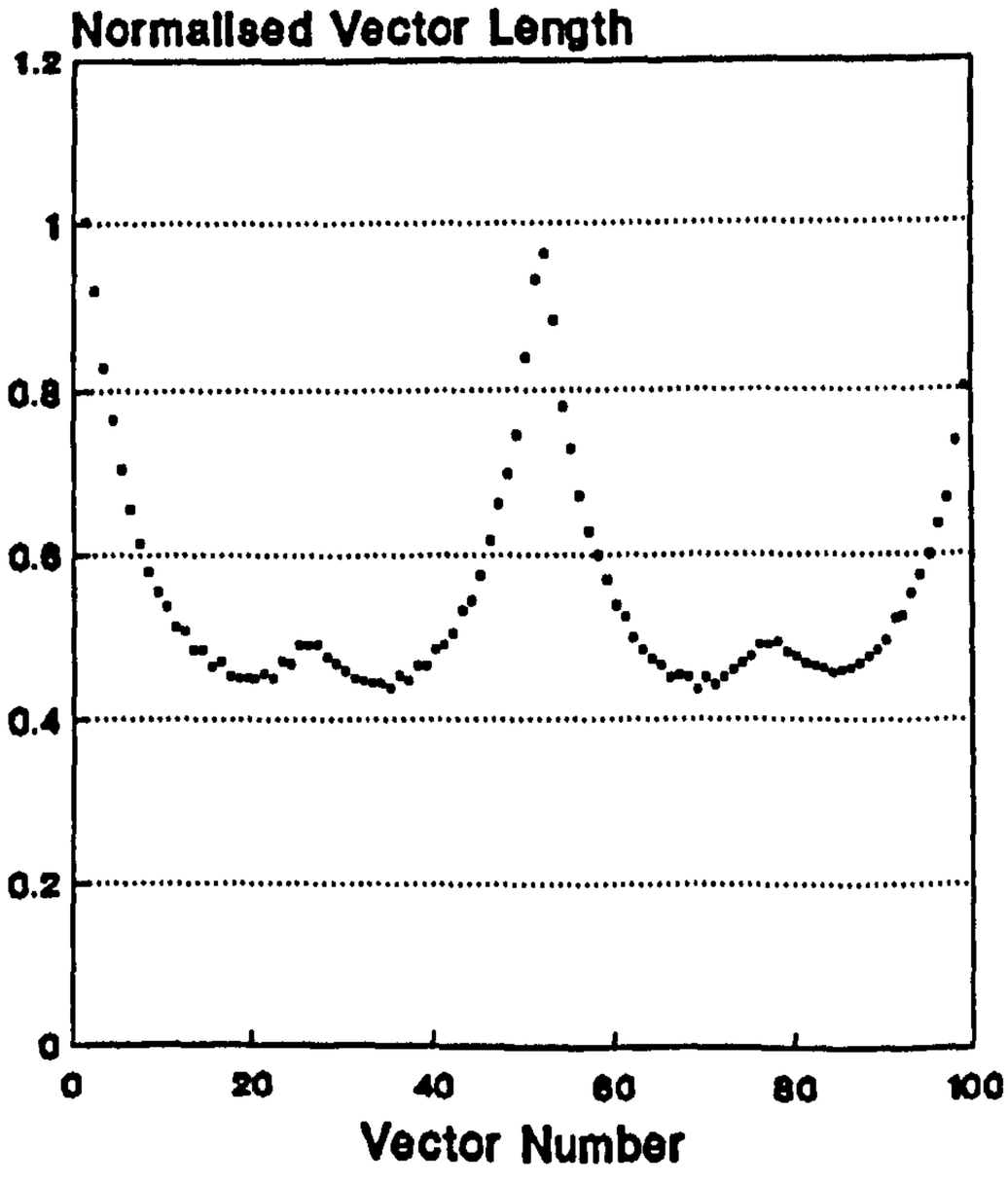


Fig.5.04(d) Rhombic acute (Diamond4)

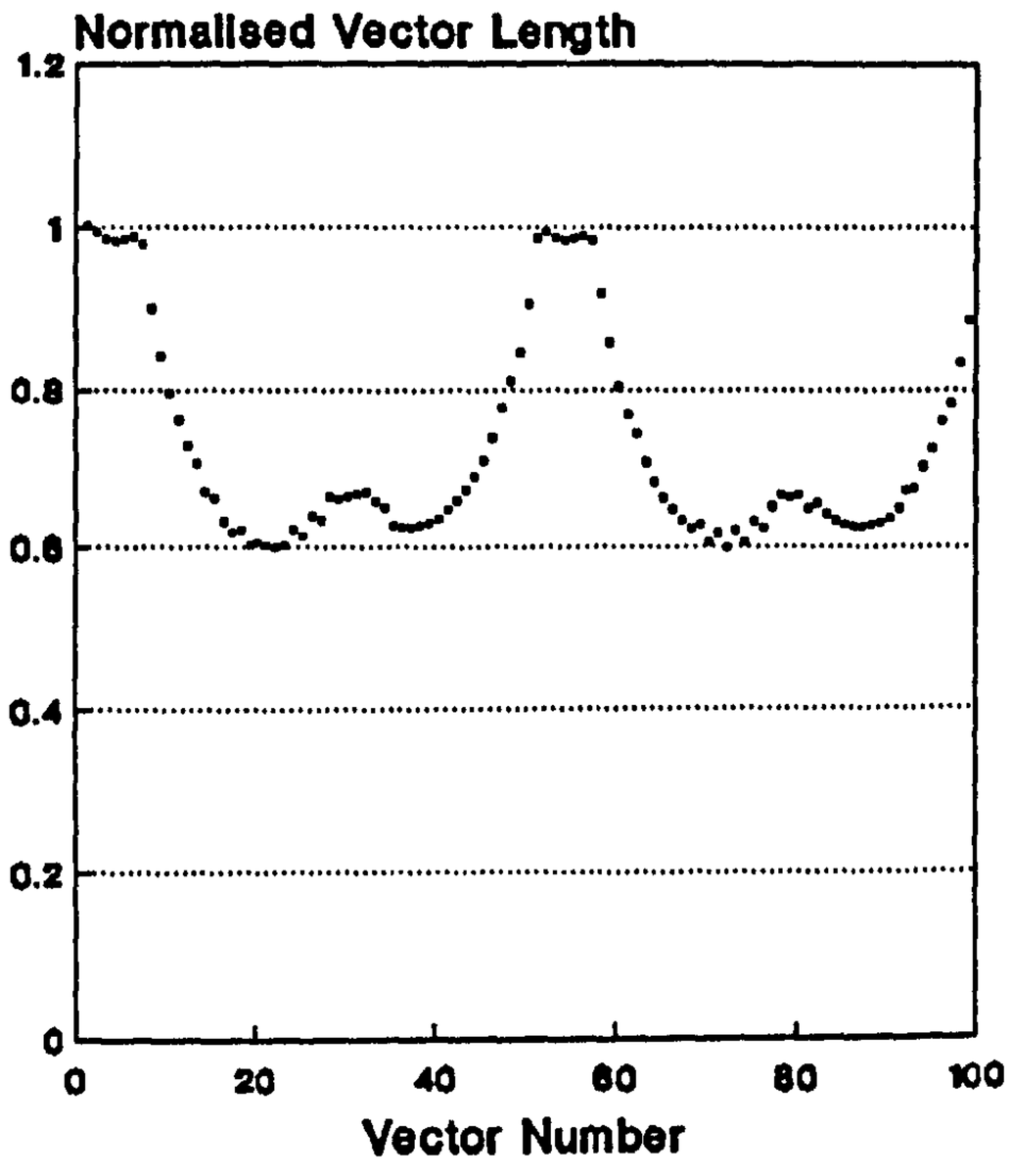


Fig.5.04(e) Rhombic acute truncate

Fig.5.04(f) Elliptic acuminate acute

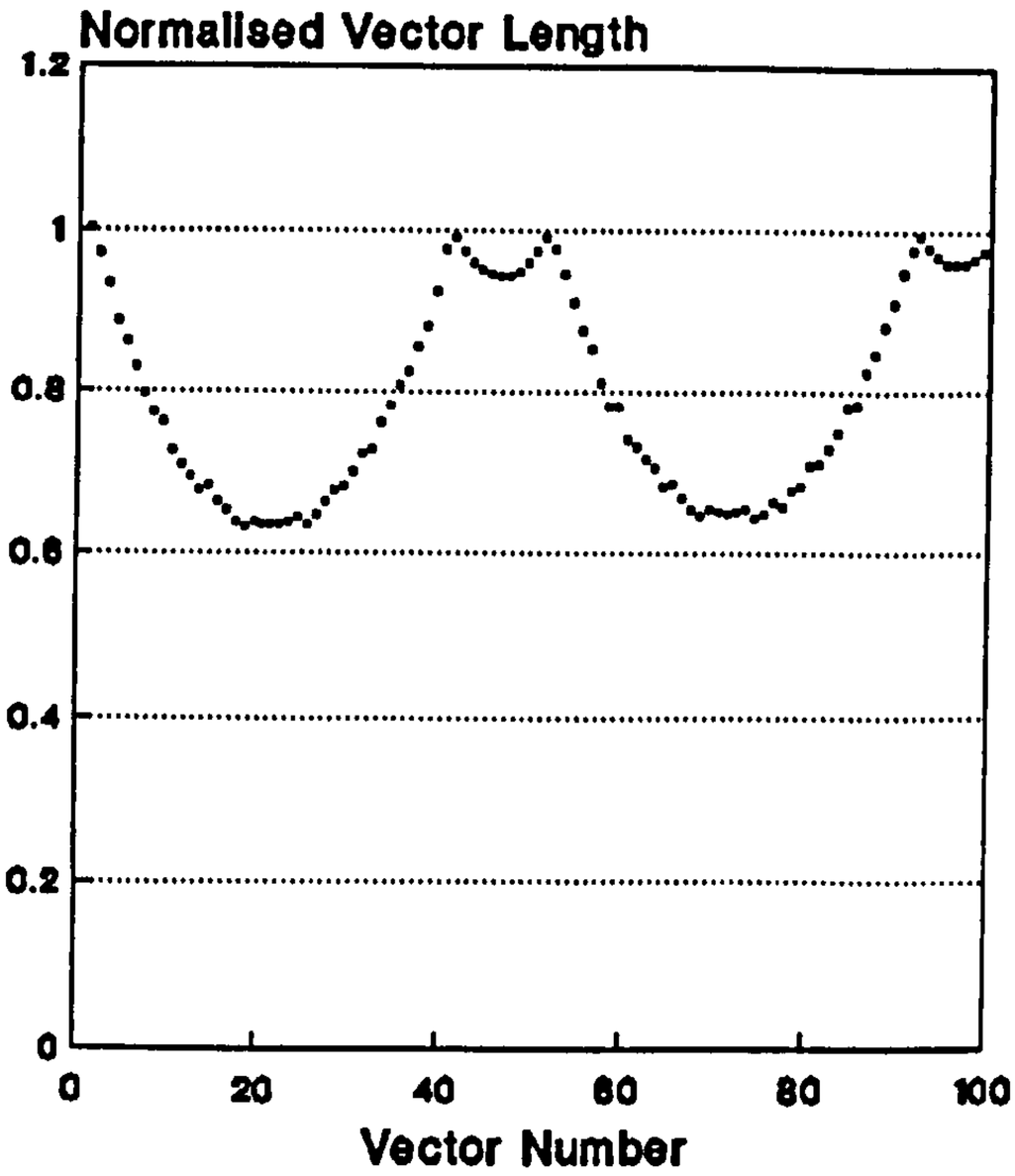


Fig.5.04(g) Elliptic truncate (Pillbox)

5.8 Chain Autocorrelation

The autocorrelation function of a chain (A) was suggested by Freeman[16] and is given by:

$$\theta_{AA}(j) = \frac{1}{n}\sum_{i=1}^{n} \cos(\min(|a_i - a_{i+j}|, |(a_i - a_{i+j}) - 8|))$$

$j = 0, 1, 2, \ldots, n$

n = number of links in chain A .

$a_i = i^{th}$ link of chain A .

When using this formula it is assumed that the chain is a continuous loop ie the link following the terminus is the initium.

As the number of elements in the autocorrelation function is dependant on the size of its original chain code the function is sampled at regular intervals. It is hoped that the sampled plot produced will be distinctive for each standard pollen shape. Fig.5.05 shows the sampled autocorrelation functions for the standard pollen shapes shown in Fig.5.02. The same classification methods used for the boundary traces in the previous section are used here.

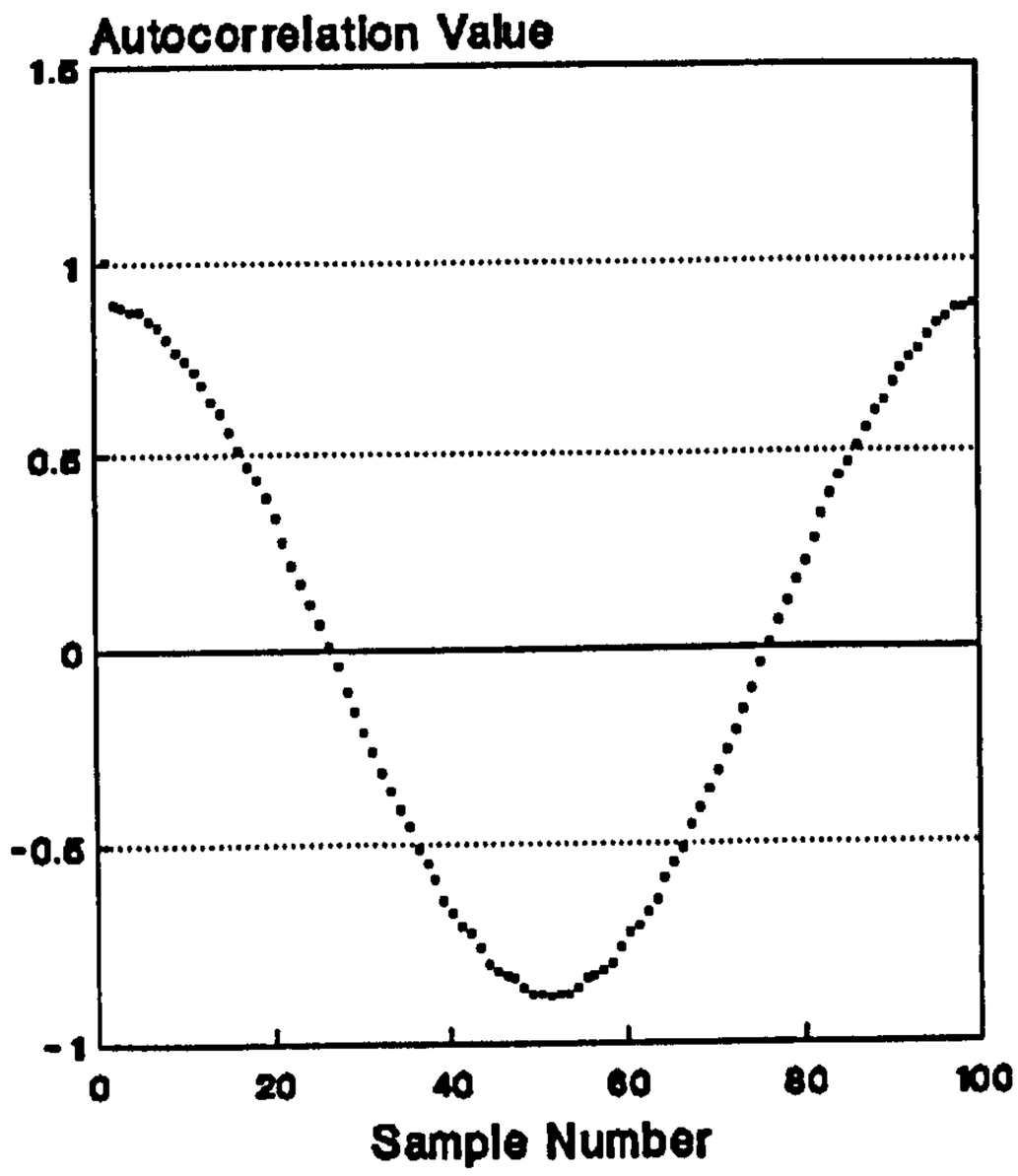


Fig.5.05(a) Circle

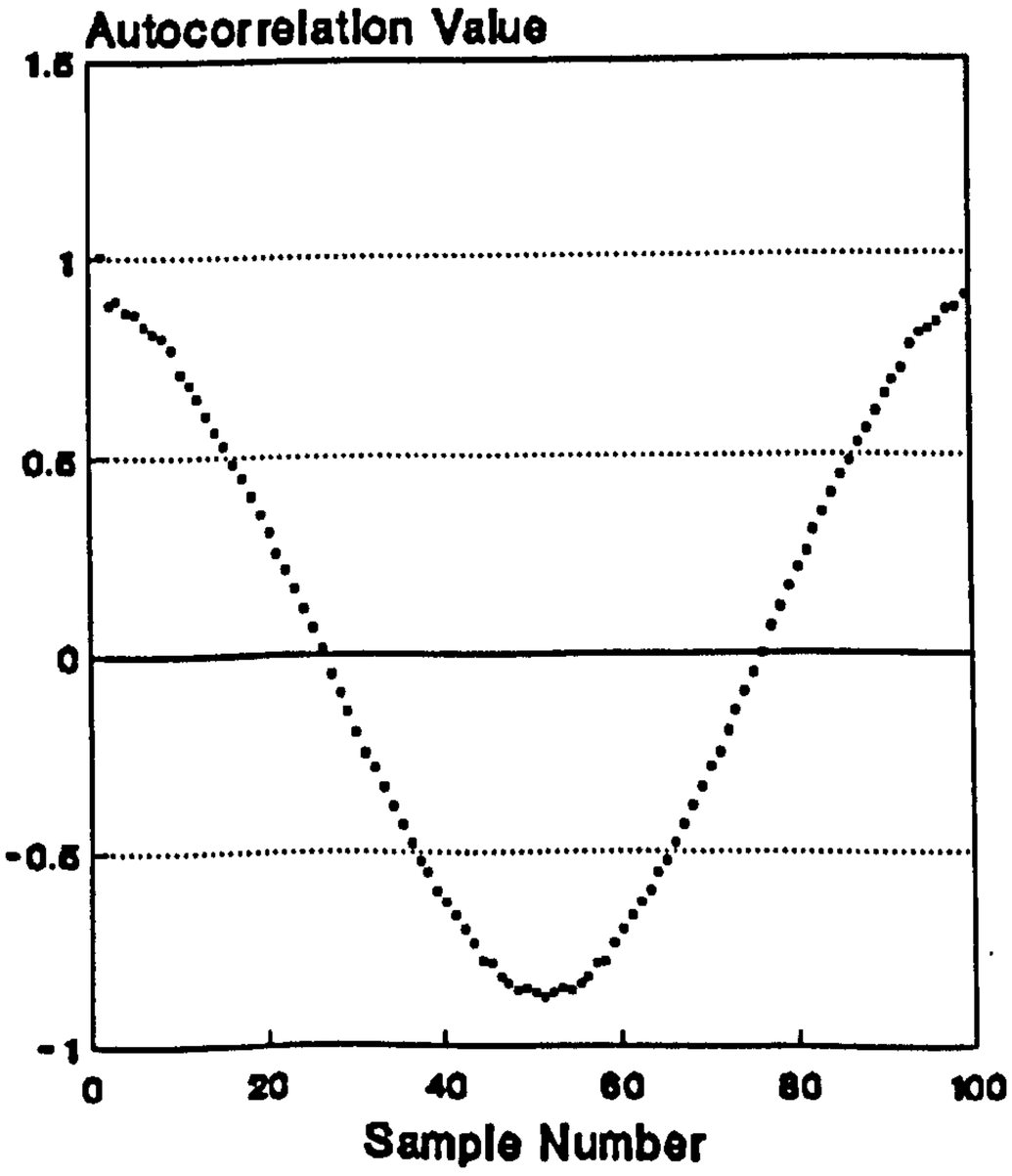


Fig.5.05(b) Ellipse

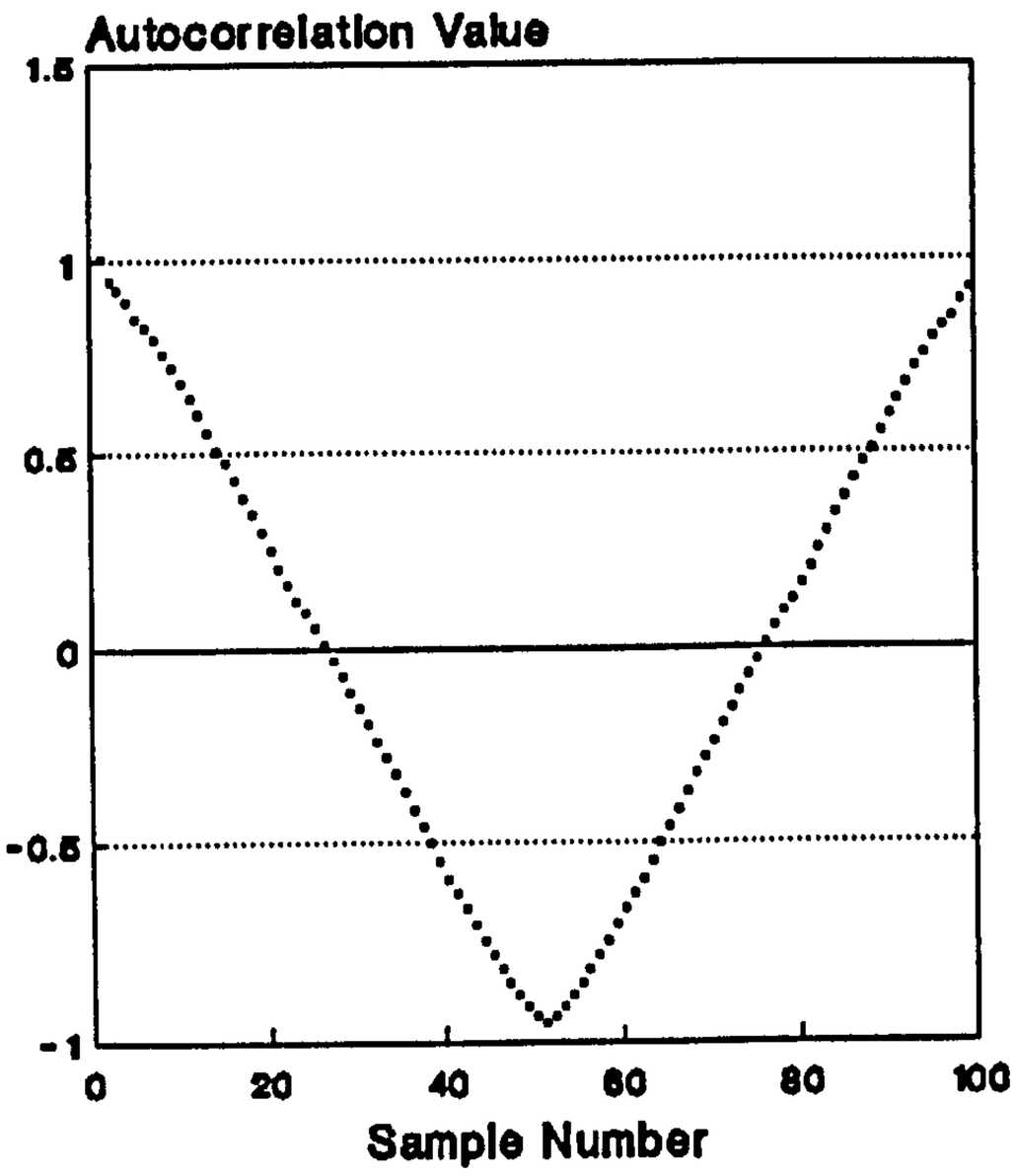


Fig.5.05(c) Rectangle

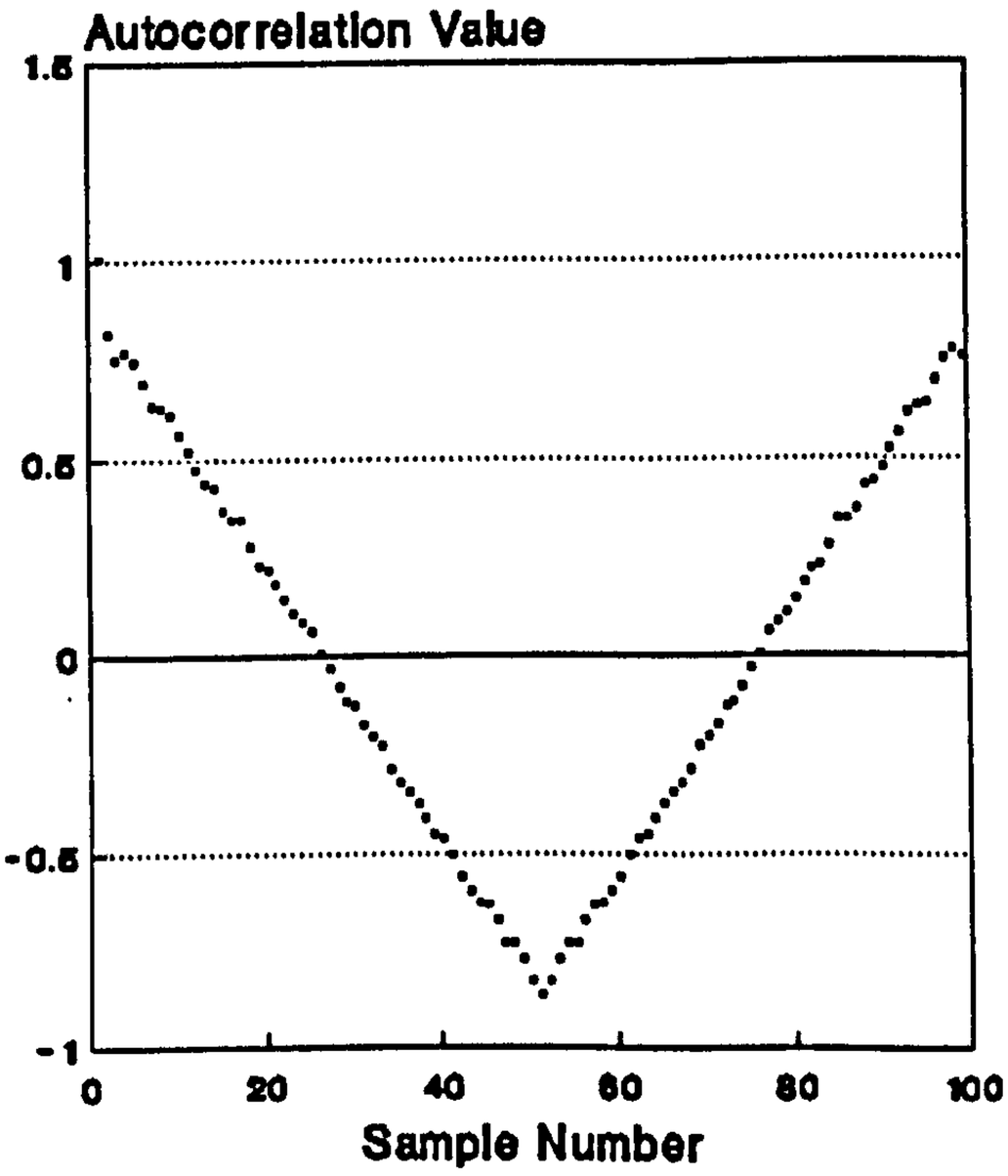


Fig.5.05(d) Rhombic acute (Diamond4)

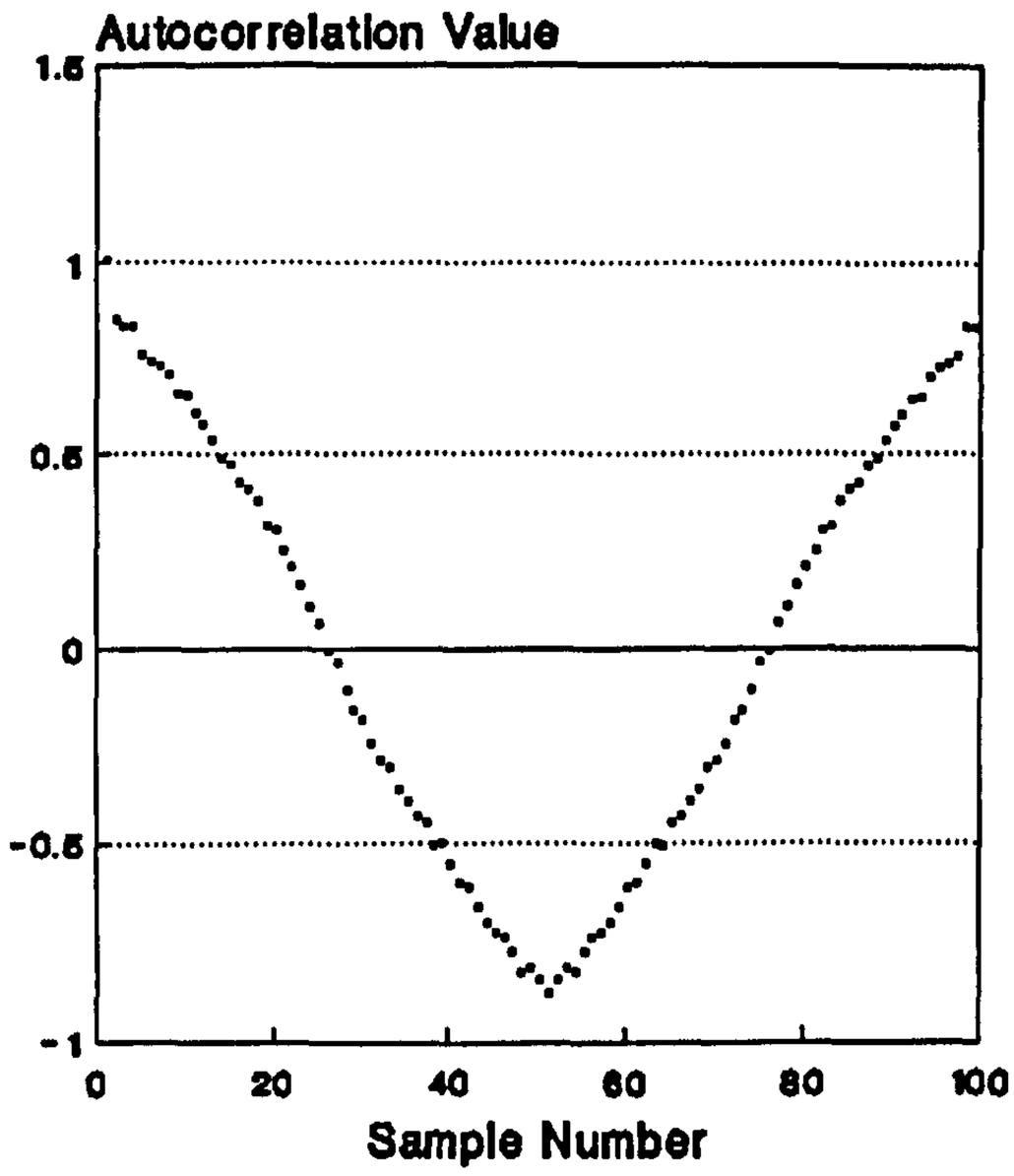


Fig.5.05(e) Rhombic acute truncate

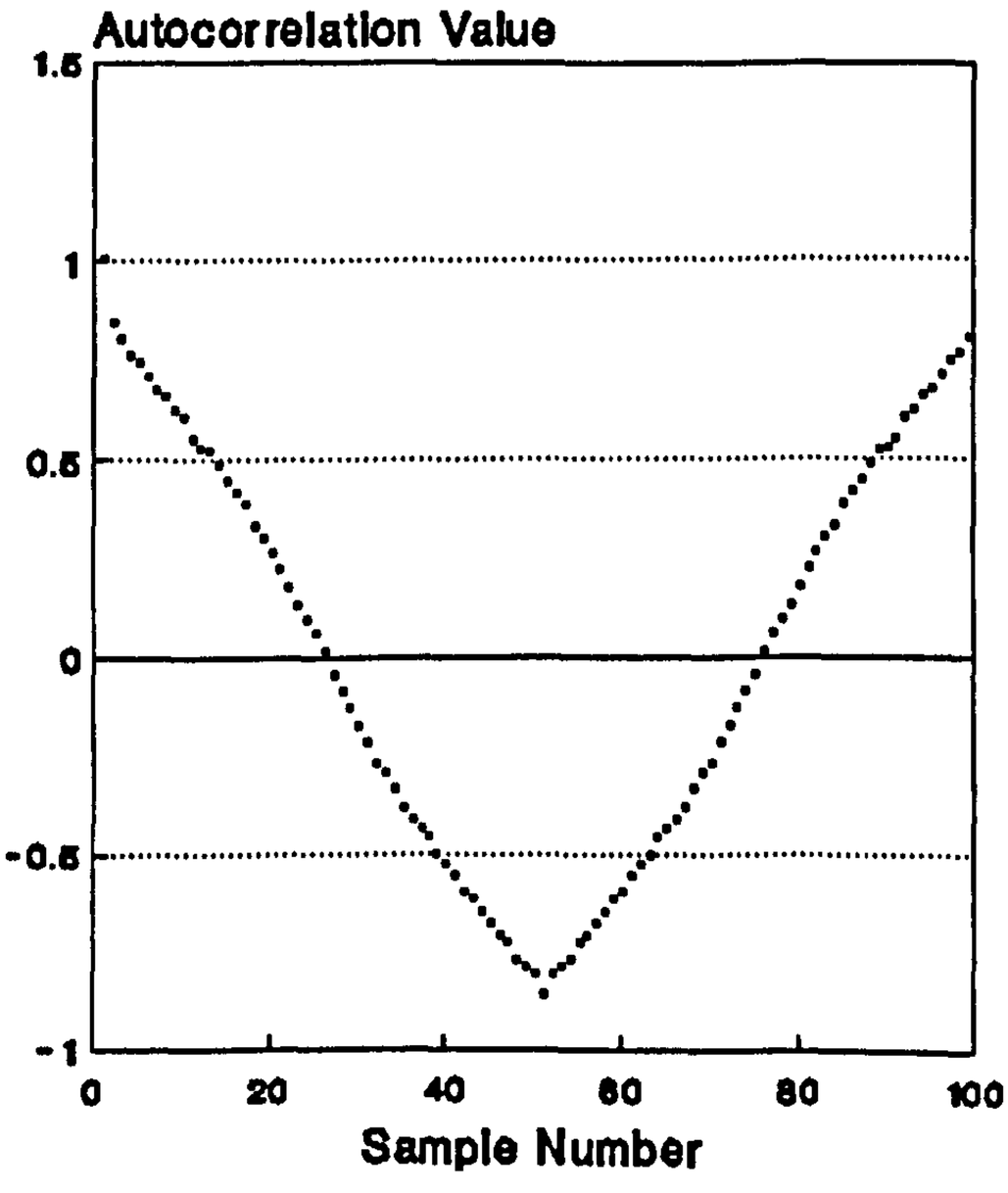


Fig.5.05(f) Elliptic acuminate acute

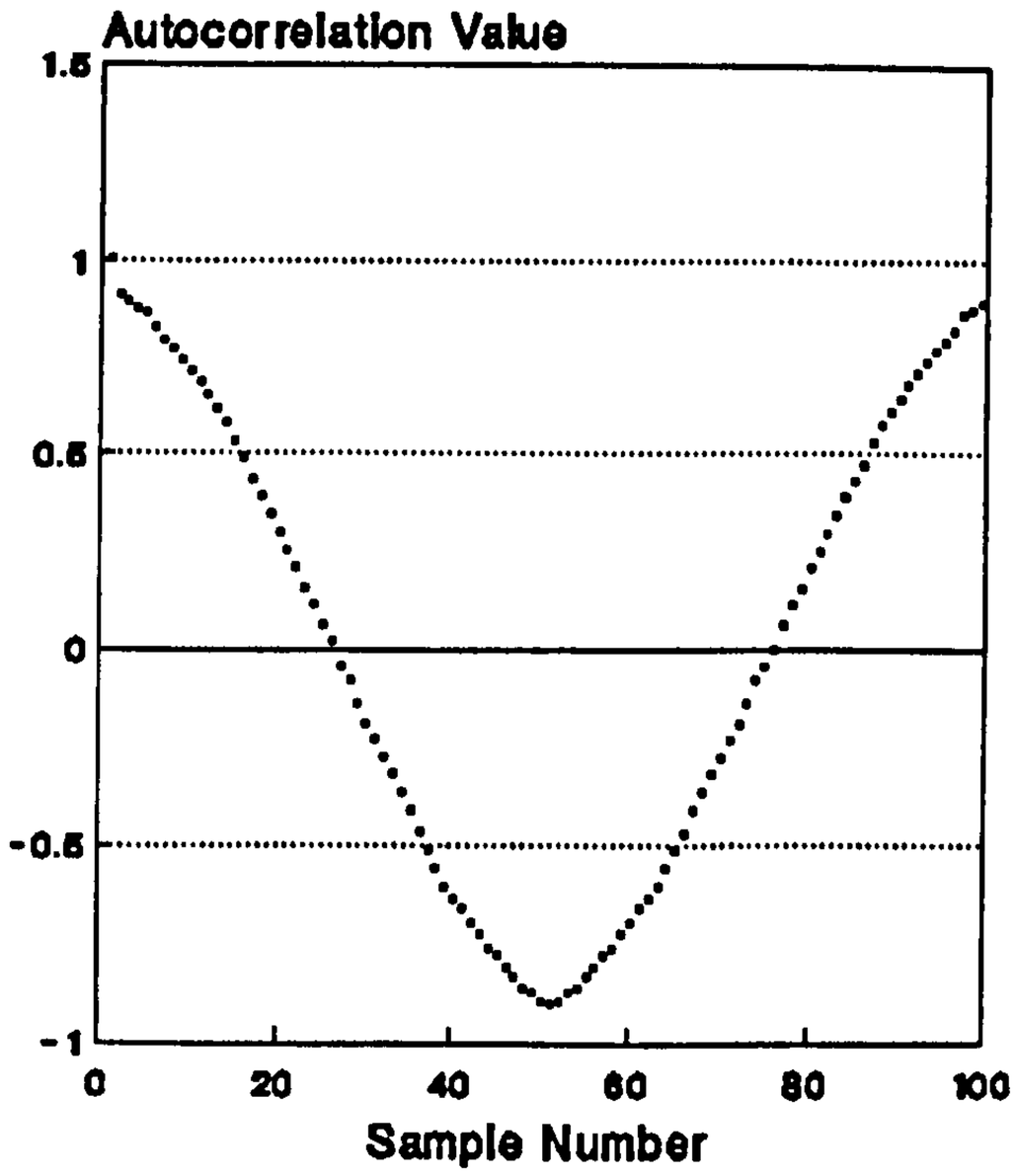


Fig.5.05(g) Elliptic truncate (Pillbox)

5.9 Specific Spine Detection and Classification

This section of the study describes a new method (Treloar and Taylor[43]) to distinguish between a group of spinose pollen. These have long, conspicuous, and generally sharp, excrescences (spines) whose length generally exceeds $3\mu m$. The usefulness of texture to identify pollen within this group is hindered as the spines make it difficult to select an area with a uniform microstructure for analysis, Waterhouse[45]. The shape parameters presented here are unconstrained by the spines and, in fact, use them to identify the pollen.

Fresh pollen from four spinose taxa were used:

Bidens hendersonensis (Plates 1),

Emilia sonchifolia (Plates 4),

Fitchia speciosa (Plates 5),

Senecio stokesii (Plates 10).

A human would use the size, distribution and shape of spines over the pollen surface to identify these pollen. Fitchia speciosa is recognised most easily as it has steeply pointed spines that have a relatively sparse distribution. The surfaces of the other pollen are dominated by the spines and are harder to separate. Emilia sonchifolia has small spines that are densely packed over the pollen surface. Bidens hendersonensis

may be distinguished from Senecio stokesii in that it has a slightly higher spine density with spines which are generally shorter with wider bases.

5.9.1 Difference Code

A close look at Fig.5.01 shows the way in which the chain code(A) changes as it traces its way around a spine. As it moves towards the peak of the spine the chain code is either 1 or 2. Once past the peak the code is either 6 or 7 with a sharp change at the peak from 2 to 6. A difference code can be used to highlight this change in code and therefore may be useful in the classification of spinose pollen.

5.9.2 Generating the Difference Code

Difference code can quickly and easily be constructed. The difference code (D_1) in Fig.5.01 is formed by recording the minimum number of 45° turns between each chain code and the next. For example, to go from code 6 to 7 a single 45° turn is required and so the difference code is 1. From codes 7 to 2, three 45° turns are required and so a difference code of 3 is recorded. This can be represented thus:

$$D_x = \min(|a_i - a_{i+x}|, N - |a_i - a_{i+x}|)$$

$$0 \leq D_x \leq \frac{N}{2}$$

N = number of pixel neighbours ($N = 8$).

x = distance between links to be compared.

The chain code used is assumed to be a closed loop where the initium and terminus coincide and therefore:

$$a_1 = a_{n+1} \text{ and } a_n = a_{1-1}$$

As the above equations suggest, each code need not be compared to its immediate successor, but may be compared to codes several steps along. For example, Fig.5.01 also shows the difference code for a step displacement of 4. This gives yet more information about the chain code being analysed. It is the potential of this property that will now be explored.

5.9.3 Difference Code Probabilities

A set of difference codes is calculated for displacement steps 1 to 20. Then the probability of each difference code occurring at each step is calculated and used as feature vectors for variable selection and leave-one-out classification (see chapter 2). Due to the large number of features (and small

number of samples) numerous sub-samples of difference code probabilities were used and the best combined into a final classification system.

5.10 Analysis of Shape and Geometric Measures

5.10.1 Results

Two sets of 18 images were captured for each of the 12 pollen taxa. The first set of images was taken at a constant magnification of 1250x. The information extracted from these images was concerned with relative pollen sizes. A second set of images was taken at the maximum magnification so that the whole pollen grain could still be seen. These images were used for shape analysis.

The first set of results was derived from the basic shape measures outlined in section 5.5. The relative pollen areas and perimeters observed at 1250x magnification are shown in Fig.5.06 and Fig.5.07. A gross perimeter measure, namely the number of links in the chain code, is shown in Fig.5.08. The measures shown in Fig.5.09 to Fig.5.13 were calculated at the maximum possible magnifications and are ratios of the basic shape measures; area, perimeter and the maximum and minimum diameters.

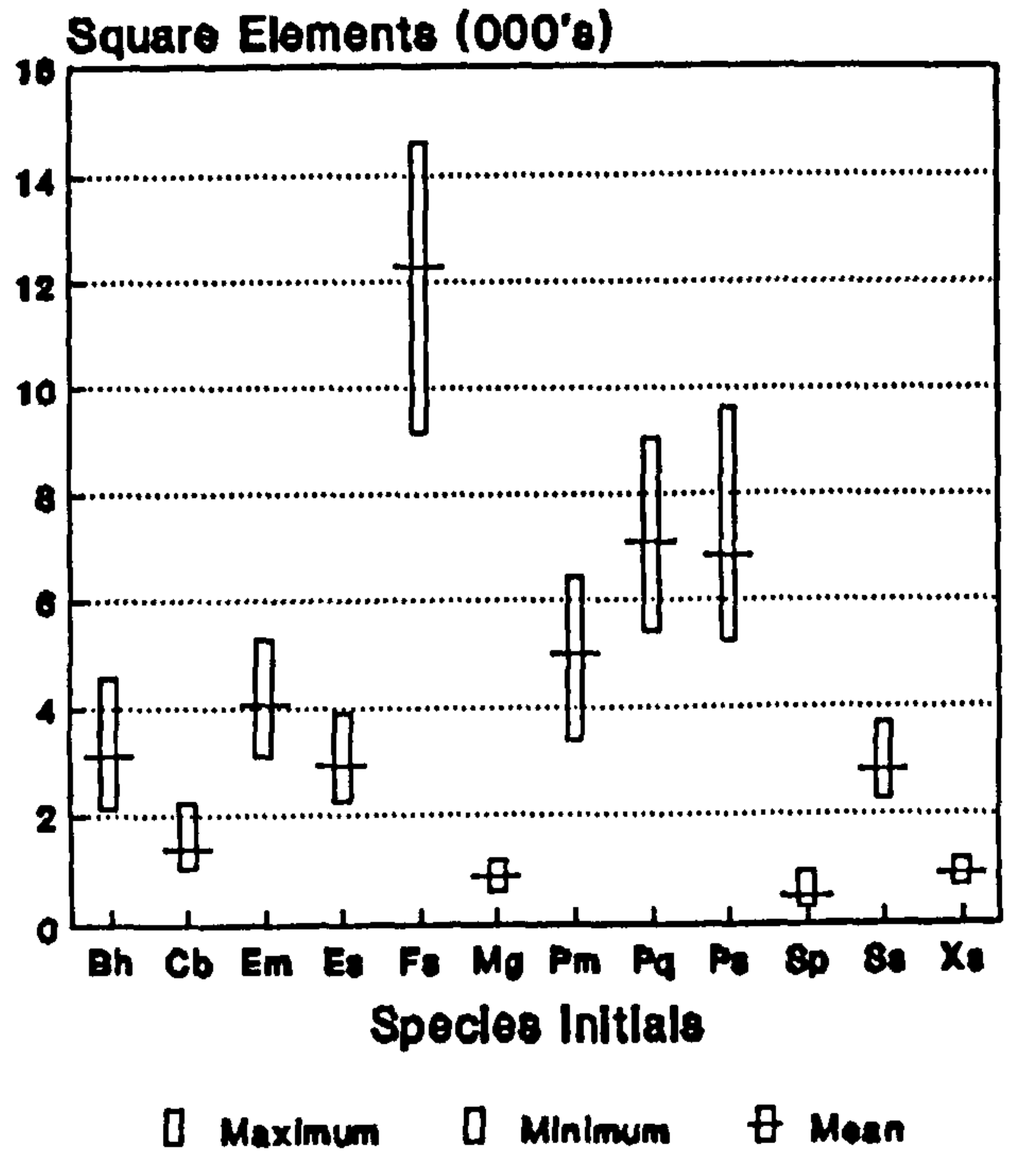


Fig 5.06 : Pollen Image Area

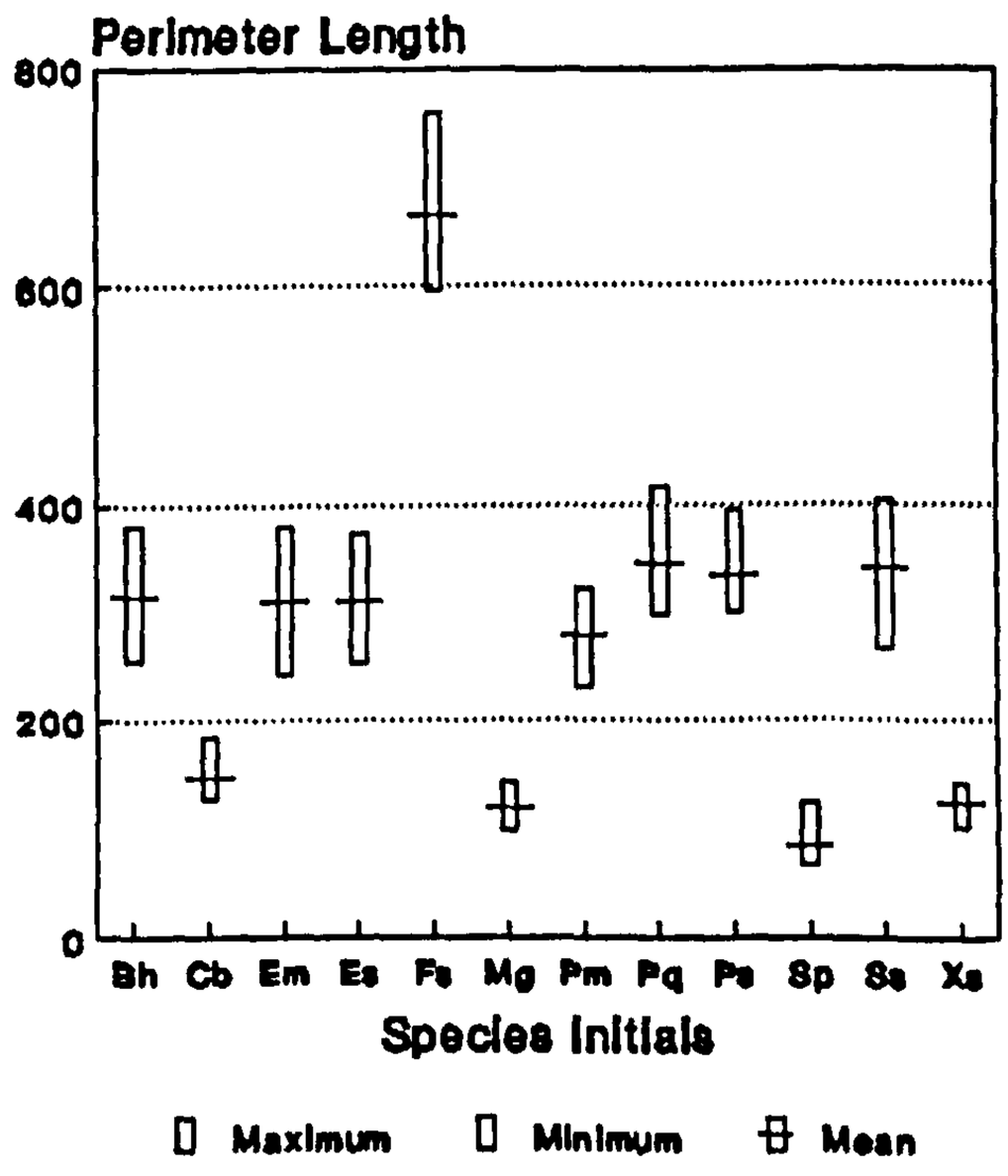


Fig 5.07 : Pollen Perimeters.

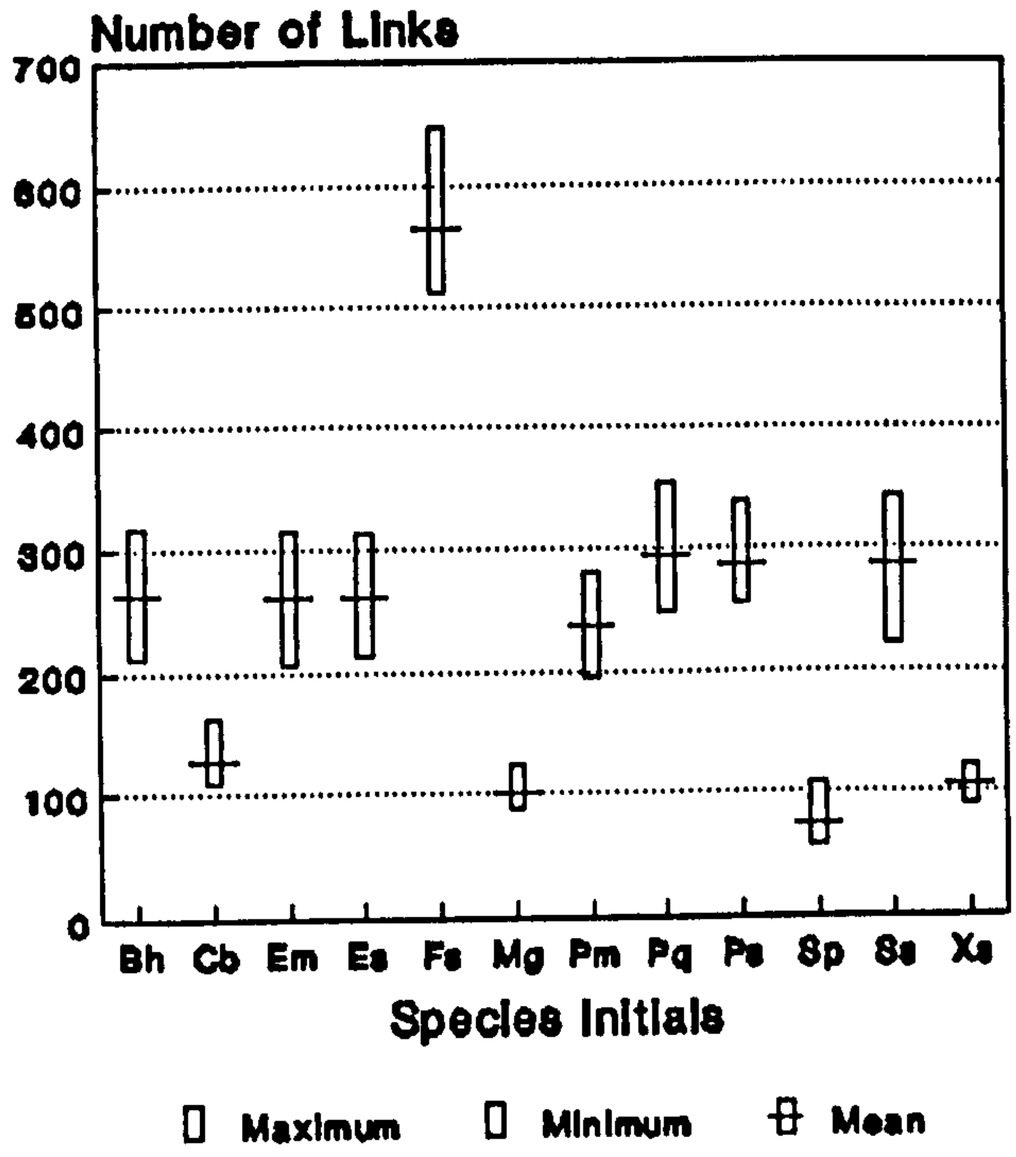


Fig 5.08 : Chain Code Links

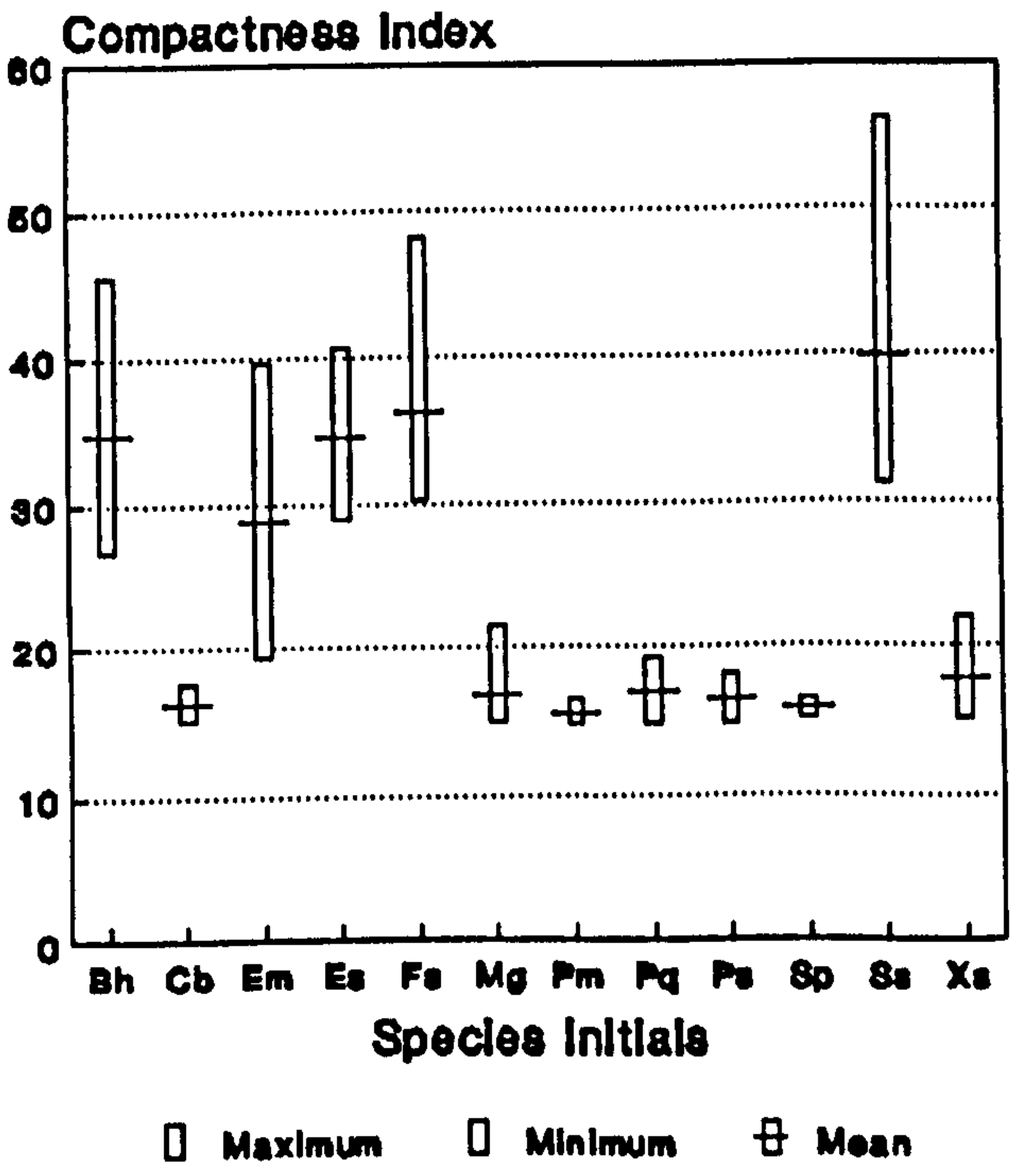


Fig 5.09 : Pollen Compactness Index

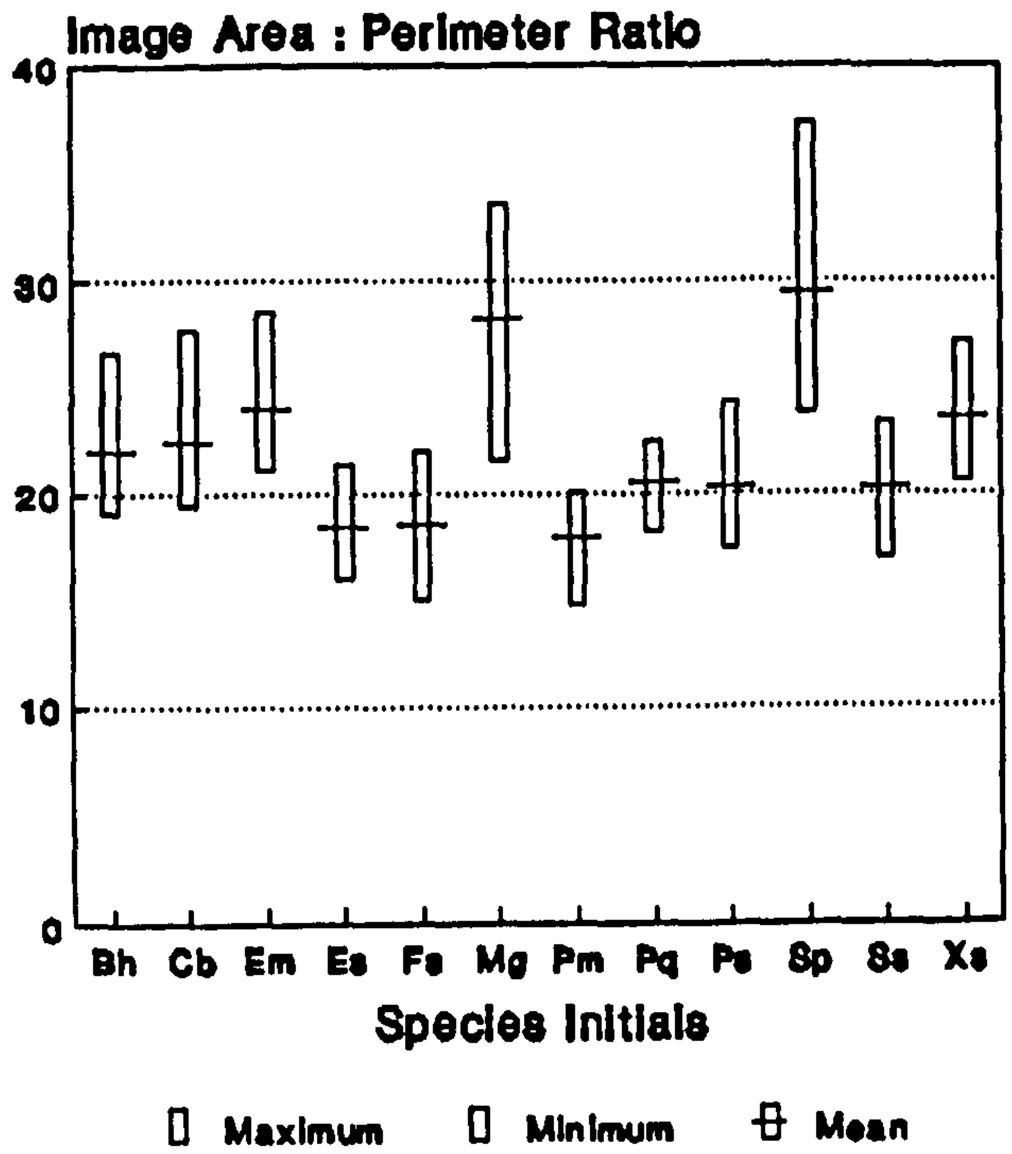


Fig 5.10 : Area : Perimeter

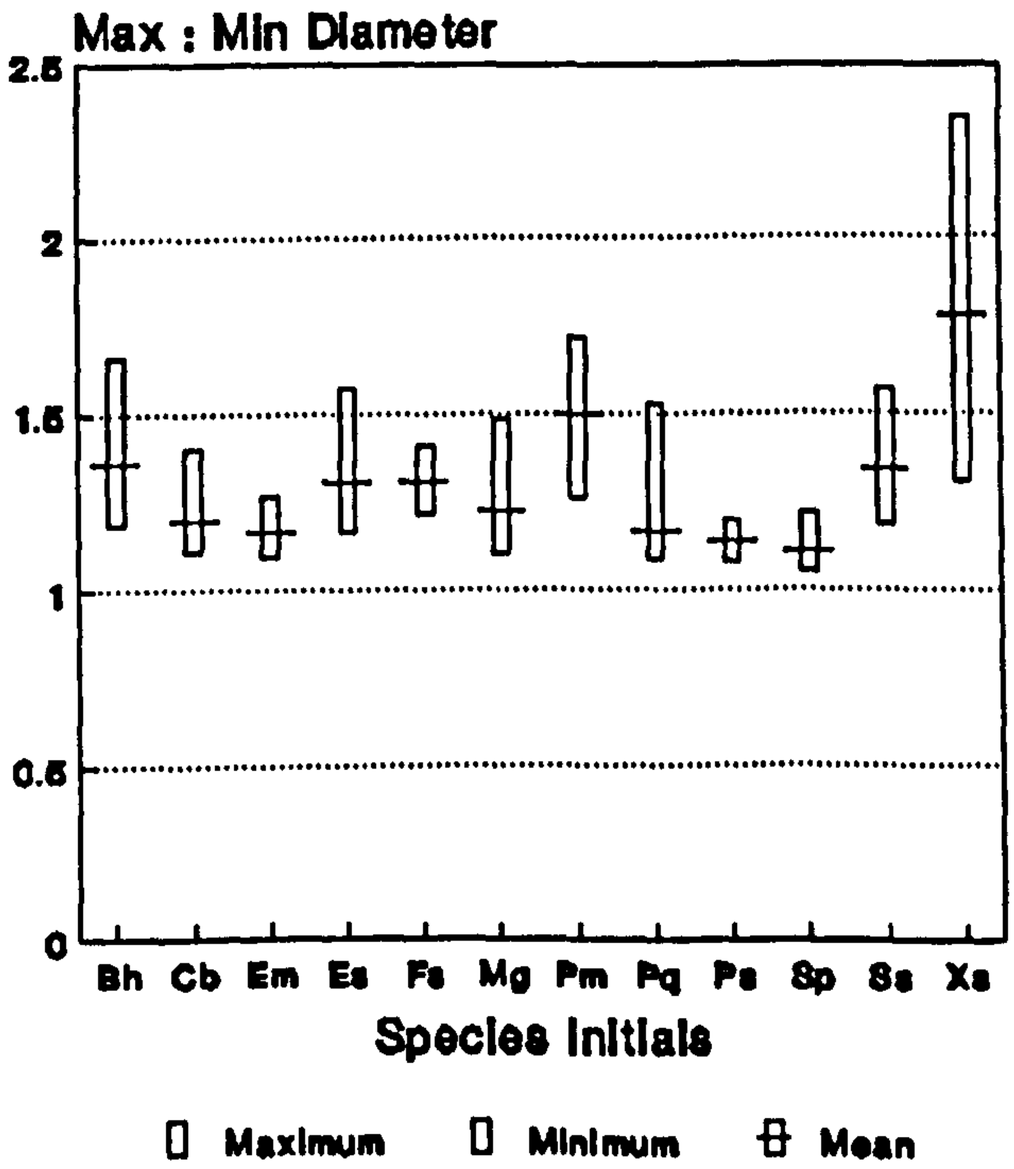


Fig 5.11 : Pollen Max : Min Diameter

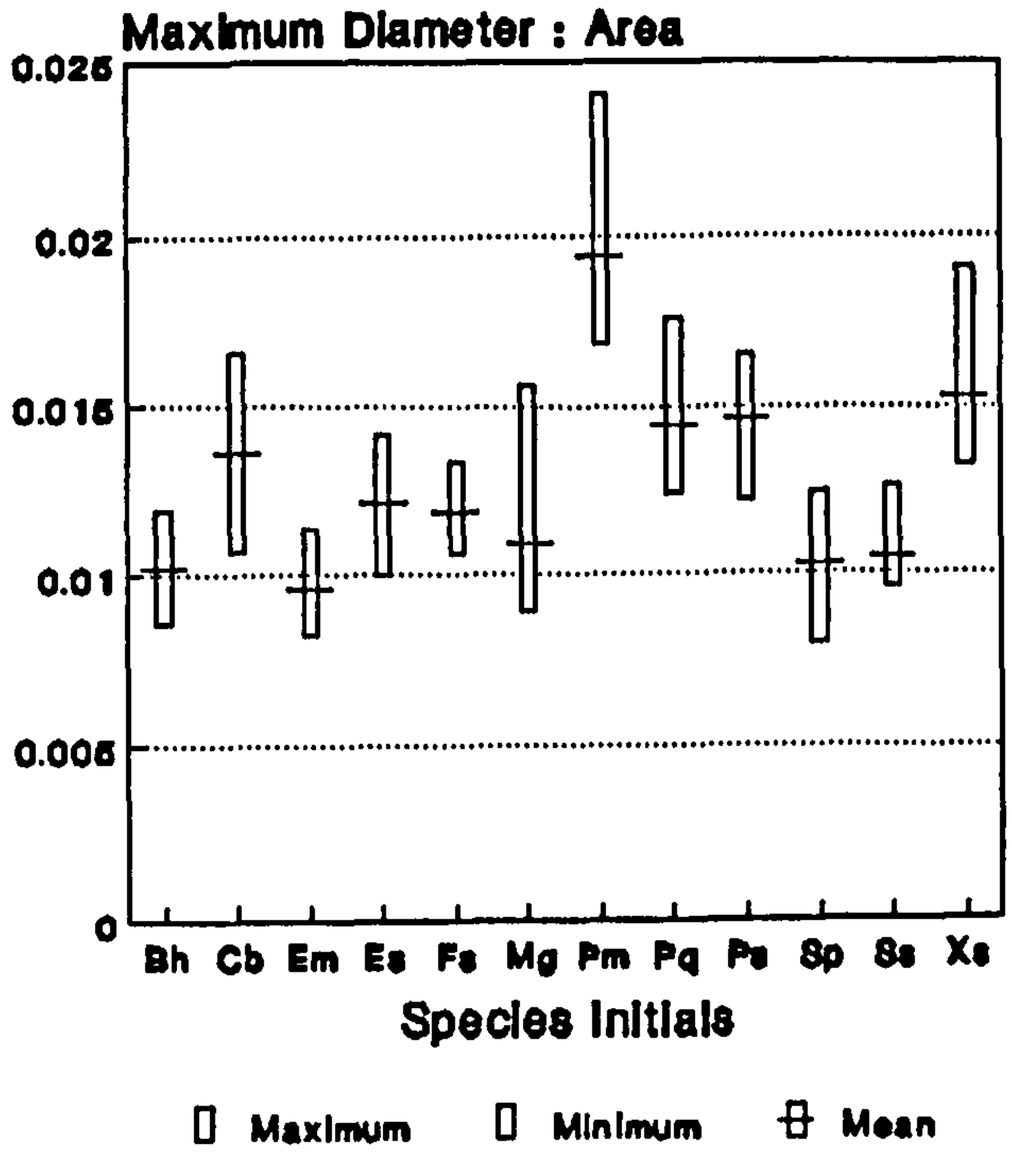


Fig 5.12 : Pollen Max. Diameter : Area

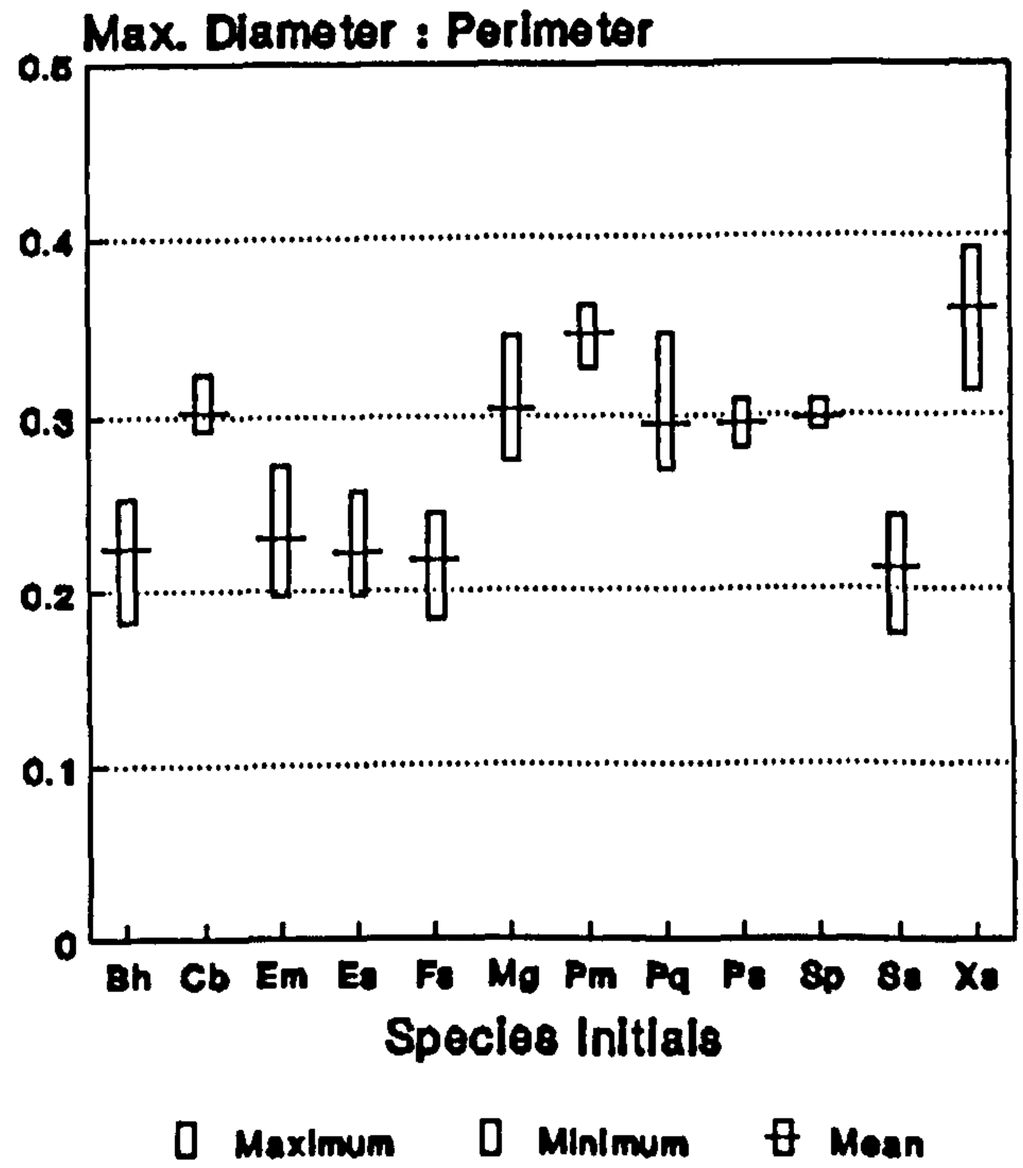


Fig 5.13 : Max. Diameter : Perimeter

As well as considering the separation of the pollen using single shape variables as described above, a whole set of variables was classified using the leave-one-out classifier described in section 2.1.3.1. The nine variables selected were; compactness index, area, perimeter, number of chain code links, area : perimeter, perimeter : area, maximum : minimum diameter, maximum diameter : area and maximum diameter : perimeter. Table 5.1(a) shows the classification rates for all 12 taxa using the variables selected. The classification rates for the non-spinose pollen are shown in Table 5.1(b) and those for the spinose pollen can be seen in Table 5.1(c).

The variable selection procedure tended to select single variables to discriminate between pollen class pairs. The variables were of varying usefulness for classification purposes and Table 5.2 shows the frequency with which each variable was selected. Also there were 5 groups of variables that tended to be associated with each other. Table 5.3 shows the relative importance of each group given by the relative frequency with which each group was selected. The groups made up of compactness index, area and either perimeter : area or area : perimeter were the most useful exclusively for separating spinose pollen.

Moment invariants were calculated for all 12 taxa and then for both the spinose and non-spinose taxa. Three different combinations of the moment invariant calculations were used. Firstly, using the logarithm of each of the 7 moments (ψ_i) applied to the object boundaries and both the raw and binary

Table 5.1(a) : Basic shape measures applied to all 12 taxa

	% Matched to Taxa								
Taxa/	Bh/Pq Ps/Ss	Cb	Em	Es	Fs	Mg	Pm	Sp	Xs
Bh/Pq/Ps/Ss	94	-	2	2	-	-	2	-	-
Cb	-	88	6	-	-	-	-	-	6
Em	-	-	94	3	-	-	-	3	-
Es	2	-	-	97	-	-	-	2	-
Fs	-	-	-	-	100	-	-	-	-
Mg	-	-	-	-	-	88	-	-	13
Pm	-	-	-	-	-	-	100	-	-
Sp	-	-	-	-	-	6	-	94	-
Xs	-	-	-	-	-	-	-	-	100

Table 5.1(b) : Basic shape measures for non-spinose pollen

	% Matched to Taxa						
Taxa/	Cb	Em	Mg	Pm	Pq/Ps	Sp	Xs
Cb	88	6	-	-	-	-	6
Em	-	86	-	2	11	-	-
Mg	-	-	88	-	-	-	13
Pm	-	-	-	100	-	-	-
Pq/Ps	-	-	-	3	97	-	-
Sp	-	-	6	-	-	94	-
Xs	-	-	-	-	-	-	100

Table 5.1(c) : Basic shape measures for spinose pollen

	% Matched to Taxa		
Taxa/	Bh/Ss	Es	Fs
Bh/Ss	94	6	-
Es	3	97	-
Fs	-	-	100

Table 5.2 : Relative occurrences of basic shape variables

Area	=	30.11%
Perimeter	=	29.55%
Number of Links	=	12.50%
Max. Diameter : Area	=	9.66%
Max. Diameter : Perimeter	=	8.52%
Perimeter : Area	=	4.55%
Area : Perimeter	=	2.84%
Compactness Index	=	2.27%
Max. : Min. Diameter	=	0.00%

Table 5.3 : Relative occurrences of basic shape variable groups

Area & Number of Links	= 41.18%
Perimeter & Area : Perimeter	= 17.65%
Links & Perimeter : Area	= 17.65%
Compactness, Area & Area : Perimeter	= 11.76%
Compactness, Area & Perimeter : Area	= 11.76%

images at the maximum magnification. The best classification results were obtained using a combination of moments calculated from raw images and object boundaries. These results may be seen in Table 5.4. The second moment combination utilised the invariant moments calculated from the normalised central moments (η_{pq}) to give scale invariant moments. These were applied to both the raw images and pollen boundaries and produced the classifications seen in Table 5.5. The final combination simply added the contrast invariance features (β_i) to the calculations. These were again applied to both raw images and pollen boundaries. The resulting features could only separate about half of the class pairs and therefore no final classification was attempted.

The standard shapes shown in Fig.5.02 and their boundary trace and autocorrelation plots (Fig.5.03 to Fig.5.05) were used as templates to classify the pollen taxa into general shape groups. The results for both the boundary tracing and auto-correlation methods are summarised in Table 5.6. As these methods are inappropriate for spinose pollen only the non-spinose pollen are used.

Finally the spine detection and classification method was tested. Due to the large number of difference code probability features produced, the variable selection procedure was applied to several sub-groups. The best variables from each sub-group were combined and reapplied to the variable selection procedure. The efficiency of the variables selected on this second pass

Table 5.4(a) : Basic moment invariants for all 12 taxa

	% Matched to Taxa										
Taxa/	Bh	Cb	Em	Es/Fs	Mg	Pm	Pq	Ps	Sp	Ss	Xs
Bh	77	-	15	-	-	-	-	-	-	8	-
Cb	-	69	-	6	-	-	13	-	13	-	-
Em	6	-	94	-	-	-	-	-	-	-	-
Es/Fs	14	-	-	73	-	-	-	-	-	14	-
Mg	-	-	13	-	78	-	-	-	9	-	-
Pm	-	-	-	-	-	75	6	6	-	-	13
Pq	-	6	-	-	6	-	81	-	6	-	-
Ps	-	16	-	-	3	-	-	81	-	-	-
Sp	-	-	19	-	-	-	-	-	81	-	-
Ss	13	-	-	3	-	-	-	-	-	84	-
Xs	-	-	-	-	13	-	-	-	-	-	88

Table 5.4(b) : Basic moment invariants for non-spinose pollen

	% Matched to Taxa							
Taxa/	Cb	Em	Mg	Pm	Pq	Ps	Sp	Xs
Cb	69	6	-	-	13	-	13	-
Em	-	88	13	-	-	-	-	-
Mg	-	13	69	-	-	-	13	6
Pm	-	-	-	75	6	6	-	13
Pq	6	-	6	-	81	-	6	-
Ps	15	-	2	-	-	81	2	-
Sp	-	21	2	-	-	-	77	-
Xs	-	-	13	-	-	-	6	81

Table 5.4(c) : Basic moment invariants for spinose pollen

	% Matched to Taxa		
Taxa/	Bh	Es/Fs	Ss
Bh/Ss	94	-	6
Es	9	72	19
Fs	13	3	84

Table 5.5(a) : Moment invariants applied to all 12 taxa

	% Matched to Taxa						
Taxa/	Cb/Em Bh	Es/Fs Pq/Ps	Ss	Mg	Pm	Sp	Xs
Bh	78	4	1	3	6	7	1
Cb/Em/Pq/Ps	2	88	-	10	1	-	1
Es/Fs/Ss	8	3	86	-	3	-	-
Mg	2	1	-	96	-	-	2
Pm	8	5	2	-	83	-	2
Sp	-	2	-	-	-	98	-
Xs	6	2	-	-	-	-	92

Table 5.5(b) : Moment invariants for non-spinose pollen

	% Matched to Taxa				
Taxa/	Cb/Em Pq/Ps	Mg	Pm	Sp	Xs
Cb/Em/Pq/Ps	87	10	-	-	3
Mg	2	93	-	-	5
Pm	5	-	94	-	1
Sp	1	1	-	99	-
Xs	2	7	-	-	91

There may be clue to contrast effects be of may somewhat

Table 5.5(c) : Moment invariants for spinose pollen

	% Matched to Taxa	
Taxa/	Bh	Es/Fs/Ss
Bh	98	2
Es/Fs/Ss	13	88

Table 5.6(a) : Equal distance boundary trace classification

	% Matched to Shape Class						
Taxa/	1	2	3	4	5	6	7
Cb	63	38	-	-	-	-	-
Em	75	25	-	-	-	-	-
Mg	56	44	-	-	-	-	-
Pm	-	75	-	-	13	-	13
Pq	88	13	-	-	-	-	-
Ps	94	6	-	-	-	-	-
Sp	100	-	-	-	-	-	-
Xs	-	13	-	-	81	6	-

Table 5.6(b) : Equal angle boundary trace classification

	% Matched to Shape Class						
Taxa/	1	2	3	4	5	6	7
Cb	-	-	-	6	75	19	-
Em	-	-	-	25	25	50	-
Mg	-	-	-	-	69	31	-
Pm	-	-	-	56	6	38	-
Pq	-	-	-	19	44	38	-
Ps	-	6	-	13	56	25	-
Sp	-	-	-	6	88	6	-
Xs	-	-	-	38	6	56	-

Table 5.6(c) : Autocorrelation classification

	% Matched to Shape Class						
Taxa/	1	2	3	4	5	6	7
Cb	-	6	-	-	63	31	-
Em	-	-	-	100	-	-	-
Mg	-	25	-	13	25	31	6
Pm	-	69	-	-	6	-	19
Pq	-	13	-	13	44	31	-
Ps	-	13	-	-	50	38	-
Sp	-	31	-	-	69	-	-
Xs	-	13	-	6	38	31	13

Key : Shape Class 1 = Circle
2 = Ellipse
3 = Rectangle
4 = Diamond 4
5 = Diamond 5
6 = Lense
7 = Pillbox

ranged from 81% to 100%. Average results from three of the best variable sets are shown in Table 5.7. Each class pair required approximately three variables to separate them.

As well as efficiently classifying the spinose pollen this method may also be used to separate spinose from non-spinose pollen. The probability of difference codes 3 and 4 occurring in spinose pollen is always greater than zero for all displacement steps. Non-spinose pollen have occasional incidents of difference code probabilities greater than zero for codes 3 and 4 but these never span more than 3 or 4 displacement steps. When they do occur the difference code probabilities never exceed 0.027 and 0.002 for codes 3 and 4 respectively. The equivalent probability values for spinose pollen are always greater than 0.12 and 0.02 respectively for displacement steps from 7 and above.

5.10.2 Discussion and Concluding Remarks

The ultimate purpose of the shape measures discussed in this chapter is to reduce the number of pollen taxa in each sub-sample at the final level of the classification scheme (Fig.1.01). Before any assessment can be made in respect to this, it is important to examine the success of the texture classification for all 8 non-spinose pollen together. It has already been shown that spinose pollen can easily be separated from the rest of the pollen; they are therefore left out at this point. The

Table 5.7 : Spinose pollen classification using difference code probabilities

	% Matched to Taxa			
Taxa/	Bh	Es	Fs	Ss
Bh	88	2	4	6
Es	6	94	-	-
Fs	3	1	93	2
Ss	6	-	-	94

Table 5.8 : Non-spinose pollen texture classification

	% Matched to Taxa							
Taxa/	Cb	Em	Mg	Pm	Pq	Ps	Sp	Xs
Cb	67	-	11	5	-	-	17	-
Em	-	96	2	-	-	-	-	2
Mg	-	-	61	-	-	-	-	39
Pm	-	-	-	91	-	-	2	7
Pq	-	6	-	-	94	-	-	-
Ps	6	-	-	-	-	94	-	-
Sp	13	-	-	7	-	-	74	6
Xs	2	6	18	6	-	-	-	68

classification of spinose pollen is discussed later. Table 5.8 shows the texture classification of non-spinose pollen. The texture features employed are Haralick measures obtained automatically from images using methods that will be described in more detail in the next chapter. While the success of the classification for some pollen is very encouraging with rates of over 90%, others do less well at less than 75%. Ideally the shape measures need to separate pollen taxa that commonly become confused, for example, Macaranga graeffeana (Mg) and Xylosma suaveolens (Xs).

The equal distance boundary tracing and matching method only met with limited success. Six non-spinose taxa were basically circular or elliptical and these are indeed the standard shape classes to which they were assigned. Both Pritchardia minor (Pm) and X.suaveolens (Xs) tend towards diamond, lens and pillbox shapes and indeed these are the only taxa that are classified this way. While this technique classifies the basic taxa shapes correctly it does not allow them to be split into sufficiently small groups for subsequent texture analysis.

The equal angle boundary tracing and matching method was even less successful than the equal distance method. The taxa are almost exclusively classified to either diamond or lens shape classes. The main problem with this method is brought about by the troughs in the boundary plot that cover a relatively short number of vectors due to the sampling frequency on edges furthest from the centroid. During the plot matching stage

even slight misalignments of the plots cause large squares of variance to occur. On reflection it may have been more advantageous to have matched these plots at their minima and not their maxima. Even with this method changed it is unlikely that any significant improvement over the equal distance method would occur.

Both these boundary tracing methods suffer when any part of the object boundary is missing. Firstly this displaces the position of the object centroid and secondly it disrupts the form of the resultant plot. This method, while more efficient than two dimensional template matching (used by Mirkin and Bagdasaryan[31]), suffers due to the natural variability of pollen shape.

The results using the autocorrelation technique are very poor for the same reasons as for the boundary tracing methods. This technique is much less flexible however, as all the plots are very similar and rely on subtle differences to distinguish between the shape classes.

The initial results obtained from the basic moment invariant calculations were very promising with relatively high classification rates and little confusion between the pollen taxa. However, subsequent results with the addition of both scale and contrast invariances show that the initial separations in Table 5.4 are probably due to contrast and scale differences in the images. It is too early however to discount the potential of

moment invariants from this problem. The classification scheme collapsed when the contrast invariant functions (β_i) were added. This is not too surprising as the contrast and brightness levels of the image were very high at the time (to enhance the pollen edge for object location). The brightness of the central regions of the pollen almost totally removed any trace of internal texture. The moments were effectively picking up the differences in illumination from taxa to taxa. To combat this problem a second image may be captured with 'normal' illumination - as used for texture analysis. After histogram equalisation and therefore dampening the contrast between images the moment invariants with scale invariance and without contrast invariance may be applied. The resulting moments will then contain information about both the shape and internal structure of the pollen.

The spine detection and classification method proved to be highly effective, producing high classification rates that are comparable to those achieved using texture analysis and requiring a small number of variables per class pair to do so. They can also be easily separated from non-spinose pollen by examining a portion of the difference codes produced.

The compactness index, one of the basic shape measures, also separates the spinose pollen from the non-spinose (Fig.5.09). Interestingly this measure also classes Elephantopis mollis (Em) at a point bridging the gap between spinose and non-spinose pollen. As Plate 3(a) shows E.mollis does have some small

spines on its surface - however these are not large enough for it to be classed as a 'true' spinose taxa. The classification of spinose and non-spinose pollen is still most effectively done using difference codes. Although as Table 5.1(c) shows the classification of Emilia sonchifolia (Es) and Fitchia speciosa (Fs) is improved when basic shape measures are combined. Indeed, F.speciosa may be left out of the difference code classification as the basic shape measures produce a 100% classification.

Of all the shape measures analysed in this chapter the basic shape measurements are by far the most effective. The three most successful single variables are; pollen area, perimeter and number of links in the chain code. All of these produced sub-samples containing no more than four pollen taxa. From the area graph in Fig.5.06, five sub-samples of pollen may be constructed with the following cut-off points:

8000+	- sub-samples of;	Pq & Ps
6-8000	- sub-samples of;	Pm, Pq & Ps
4-6000	- sub-samples of;	Em, Pm, Pq & Ps
2-4000	- sub-samples of;	Cb, Em & Pm
0-2000	- sub-samples of;	Cb, Mg, Sp & Xs

The resulting texture classifications of these sub-samples are given in Table 5.9(a) to (e).

Table 5.9(a) : Texture classification of Pq & Ps

	% Matched to Taxa	
Taxa/	Pq	Ps
Pq	100	-
Ps	-	100

Table 5.9(b) : Texture classification of Pm, Pq & Ps

	% Matched to Taxa		
Taxa/	Pm	Pq	Ps
Pm	100	-	-
Pq	6	94	-
Ps	-	-	100

Table 5.9(c) : Texture classification of Em, Pm, Pq & Ps

	% Matched to Taxa			
Taxa/	Em	Pm	Pq	Ps
Em	100	-	-	-
Pm	-	100	-	-
Pq	7	2	91	-
Ps	-	-	-	100

Table 5.9(d) : Texture classification of Cb, Em & Pm

	% Matched to Taxa		
Taxa/	Cb	Em	Pm
Cb	83	-	17
Em	-	100	-
Pm	-	-	100

Table 5.9(e) : Texture classification of Cb, Mg, Sp & Xs

	% Matched to Taxa			
Taxa/	Cb	Mg	Sp	Xs
Cb	72	11	17	-
Mg	-	61	-	39
Sp	24	-	74	2
Xs	4	26	-	70

It must be noted that the variables used to separate these sub-samples have been selected individually for each sub-sample. The classification of non-spinose pollen in Table 5.8 used 17 different Haralick measures. The number of measures used for these sub-samples are considerably smaller. Also the measures produced by Laws[26] play a greater role as the number of classes to separate declines. When only two taxa are to be separated Laws measures are used on about 70% of occasions.

Both the perimeter and number of chain code links produce identical sub-samples at slightly different cut-off points. The most distinct separations may be seen in the perimeter graph in Fig.5.07:

550+ - sub-samples of; Fs (spinose)

200-550 - sub-samples of; Em, Pm, Pq & Ps

0-180 - sub-samples of; Cb, Mg, Sp & Xs

The subsequent texture analysis of the lower groups may be seen in Table 5.9(c) and (e). The good classifications achieved using these sub-samples are marred by the classification shown in Table 5.9(e) where M.graeffeana (Mg), X.suaveolens (Xs), Canthium barbatum (Cb) and the fungal spores (Sp) are all poorly classified.

In an attempt to improve this situation several of the basic shape measures were combined and produced the classification shown in Table 5.1(b). Using this table it can be seen that

if an 'unknown' sample is matched to X.suaveolens (Xs) then there is a small chance that it could also be either C.barbatum (Cb) or M.graeffeana (Mg). Using this principle the following 7 sub-samples were constructed:

Cb match - sub-samples of; Cb (100% classification)
Em match - sub-samples of; Cb & Em
Mg match - sub-samples of; Mg & Sp
Pm match - sub-samples of; Em, Pm, Pq & Ps
Pq/Ps match - sub-samples of; Em, Pq & Ps
Sp match - sub-samples of; Sp (100% classification)
Xs match - sub-samples of; Cb, Mg & Xs

The subsequent classifications of these sub-samples may be seen in Table 5.10(a) to (e). Sub-samples Cb and Sp however can be classified immediately. These sub-samples again have classification rates over 90% and indeed more often than not at 100%. However, X.suaveolens and M.graeffeana are still poorly classified in sub-sample Xs but this is less significant as both C.barbatum and M.graeffeana are represented separately in other sub-samples (Table 5.10(a) and (b) respectively) with high classification rates. Here it may be plausible to ignore M.graeffeana from this classification as there is only a small chance that the 'unknown' is in this class. This is not an ideal situation. There will inevitably be pollen taxa in the same sub-sample with very similar texture. It is quite clear that more work is required on sensitive texture measures to separate such taxa.

Table 5.10(a) : Texture classification of Cb & Em

	% Matched to Taxa	
Taxa/	Cb	Em
Cb	100	-
Em	-	100

Table 5.10(b) : Texture classification of Mg & Sp

	% Matched to Taxa	
Taxa/	Mg	Sp
Mg	94	6
Sp	-	100

Table 5.10(c) : Texture classification of Em, Pm, Pq & Ps

	% Matched to Taxa			
Taxa/	Em	Pm	Pq	Ps
Em	100	-	-	-
Pm	-	100	-	-
Pq	7	2	91	-
Ps	-	-	-	100

Table 5.10(d) : Texture classification of Em, Pq & Ps

	% Matched to Taxa		
Taxa/	Em	Pq	Ps
Em	100	-	-
Pq	6	94	-
Ps	-	-	100

Table 5.10(e) : Texture classification of Cb, Mg, & Xs

	% Matched to Taxa		
Taxa/	Cb	Mg	Xs
Cb	81	15	4
Mg	2	57	41
Xs	6	28	67

In conclusion it must be noted that the pollen used in this classification scheme are all fresh pollen. The usefulness of this type of scheme in areas of palynology such as the composition of modern pollen rain is more or less proven. However, the usefulness of this scheme for the identification of fossil pollen is still unknown. This system is probably applicable to the younger fossil samples, but the pollen in older fossil sediments having been compressed and damaged tend to lose their shape and therefore this classification scheme will require modifications. Even with the loss of shape information the size of the grains may still be of some use. This is mainly speculation and awaits further investigation.

Chapter 6

Other Techniques and Future Work

6.1 Location of Optimum Areas for Texture Analysis

In all previous research (Langford[25], Waterhouse[45] and Rutter[40]) regions of the pollen sampled for texture analysis have been selected manually. This would not be acceptable in any final automated system.

Before outlining any technique for automatic texture region location it is necessary to define how a texture region is selected manually. Firstly, it is important to avoid sampling macro-elements of the pollen grain such as colpi or pori. Secondly the area selected must have as wide a range of grey-levels as possible, thus maximising the information content of the image. However, areas with rapid changes in grey-levels may indicate the edges of colpi or pori and must also be avoided. Also areas to which small particulate material have become attached must be excluded as the texture is obscured. Overly dark or bright areas of the image must also be avoided. Waterhouse[45] also noted that different areas of the pollen surface have different textures. Also texture changes as the pollen surface curves away from the viewer and therefore samples taken towards the pollen edges are different from the centre.

As already stated in chapter 5 the texture measures used are taken from pollen regions selected automatically. So how are the worst texture areas outlined above excluded? The simple answer is that they are not excluded! The basic assumption being that the worst texture areas are as much a part of the

distinctive texture of a particular pollen taxa as the best areas. The area selected is quite simply the whole image after the SEM has magnified the pollen to such an extent that its edges are lost. It effectively samples the largest square region that can fit completely inside the pollen grain. This eliminates some of the problem areas of a pollen grain such as the edges that are curving away and averages out the changes in texture from region to region over the pollen surface. For this study the selection of the square region is done manually. As the location of the pollen edge is known (see chapter 4 and 5) it should be a relatively easy task to do this with computer controlled magnification and stage movement.

Although the brightness and contrast used for texture analysis are more or less constant an automatic tuner would be useful. To adjust these settings manually the histogram of the image is used and the balances changed until there are few maximum or minimum grey-levels. No subjective visual estimation of the image balance is used. Again these settings should be achieved easily by a computer that has been given some initial settings and boundaries.

To use the whole (256x256) image to construct the texture measures is computationally very time consuming. About 7225 pixels (equivalent to an 85x85 image) are sampled within the whole image to produce the texture measures used in chapter 5. This reduced the significance of the overall classification only slightly. It must be noted here that the displacement

vectors used in the construction of the Haralick measures are not limited to 1, 4 and 8 as used by all other researchers, but range from 1 to 17. The variables that worked best for each sample are subsequently used.

As pollen size is relatively constant there is little concern that the subtle differences in magnification of the texture within a taxa would effect the results significantly. Indeed the taxa most often confused, namely Macaranga graeffeana (Mg) and Xylosma suaveolens (Xs) have the smallest size ranges (Fig.5.06 to Fig.5.08). Fossil pollen may present a problem if their sizes fluctuate greatly due to preservation damage. This has yet to be demonstrated.

The quality of the texture measures may be improved further by using an initial scan of the whole image to eliminate areas of extreme brightness and darkness. The image could be split into small overlapping regions and the grey-level histograms of each analysed. The regions with extremes of grey-level intensities (possibly taken as a mean) would be ignored. This may then exclude dark areas such as colpi, pori or even damaged areas. Bright areas such as the regions close to the pollen edge that tend to be highly illuminated may also be detected and ignored. Such calculations may prove to be computationally too time consuming.

6.2 Neural Networks

Rumelhart et al[38] begin their book by posing the question: 'What makes people smarter than machines?' Even though people are not as fast or precise as computers for numerical calculations they are far better at perceiving objects and noting their relationships to each other, understanding language, as well as a wide variety of other cognitive tasks. People have the ability to learn tasks so that they can improve their accuracy and fluency at doing them, machines cannot.

A classic response to the above question is that it is the 'software' that brains follow that is important. Therefore all that is needed to produce a learning and thinking machine is the correct computer program.

However, Rumelhart et al[38] take the view that:

"...people are smarter than today's computers because the brain employs a basic computational architecture that is more suited ... to processing tasks that people are good at."

So what is this computational architecture that the brain employs? The brain is a highly complex parallel mechanism composed of about 8 billion processing units (neurons), each one of which is connected to thousands of others. The parallel nature of the brain explains its great ability to solve pattern recognition problems. Speed is not the important factor -

parallelism is. The serial nature of an electronic computer lends itself to rapid numerical computations but not to pattern recognition.

The approach of neural computing is to apply the basic principles of brain architecture to computer systems. By modelling the brain's basic systems it should be possible to make computers learn solutions to parallel problems such as pattern recognition.

There are numerous models that have been proposed for work on neural networks. The network explored here is the layered feed-forward network model which is one of the most popular (Rangwala & Dornfeld[37]). This provides a very flexible model able to coup with more complex tasks and also has a sophisticated learning scheme for network training.

Fig.6.01 shows the structure of a simple layered feed-forward network. It consists of an input layer, any number of hidden layers and an output layer, each of which is made up of individual processing units or nodes. The input layer serves only to distribute input values received to the hidden layer(s). An input signal is propagated through a network a layer at a time and is modified as it goes by simple equations. The input signal to each node is dependent on the output signal from each node in the layer below, combined with the weight (or strength)

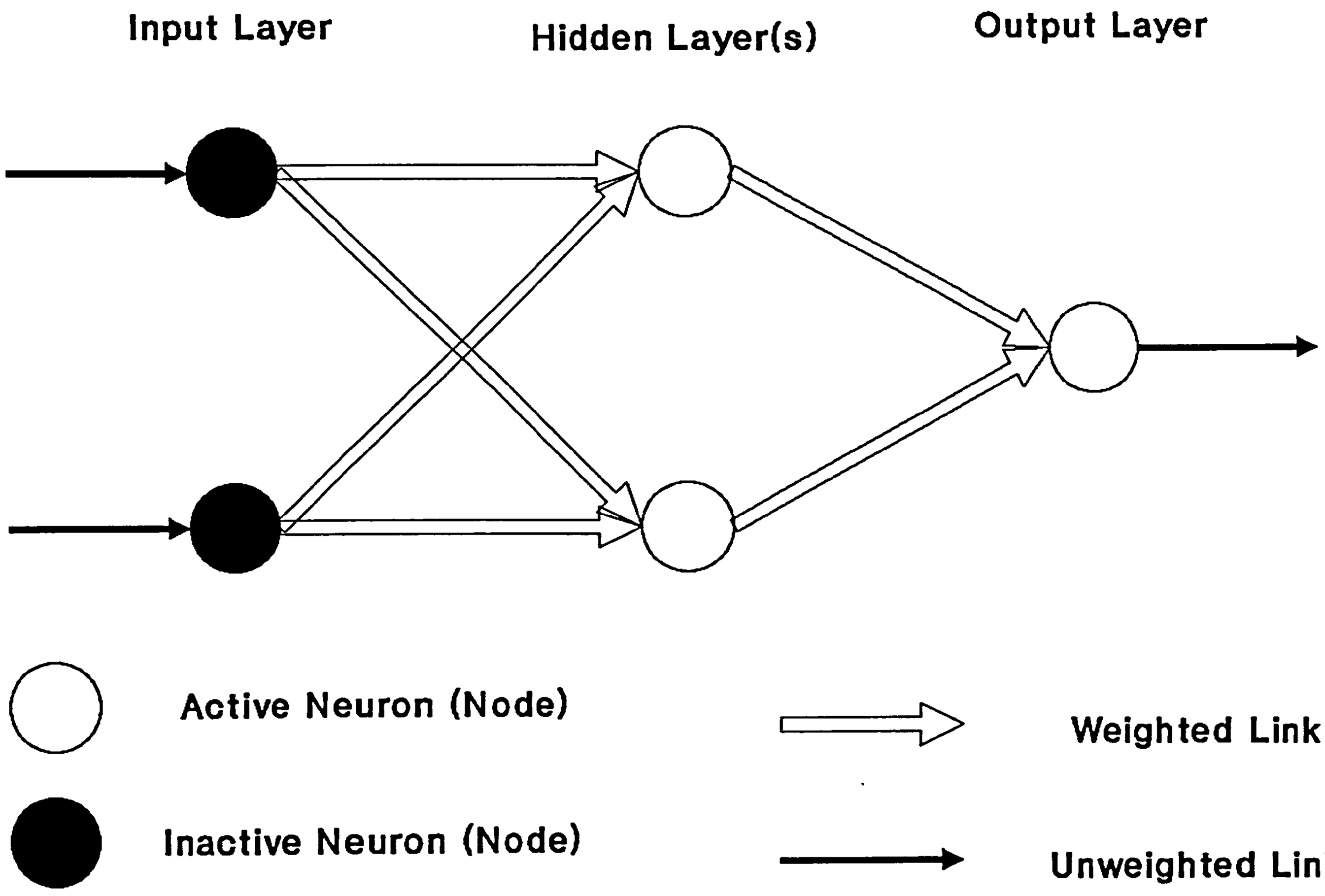


Fig.6.01 Feedforward Neural Network Structure

of the connection between them. A node also has an activity level (or threshold) that is added to the total input signal. This may be expressed using the following equation:

$$net_{ik} = \sum_{j=1}^{N_{k-1}} [W_{ijk} \cdot O_{jk-1}] + t_{ik}$$

net_{ik} = input to the i^{th} node in the k^{th} layer.

N_{k-1} = number of nodes in the $(k-1)^{th}$ layer.

W_{ijk} = weight of connection between j^{th} node in $(k-1)^{th}$ layer to the i^{th} node in the k^{th} layer.

O_{ik} = output from the i^{th} node in the k^{th} layer.

t_{ik} = activity level of the i^{th} node in the k^{th} layer.

$k = 1$ for input layer.

The output signal from a node is then calculated by applying a nondecreasing nonlinear function to the inputs of that node. The function used in this case is a sigmoid function. The output from a node may therefore be expressed thus:

$$O_{ik} = \frac{1}{1 + e^{-net_{ik}}}$$

To produce a working neural network the values of all the weights and activity levels have to be set to small positive and negative random numbers. Then the neural network is taught by presenting it with a set of input and desired output patterns. Firstly, an input is transmitted through the network (as

described above) and the resulting signals at the output layer are compared to the desired output patterns. If the actual output is identical to the desired output then no learning need occur, otherwise the weights between each node (W_{ij}) and activity levels of each node (t_{it}) are changed to reduce the difference between actual and desired outputs (ie the error). Further input/output patterns are applied until the error is driven close to zero. Thus the network is taught to map a set of input patterns to a set of output patterns. The network should then be able to map similar patterns to the correct output pattern.

The learning rule used to change the weights during the learning phase is the generalised delta rule. The error at the output layer of the network is propagated backwards and used to change the weights and activity levels of the nodes as it goes. In the following text the computations for changing weights are used. To apply this method to change the activity level of a node the activity level must be treated as a weight into that node from an imaginary node which has a constant output of 1.

All the weights are changed a layer at a time from the output to input layer using the following formula:

$$\Delta W_{ji}(x+1) = \eta(\delta_{jk} O_{ik-1}) + \alpha \Delta W_{ji}(x)$$

x = the number of pattern presentations.

η = learning rate $(0 < \eta \leq 1)$.

α = momentum rate $(0 \leq \alpha \leq 1)$.

δ_{jk} = error signal of j^{th} node in k^{th} layer.

O_{ik-1} = output from i^{th} node $(k-1)^{th}$ layer.

The function $\alpha \Delta W_{ji}(x)$ is a momentum term that prevents the network from falling into a local minima by relating the present weight change to the last.

There are two equations for the error signal (δ_{jk}), the first is used for the output layer:

$$\delta_{jk} = (d_{jk} - O_{jk}) \cdot O_{jk} \cdot (1 - O_{jk})$$

d_{jk} = desired output of the j^{th} node in output layer k.

The delta value of each output node is used to calculate all the weight changes for all the connections feeding into that node from the layer above. The second delta calculation for all hidden layers is:

$$\delta_{jk} = O_{jk} \cdot (1 - O_{jk}) \cdot \sum_{i=1}^{N_{k+1}} \delta_{ik+1} W_{ijk+1}$$

N_{k+1} = number of nodes in layer $(k+1)$.

δ_{ik+1} = delta value for i^{th} nodes in previous layer $(k+1)$.

W_{ijk+1} = weight between j^{th} node in present layer and i^{th} node in previous layer $(k+1)$.

Once all the weights and activity levels have been modified in this back propagation stage, a new set of input patterns are presented to the network. These forward and backward passes are repeated until the error at the output layer falls below a predefined level for all input patterns. The total error of the network for pattern, p, is given by:

$$E_p = \frac{1}{2} \sum_{j=1}^{N} (d_{jk} - O_{jk})^2$$

d_{jk} = desired output for pattern p of the j^{th} node in output layer k.

O_{jk} = actual output for pattern p of the j^{th} node in output layer k.

N = number of nodes in output layer.

Now that the basic principles of neural networks have been explained their usefulness for pollen texture classification can be explored. The idea is to match a set of input texture

measures to a set of output nodes each of which represent a single pollen taxa. The desired output is 1 at the node representing the present input and 0 at the rest.

Three training sets of input measures are used for each pollen taxa. These are the mean together with the mean plus the standard deviation and the mean minus the standard deviation, of each variable. As there are 18 samples of each variable, each sample only has a limited influence on the overall mean and standard deviation. The maximum value of the mean plus standard deviation for each variable in all taxa is then used to normalise both the training set and the original samples (Davalo and Naim[8]). The training variables all lie in the range 0 to 1. Some of the original samples however can take on larger values as they can exceed the normalising value. The particular variables used were those selected by the variable selection procedure used with the linear classifier (see chapter 2).

The first test of the neural network was on two taxa; Canthium barbatum (Cb) and Elephantopis mollis (Em). Using the linear classifier with only a single variable both these taxa had a 100% classification rate (Table 5.10(a)). The neural network constructed had a structure of 1-1-2 ie one input node (for the single variable), a single hidden node and two output nodes (one for each taxa).

This network learnt very quickly and achieved the target error of 0.001 in less than 100 iterations. A 100% classification was achieved when the original texture samples were passed through the network. This is not at all surprising as this had already been attained using a linear classifier.

The next test was to try and separate Pritchardia minor (Pm), Passiflora quadrangalaris (Pq) and Pseudoelephantopis spicatur (Ps). Using the linear classifier both Pm and Ps had achieved a 100% classification rate. However, Pq had only a 94% success rate due to confusion with Pm. A network structure of 4-8-3 was used and took just 445 iterations to reach a target error of 0.01. Fig.6.02 shows the learning curve of this network. All the original samples passed through the trained network were correctly classified. Even the worst classification is very convincing having the following output pattern:

```
Output node 1 : 0.903611        desired : 1.000000
Output node 2 : 0.011722        desired : 0.000000
Output node 3 : 0.219977        desired : 0.000000
```

A final test was done on the very poor classifications of Canthium barbatum (Cb), Macaranga graeffeana (Mg), Fungal spore (Sp) and Xylosma suaveolens (Xs) (see Table 5.09(e)). A network structure of 16-80-4 was constructed with a target error of 0.01. Although the target error was achieved the classification of original samples was particularly poor for Mg and Xs. The network failed completely to identify Mg and only classified

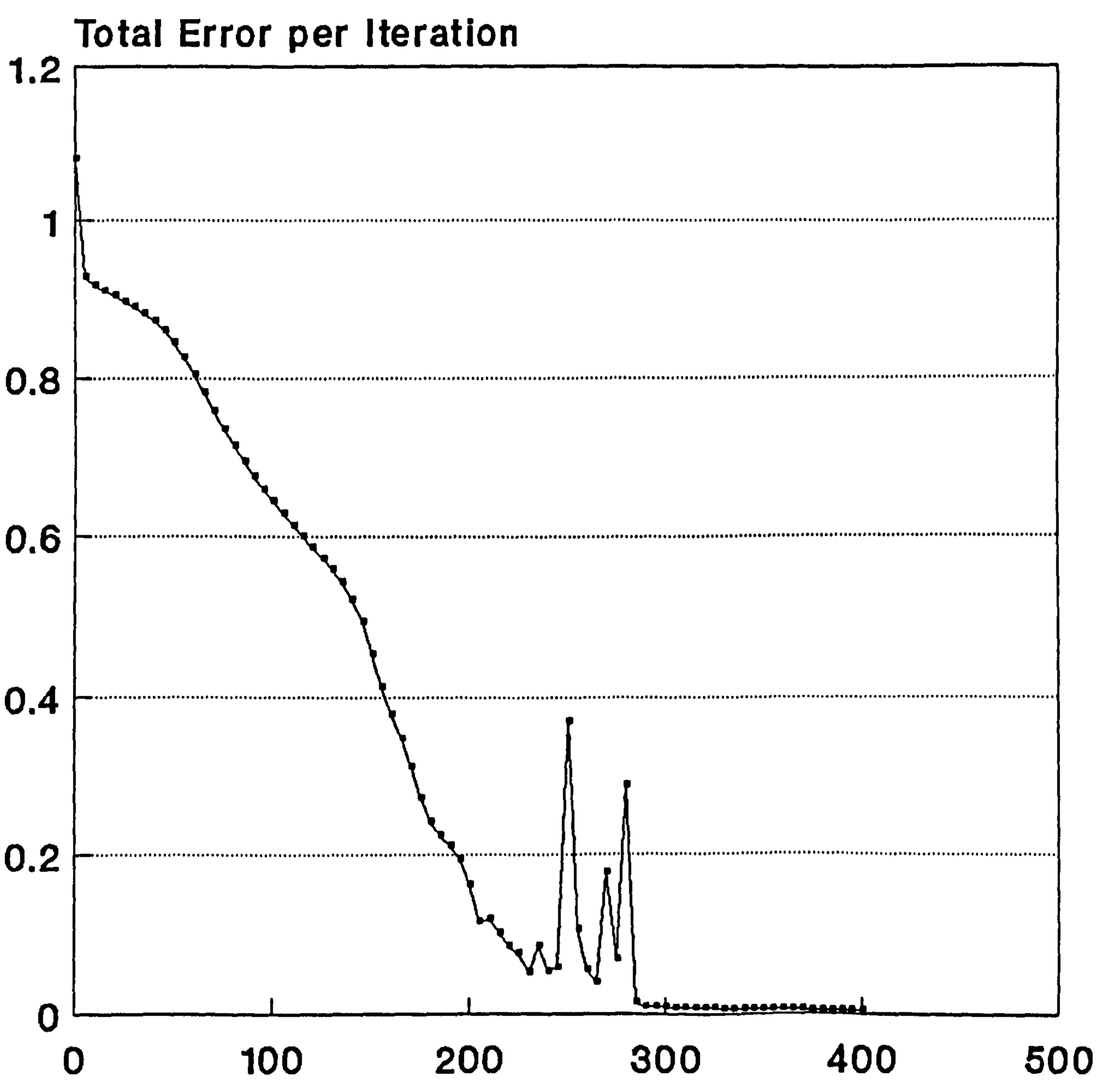


Fig.6.02 Learning Curve for Pq, Ps & Pm

Xs correctly 33% of the time. The classification rates of Cb and Sp both increased by about 12% however. The structure required to achieve these classifications was also overly bulky.

These results serve to illustrate further that the variables used to separate Mg from Xs are completely inadequate and further research is required to improve the situation.

The use of neural networks needs further exploration to discover other areas within this project to which they may be applied.

6.3 Future Work and Concluding Remarks

This thesis has demonstrated the success of the classification scheme proposed in chapter 1. Pollen may now be located in the image and their edges found. Optimum areas of the pollen exine for texture analysis have now been selected automatically and the pollen that was initially poorly classified have been split into small highly classified sub-groups by shape analysis. A study of neural networks has shown a way that these classifications may be improved still further.

There are several areas however that require further investigation. The hardware control of the microscope itself

still awaits work - similar systems are already available, for example, the automatic analysis of cervical smears (Watanabe et al[46]).

As well as testing other methods of texture analysis it may be useful to see whether detrital material has a distinctive texture that may be used to exclude it from a sample. Also the assumption that spinose pollen cannot be separated using texture measures becomes more doubtful as other macro-elements of the pollen exine have been shown to have little effect on their classification. The texture analysis of fossil pollen also requires further investigation.

The final problem is in the use of the SEM. Every image used in this study took at least 45 seconds to construct in the SEM frame-store due to the slow scan rate of the electron beam. Images captured using faster scan rates are of poorer quality. The slow scan is less of a problem at the moment as processing speeds of the images are slower than the image capture. However for a final system this will be less acceptable. Therefore it may be advisable to attempt pollen analysis using either a laser-scan or conventional light microscope where image capture is instantaneous.

References

[1] Aleksander, I., & Morton, H. (1990)
"An introduction to neural computing,"
Chapman and Hall, London.

[2] Beale, R., & Jackson, T. (1990)
"Neural computing: an introduction,"
Adam Hilger, Bristol.

[3] Beney, J.E. (1991)
"The automation of palynology: locating the object,"
Final Year Research Project #9106, Dept. of Electronic Engineering, University of Hull.

[4] Casey, R.G. (1970)
"Moment normalization of hand printed characters,"
IBM Journal of Research and Development, vol.14, pp.548-557.

[5] Data Translation Ltd (1990)
"1990 product catalogue,"

[6] Data Translation Ltd (1985)
"DT2803 user's manual,"

[7] Data Translation Ltd (1985)
"VideoLab user's manual,"

[8] Davalo, E., & Naim, P. (1991)
"Neural networks,"
translated by Rawthorne, A., MacMillan Education, London.

[9] Don, M., et al (1984)
"Metal surface inspection using image processing techniques,"
IEEE Trans. Systems, Man & Cybernetics, vol.SMC-14, pp.139-146.

[10] Dudani, S.A., Breeding, K.J. & McGhee, R.B. (1977)
"Aircraft identification by moment invariants,"
IEEE Trans. Comput., vol.C-26, pp.39-45.

[11] Flenley, J.R. (1968)
"The problem of pollen recognition,"
in Clowes, M.B., & Penny, J.P. (ed.) 'Problems in picture interpretation,' C.S.I.R.O., Canberra. pp.141-145.

[12] Forster, R.M. (1986)
"Palynomorph purification and fractionation by equilibrium density gradient centrifugation,"
Unpublished M.Sc. Thesis, University of Hull.

[13] Forster, R.M., & Flenley, J.R. (1989)
"The application of density gradient centrifugation to palynology,"
School of Geography & Earth Resources, University of Hull, Miscellaneous Series No.35

[14] Freeman, H. (1961)
"On the encoding of arbitrary geometric configurations,"
IRE Transactions on Electronic Computers, EC-10, pp.260-268.

[15] Freeman, H. (1961)
"Techniques for the digital computer analysis of chain-encoded arbitrary plane curves,"
Proceedings of the National Electronics Conference, vol.17, pp.421-432.

[16] Freeman, H. (1961)
"A technique for the classification and recognition of geometric patterns,"
Proceedings of the 3rd International Congress on Cybernetics, Namur, Belgium, pp.348-368.

[17] Freeman, H. (1974)
"Computer processing of line-drawing images,"
Computing Surveys, 6, #1, pp.57-97.

[18] Freeman, H. (1980)
"Lines, curves and the characterization of shape,"
Report no. IPL-TR-80-004, Image Processing Laboratory, Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York.

[19] Gundersen, K. (1990)
"Pollen analysis,"
Final Year Research Project #9024, Dept. of Electronic Engineering, University of Hull.

[20] Guppy, J. et al (1973)
"Further developments in computer assistance to pollen identification,"
Special Publication #4, Geological Society of Australia, pp.201-206

[21] Hand, D.J. (1981)
"Discrimination and classification,"
Wiley, Chichester.

[22] Haralick, R.M., Shanmagan, K., & Dinstein, I. (1973)
"Textural features for image classification,"
IEEE Trans. Sys., Man & Cybern., vol.SMC-3, pp.610-621.

[23] Hu, M.K. (1962)
"Visual pattern recognition by moment invariants,"
IRE Trans. Information Theory, vol.IT-8, pp.179-187.

[24] Kittler, J., & Illingworth, J. (1985)
"On threshold selection using clustering criteria,"
IEEE Trans. Sys., Man and Cybern., vol.SMC-15, no.5, pp.652-655

[25] Langford, M., (1988)
"Some applications of digital image processing for automation in palynology,"
Ph.D thesis, Department of Electronic Engineering, University of Hull.

[26] Laws, K.I. (1979)
"Texture energy measures,"
Proc. Image Understanding Workshop, Nov. 1979, pp.47-51.

[27] Levine, M.D., & Nazif, A.M. (1985)
"Dynamic measurement of computer generated image segmentations,"
IEEE Trans. Pattern Anal. Machine Intell., vol.PAMI-7, pp.155-164

[28] Lowe, J.J., & Walker, M.J.C., (1984)
"Reconstructing quaternary environments,"
Longman.

[29] Maitra, S. (1979)
"Moment invariants,"
Proc. IEEE, vol.67, pp.697-699.

[30] McKay, R.J. (1977)
"Simultaneous procedures for variable selection in multiple discriminant analysis,"
Biometrika, vol.64, pp.283-290.

[31] Mirkin, G.R., & Bagdasaryan, L.L. (1972)
"The feasibility of identifying palaeontological objects with the aid of optical analysing systems,"
Palaeontological Journal, vol.6, pp.103-108

[32] Moore, P.D., & Webb, J.A. (1978)
"An illustrated guide to pollen analysis,"
Hodder & Stoughton, London.

[33] Otsu, N., (1979)
"A threshold selection method from gray-level histograms,"
IEEE Trans. Sys., Man, Cybern., vol.SMC-9, no.1, pp.62-66

[34] Pietikainen, M., Rosenfeld, A., & Davis, L.S. (1983)
"Experiments with texture classification using averages of local pattern matches,"
IEEE Trans. Sys., Man & Cybern., vol.SMC-13, pp.421-426.

[35] Prewitt, J.M.S. (1970)
"Object enhancement and extraction,"
in Lipkin, B.S., and Rosenfeld, A. (eds.), 'Picture processing and psychopictorics,' Academic, NewYork, pp.75-149.

[36] Psenner, R. (1991)
"Detection and sizing of aquatic bacteria by means of epifluorescence microscopy and image analysis,"
Image Enhancement and Analysis, October 1991, pp.13-15.

[37] Rangwala, S.S. & Dornfeld, D.A. (1989)
"Learning and optimization of machining operations using computing abilities of neural networks,"
IEEE Trans. Sys., Man & Cybern., vol.SMC-19, pp.299-314.

[38] Rumelhart, D.E., McClelland, J.L., et al (1986)
"Parallel distributed processing, explorations in the microstructure of cognition, volume 1: foundations,"
MIT Press, London.

[39] Rutovitz, R., et al (1978)
"Computer-assisted measurement in the cytogenetic laboratory,"
pp.301-329, in Batchelor, B.G. (ed.), 'Pattern recognition: ideas in practice,' Plenum Press, New York.

[40] Rutter, B.W. (1989)
"Automatic identification of pollen grains by the analysis of surface texture: a review,"
Final Year Research Project #1876, Dept. of Electronic Engineering, University of Hull.

[41] Squires, R.H., (1970)
"A computer program for the presentation of pollen data,"
University of Durham.

[42] Tien, C.H. (1981)
"A note on invariant moments in image processing,"
IEEE Trans. Sys., Man & Cybern., vol.SMC-11, pp.831-834.

[43] Treloar, W.J. & Taylor, G.E. (1992)
"The application of difference code to the identification of pollen,"
Proceedings of the 11th IASTED Conference on Modelling, Identification and Control. pp.487-490.

[44] Walker, D., et al (1968)
"The computer assisted storage and retrieval of pollen morphological data,"
Pollen et Spores, vol.10, pp.251-262

[45] Waterhouse, H. (1988)
"Automated identification of fossil pollen grains by texture analysis : a preliminary investigation,"
M.Sc Thesis, Department of Geology, University of Hull.

[46] Watanabe, S., & the CYBEST Group (1974)
"An automated apparatus for cancer prescreening: CYBEST,"
Computer Graphics and Image Processing, vol.3, pp.350-358

[47] Weszka, J., Dyer, C., & Rosenfeld, A. (1976)
"A comparative study of texture measures for terrain classification,"
IEEE Trans. Sys., Man & Cybern., vol.SMC-6, pp.269-285.

[48] Whittow, J., (1984)
"Dictionary of physical geography,"
Penguin Books.

[49] Wong, A.K.C., & Sahoo, P.K., (1989)
"A gray-level threshold selection method based on maximum entropy principle,"
IEEE Trans. Sys., Man, Cybern., vol.SMC-19, no.4, pp.866-871

Plates

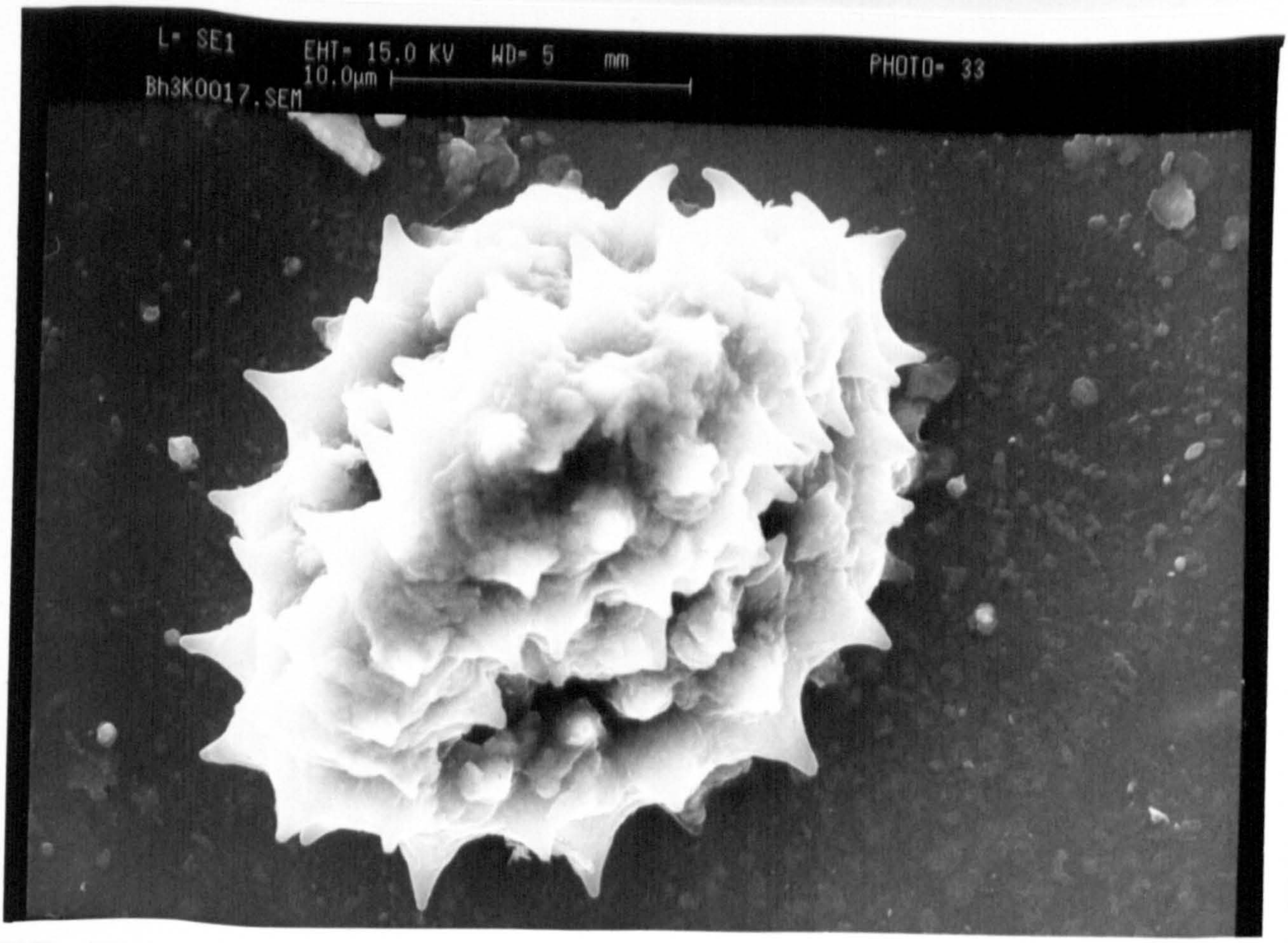


Plate 1 : Bidens hendersonensis

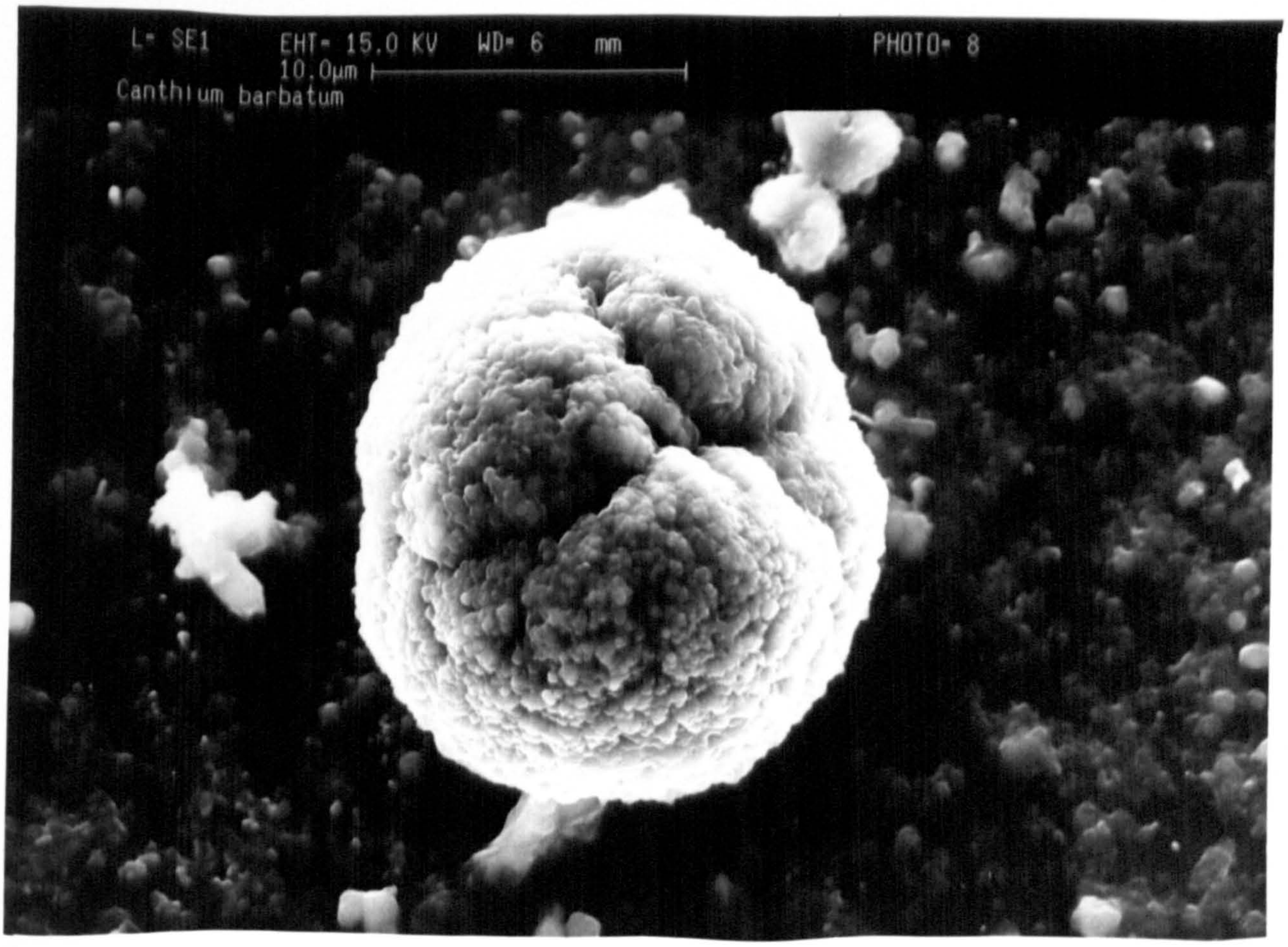


Plate 2(a) : Canthium barbatum

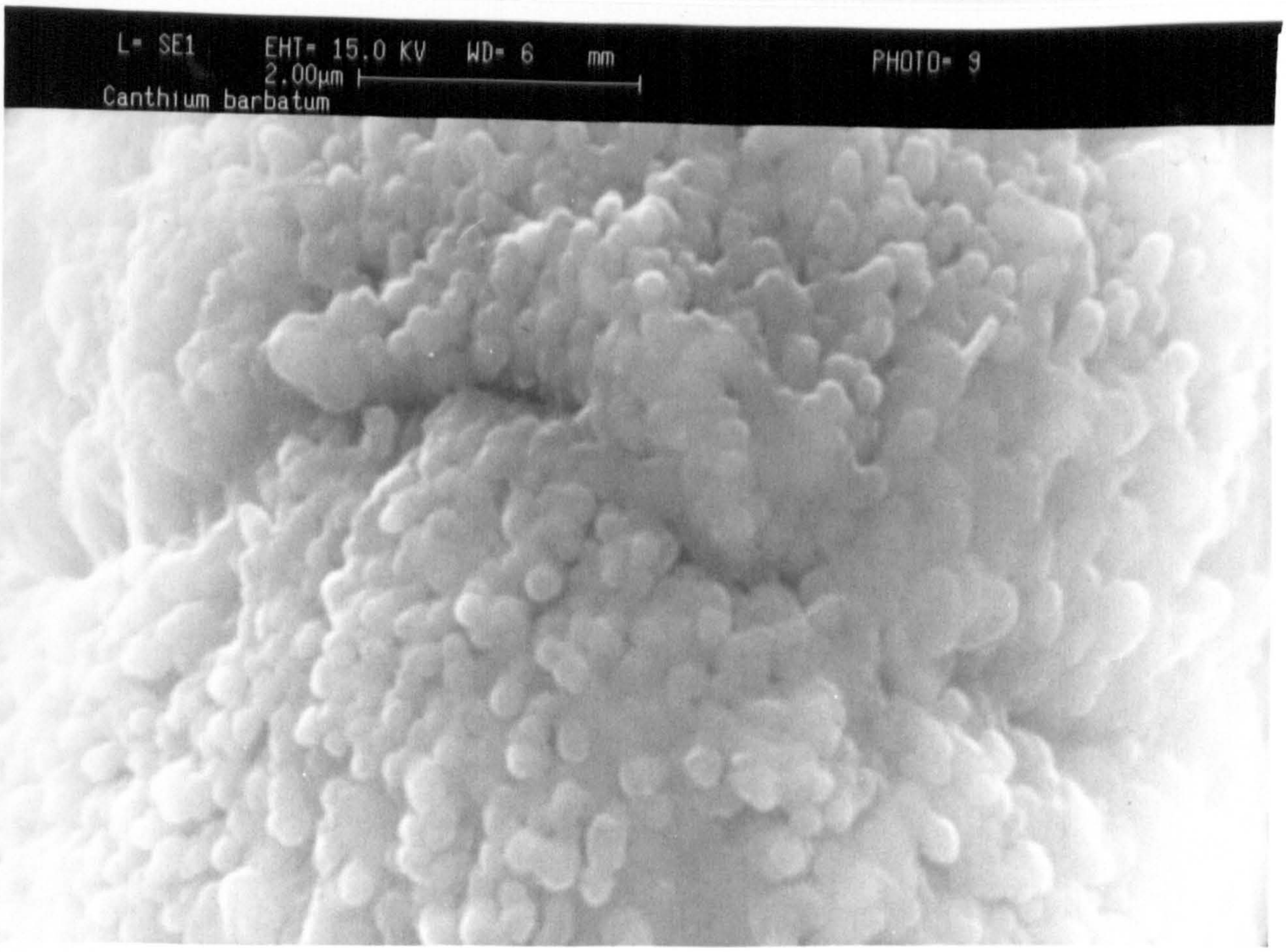


Plate 2(b) : Canthium barbatum (Texture)

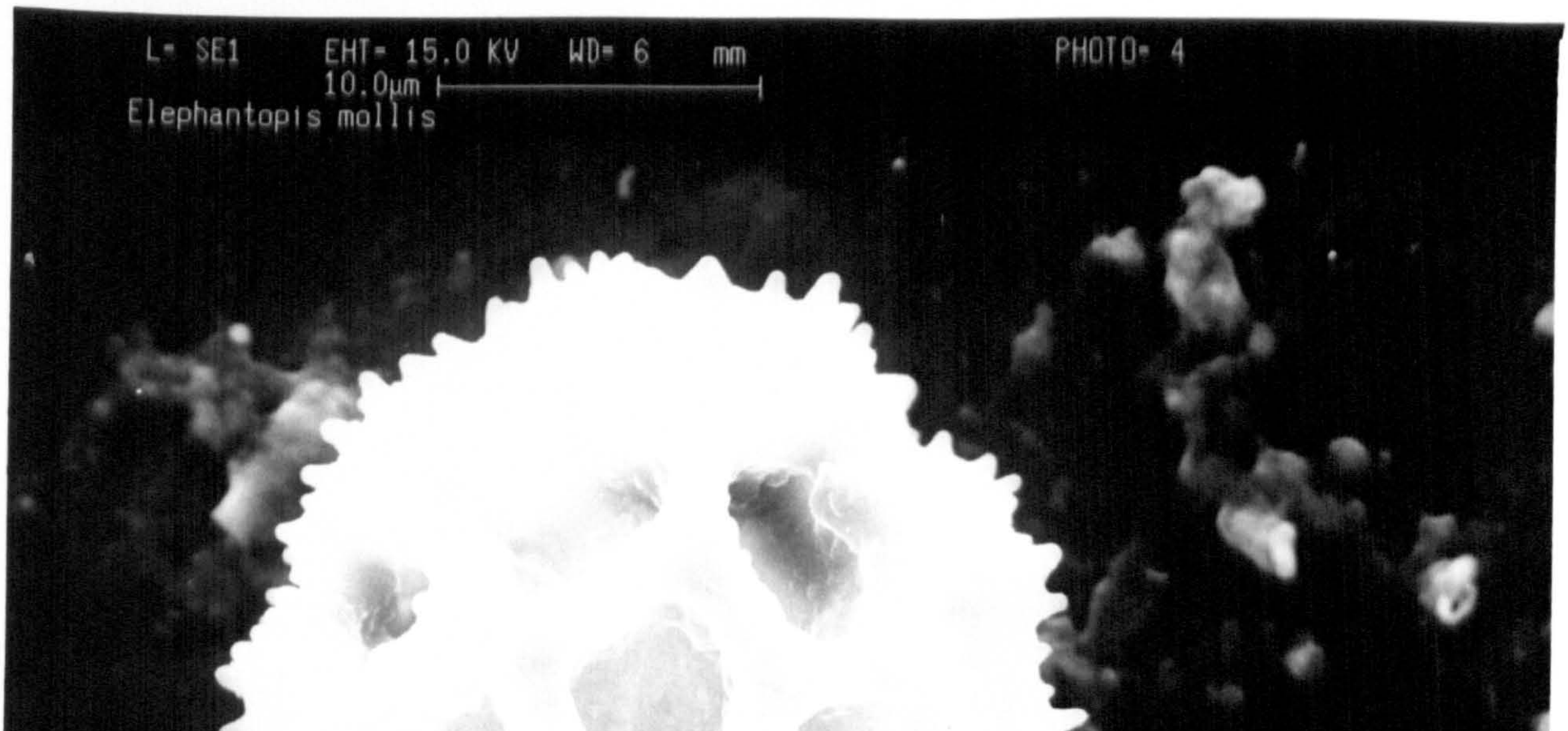


Plate 3(a) : Elephantopsis mollis

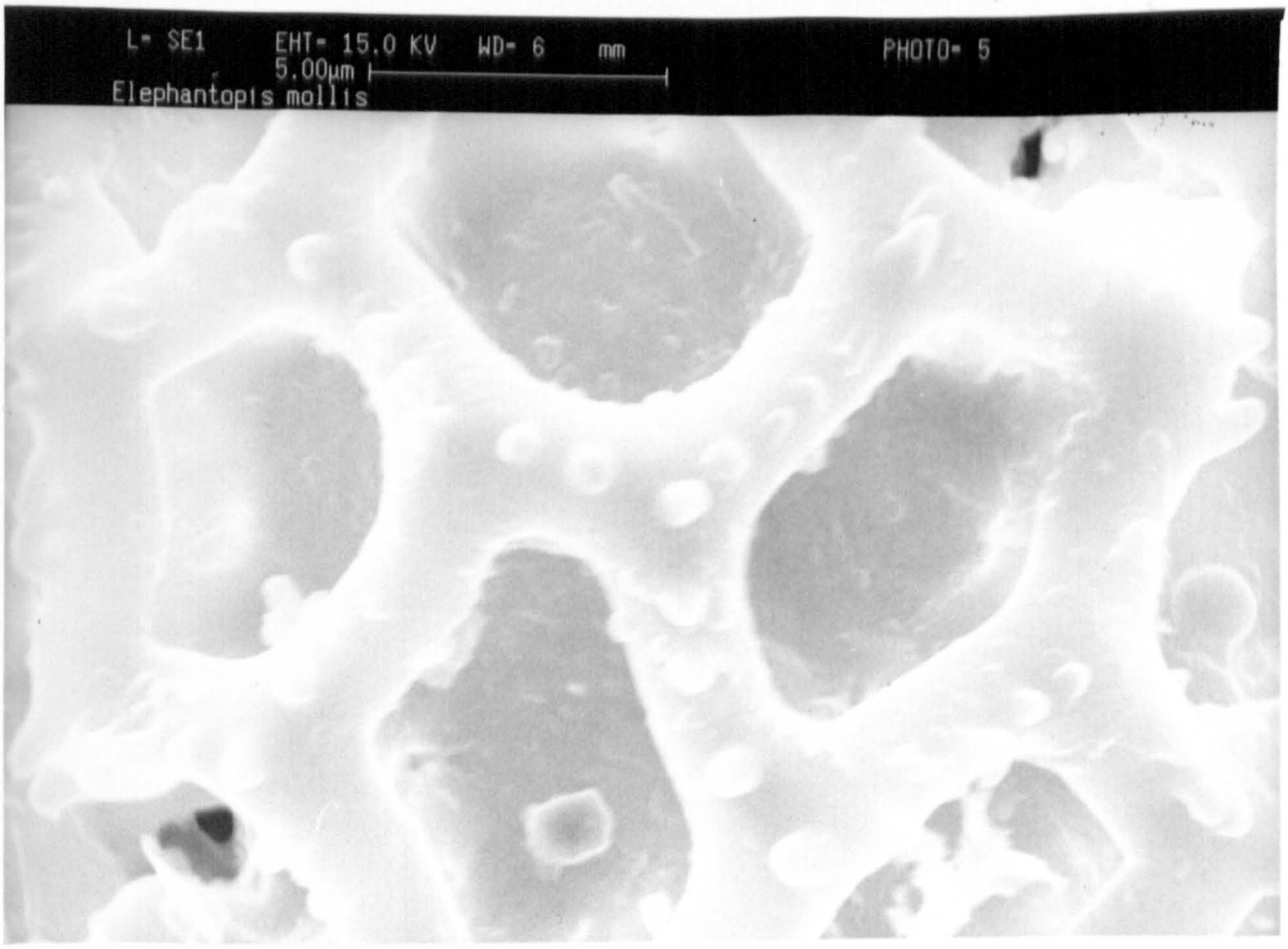


Plate 3(b) : Elephantopsis mollis (Texture)

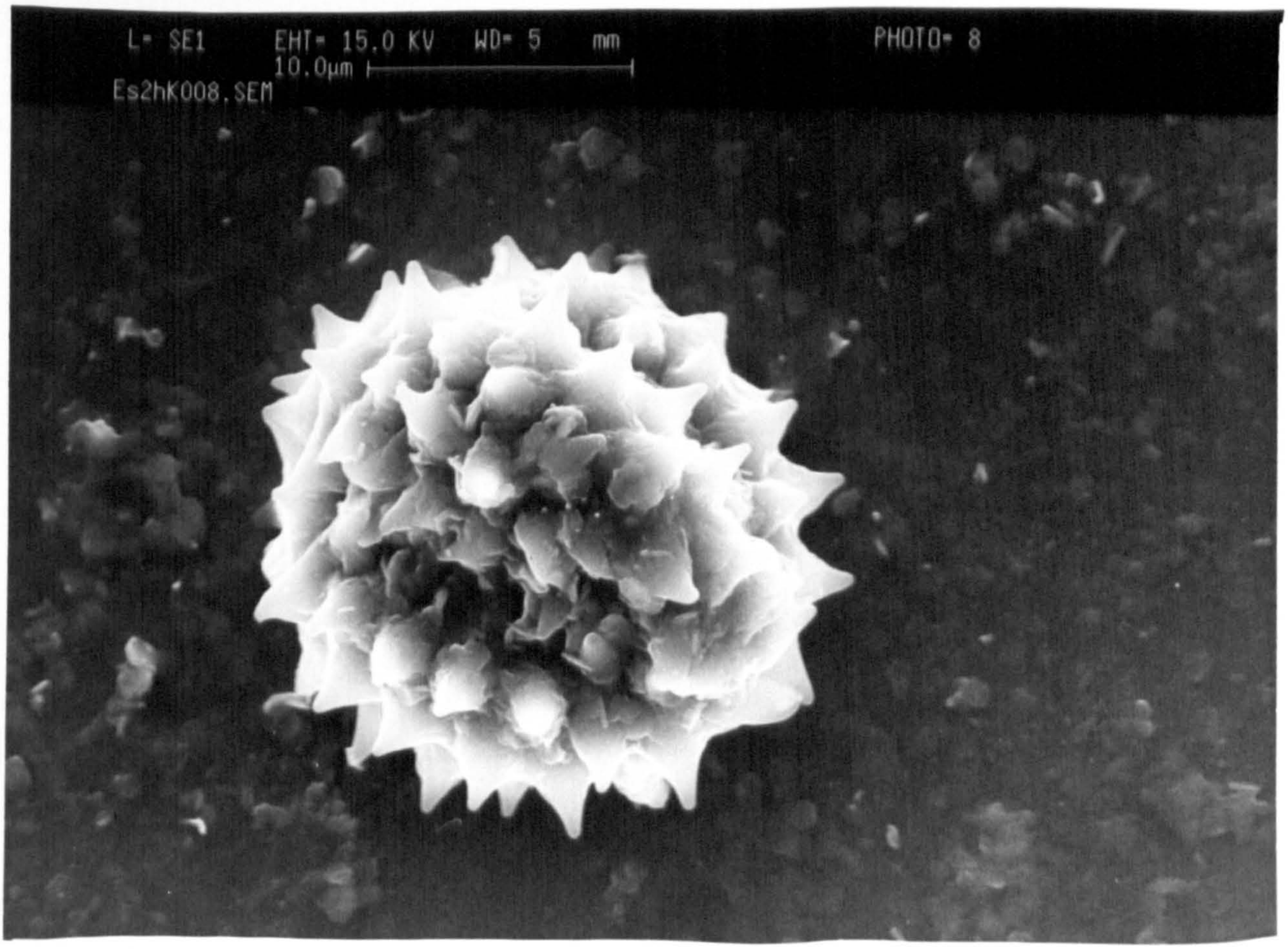


Plate 4 : Emilia sonchifolia

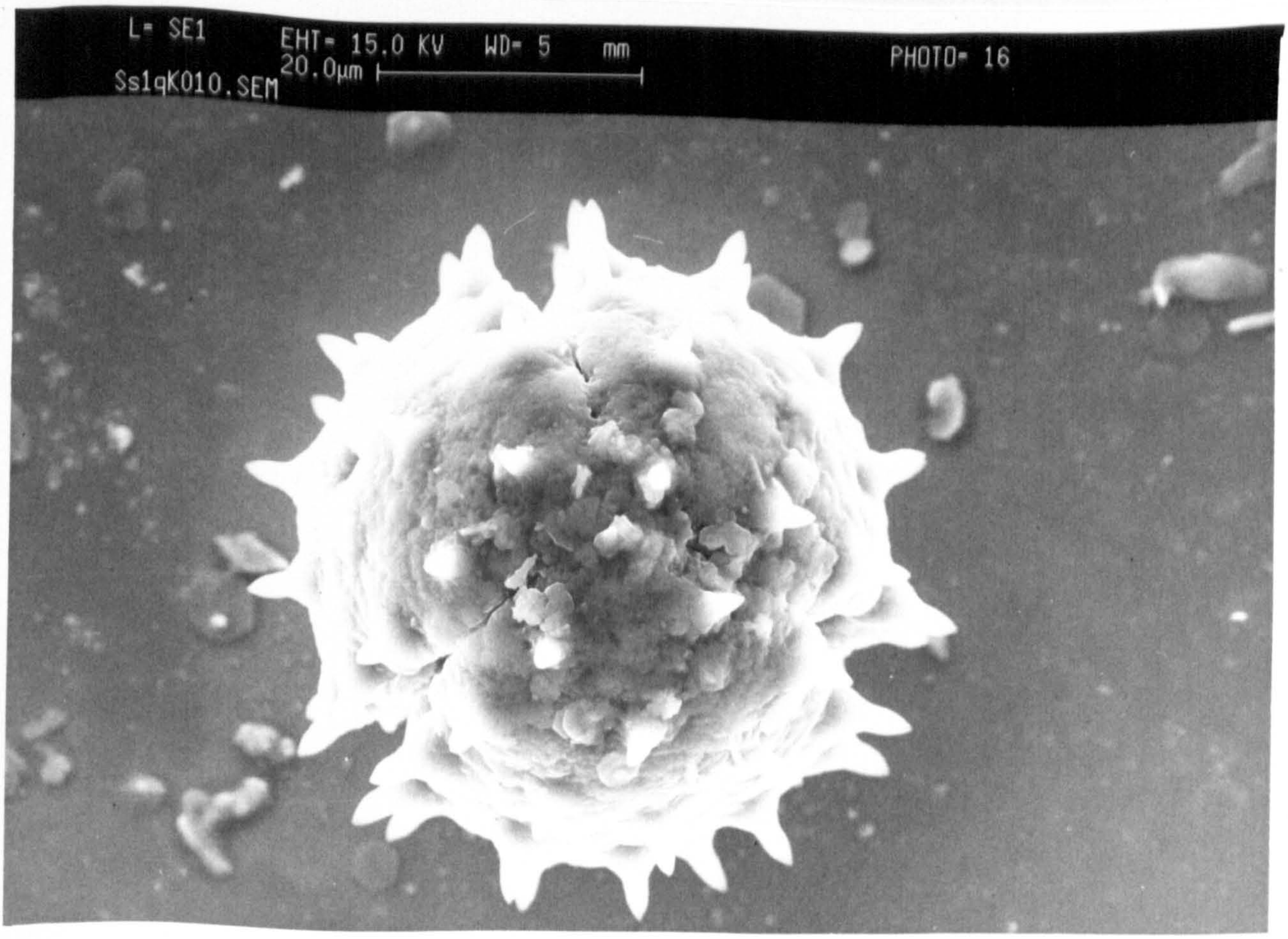


Plate 5 : Fitchia speciosa

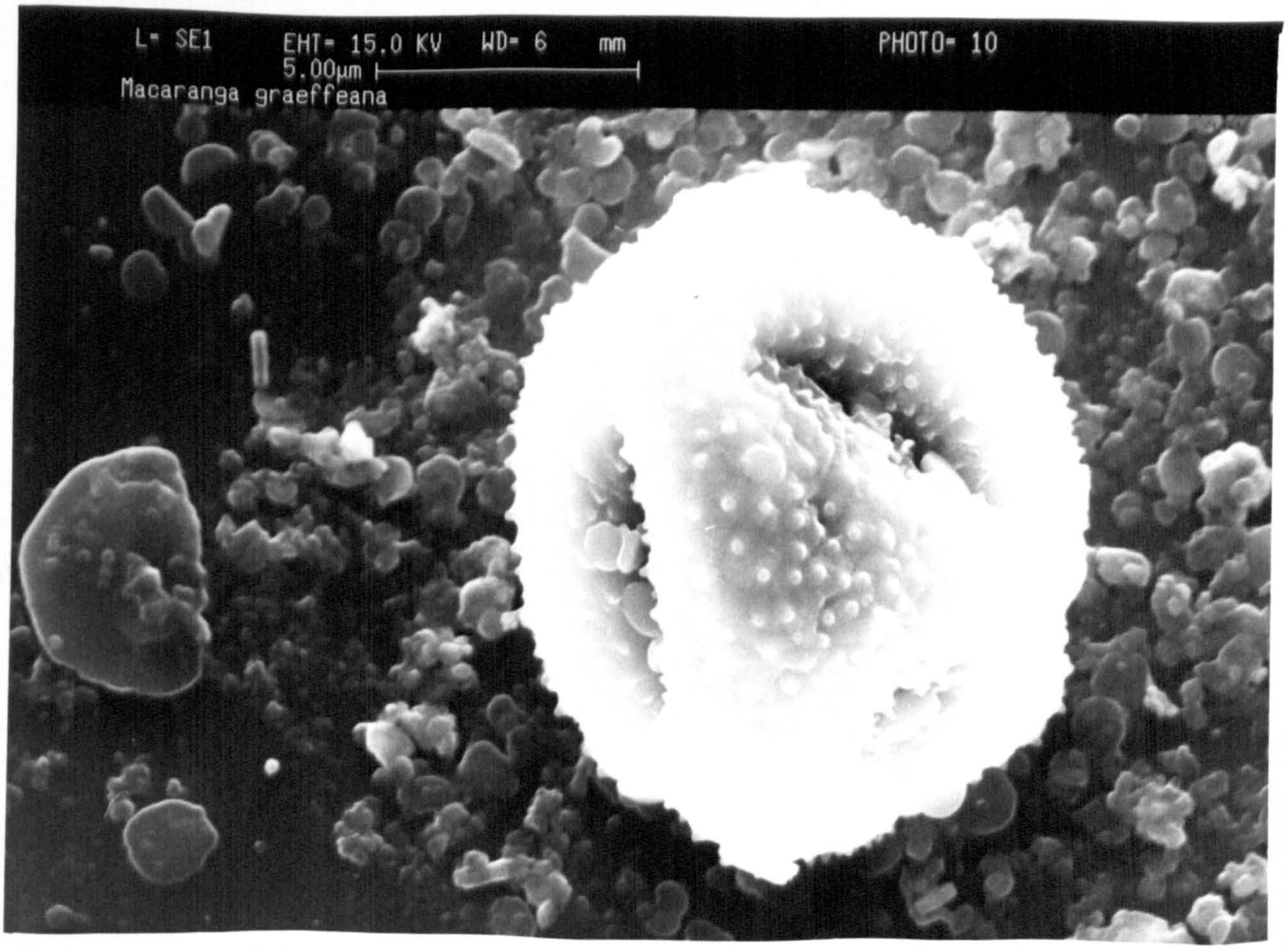


Plate 6(a) : Macaranga graeffeana

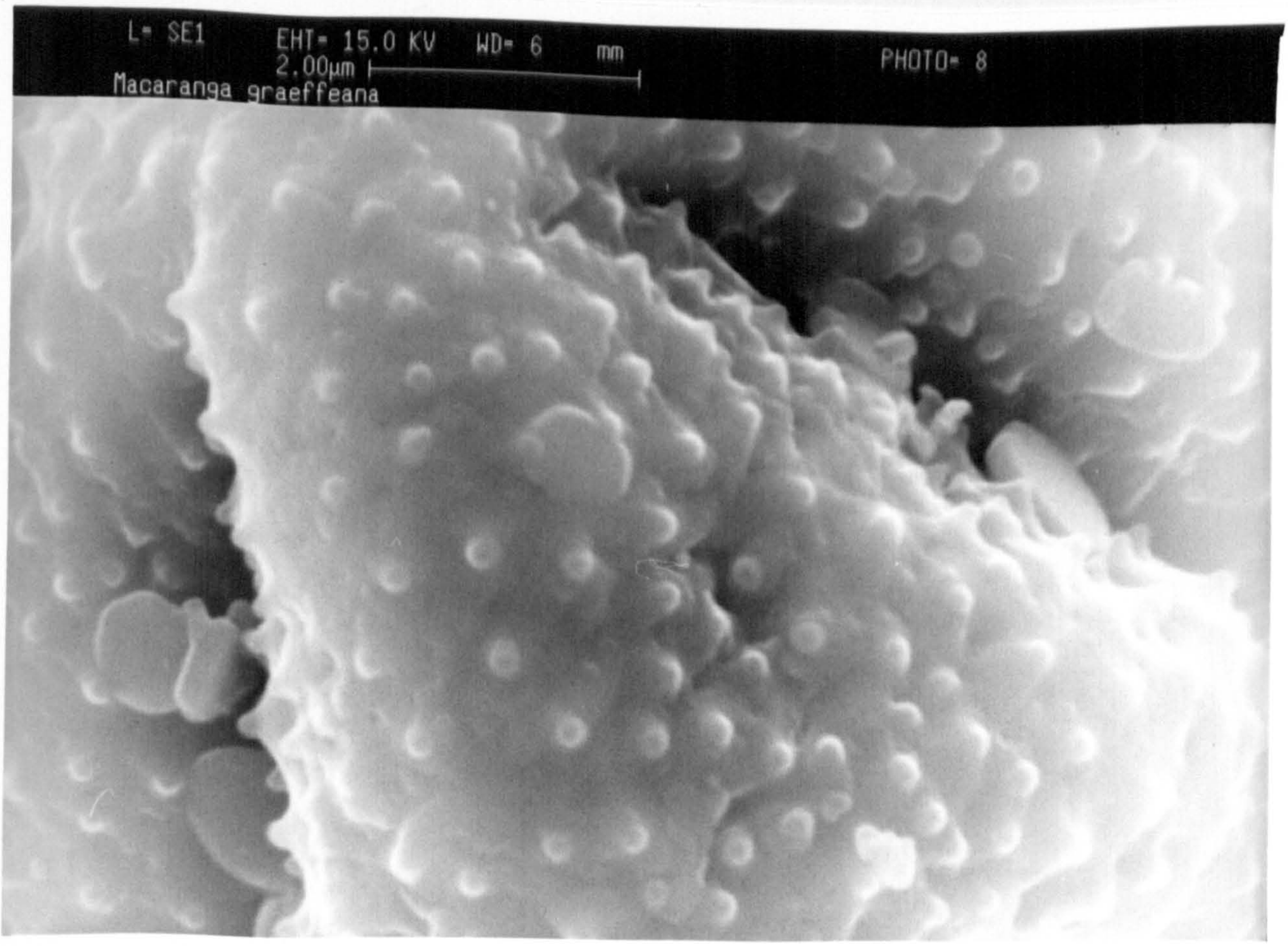


Plate 6(b) : Macaranga graeffeana (Texture)

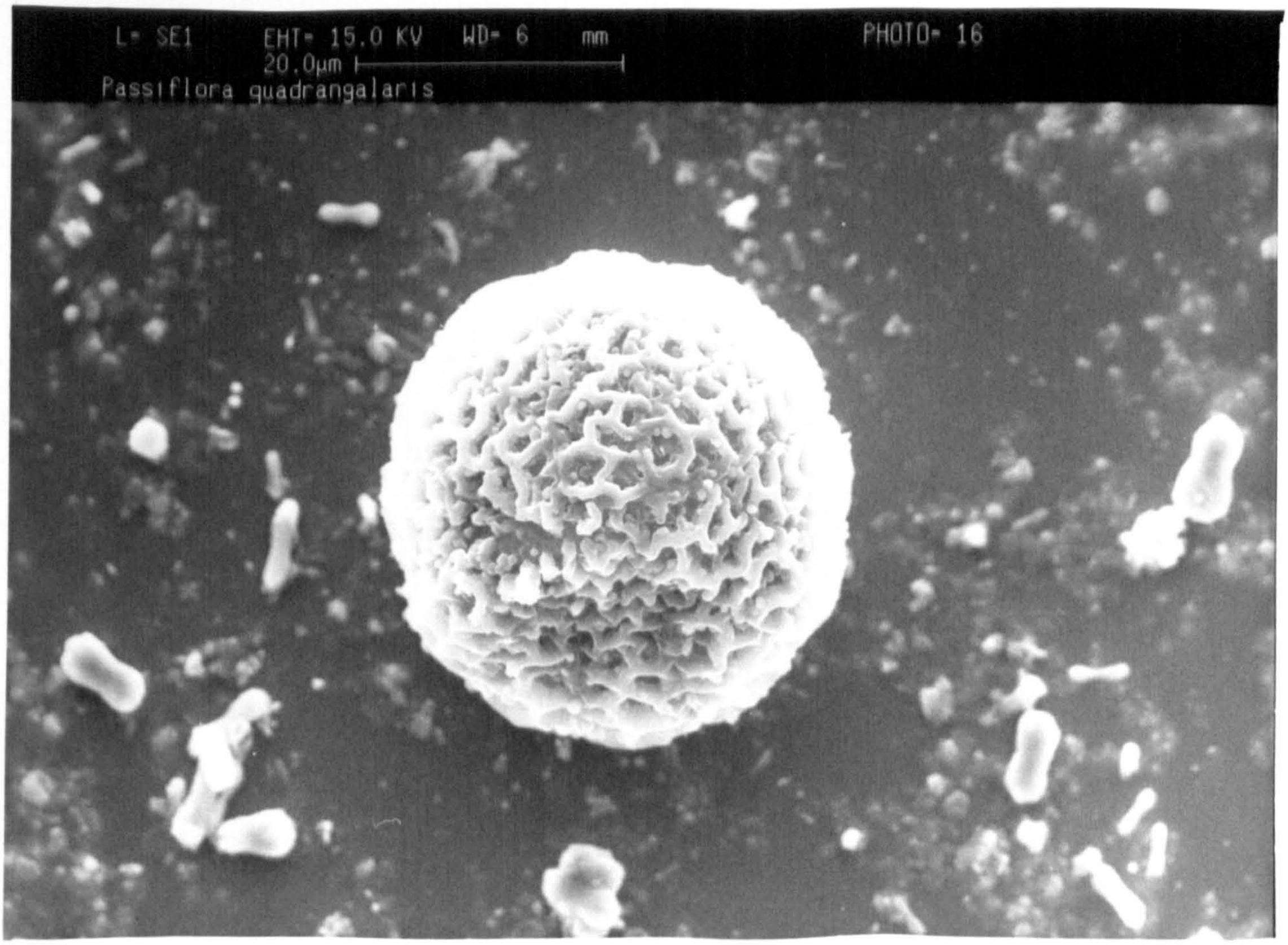


Plate 7(a) : Passiflora quadrangalaris

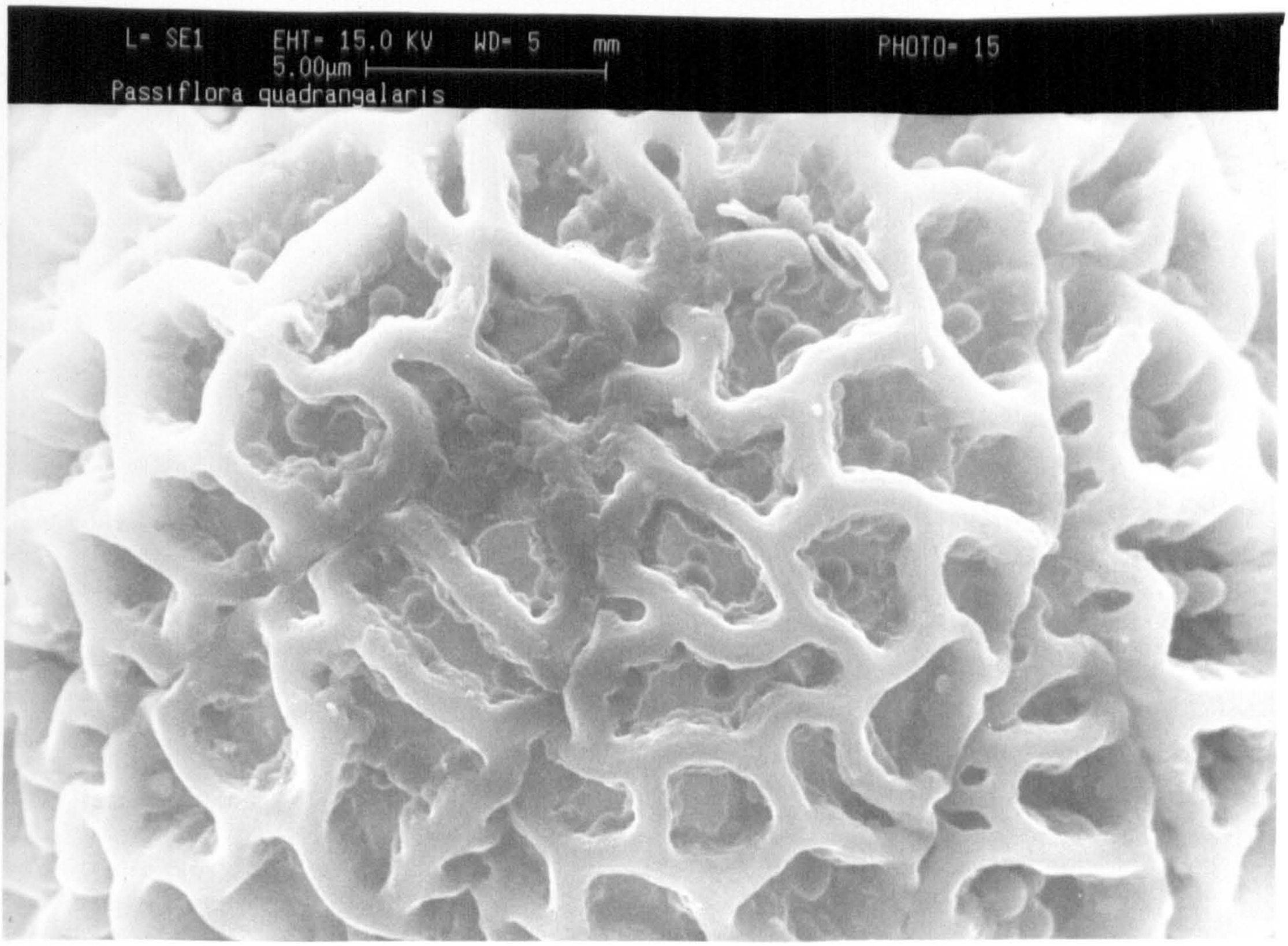


Plate 7(b) : Passiflora quadrangalaris (Texture)

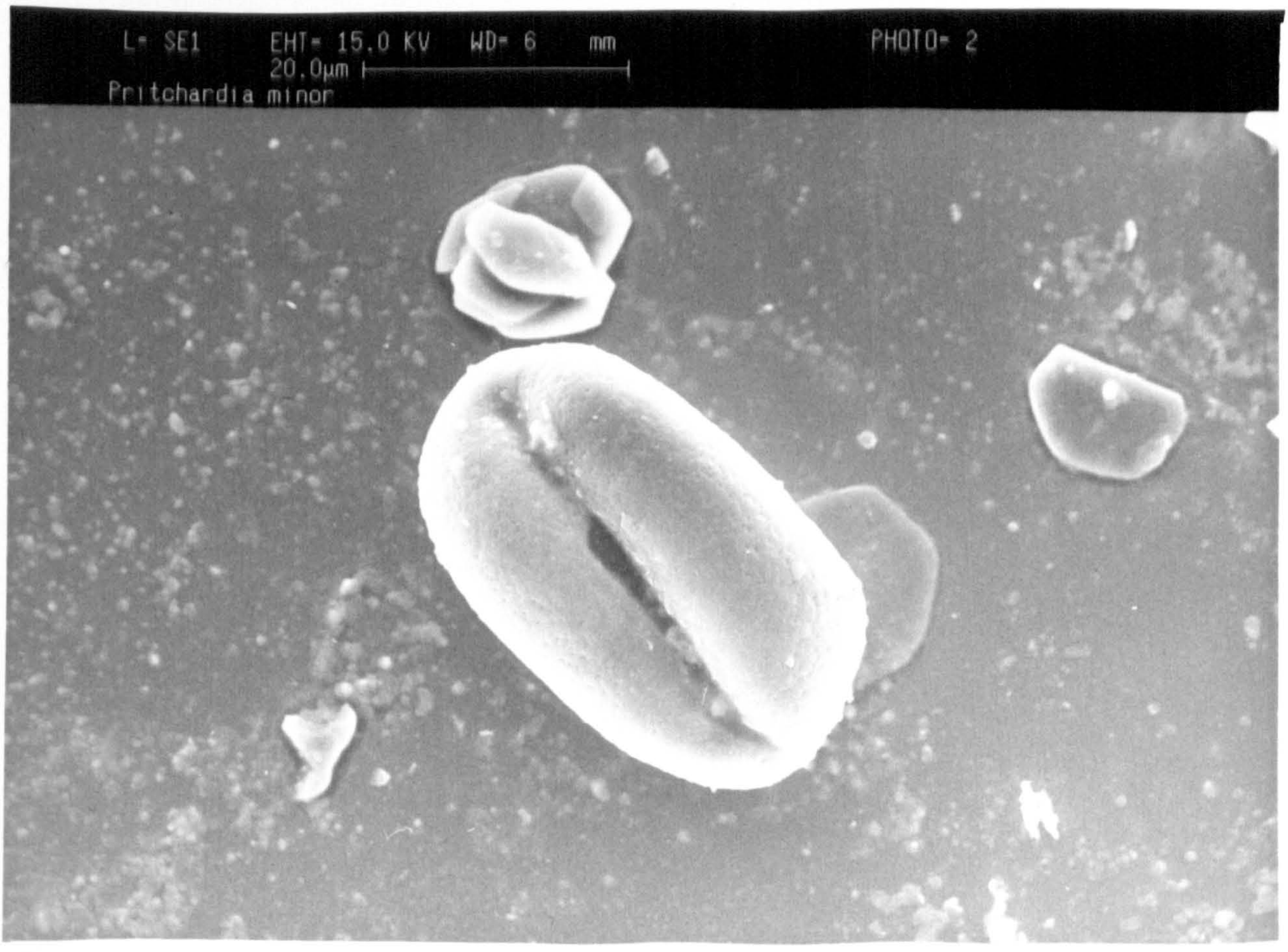


Plate 8(a) : Pritchardia minor

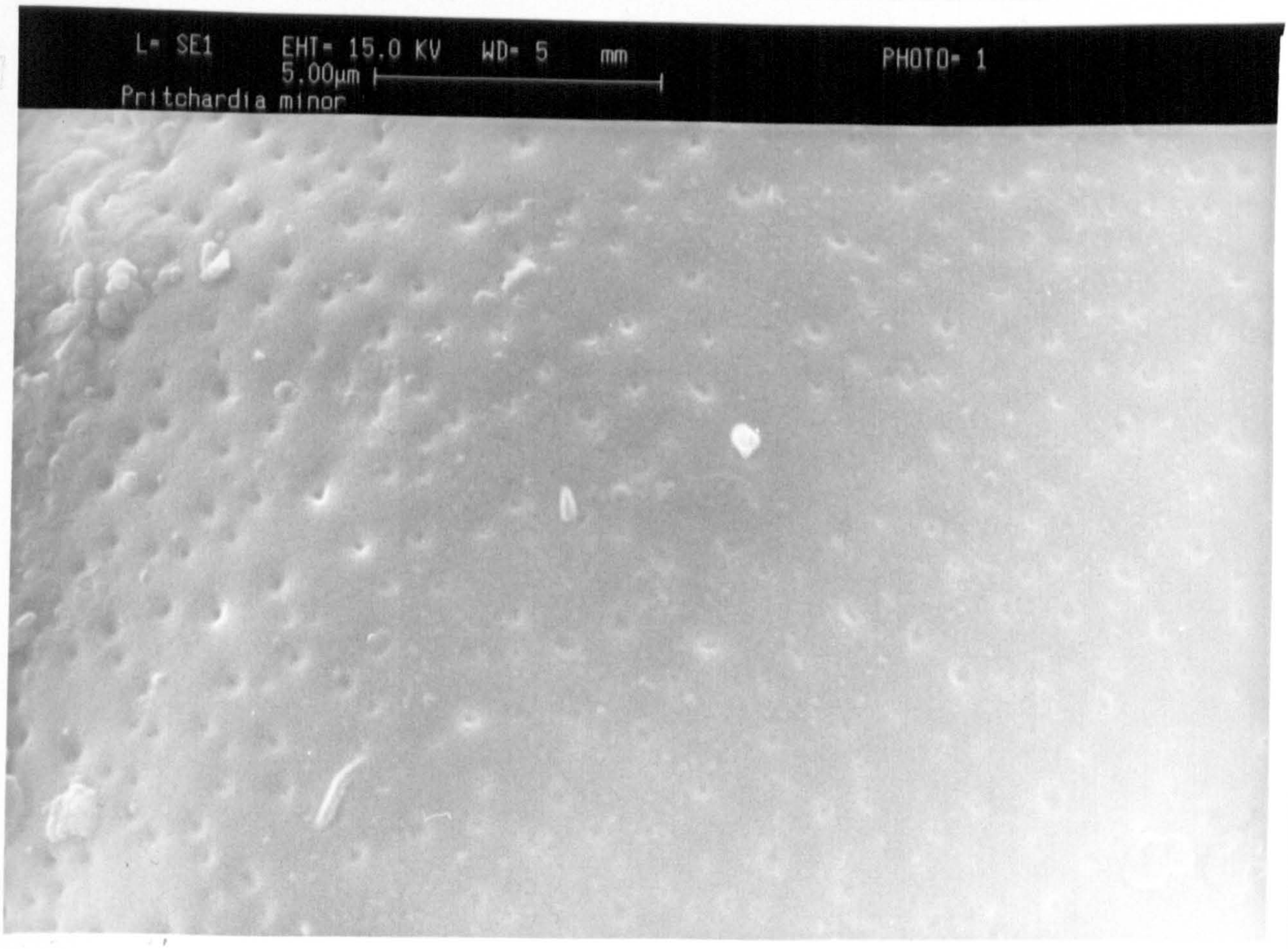


Plate 8(b) : Pritchardia minor (Texture)

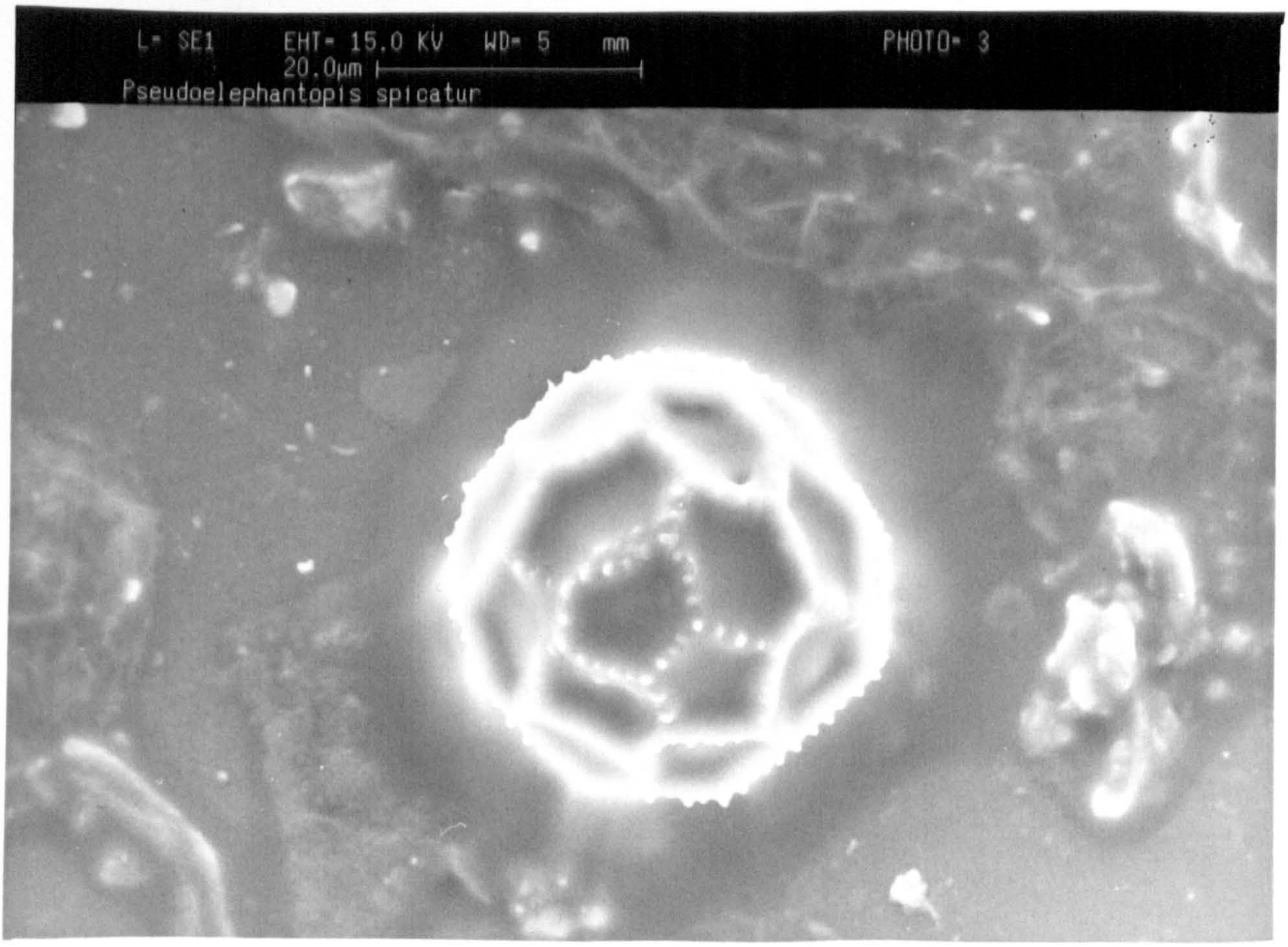


Plate 9(a) : Pseudoelephantopis spicatur

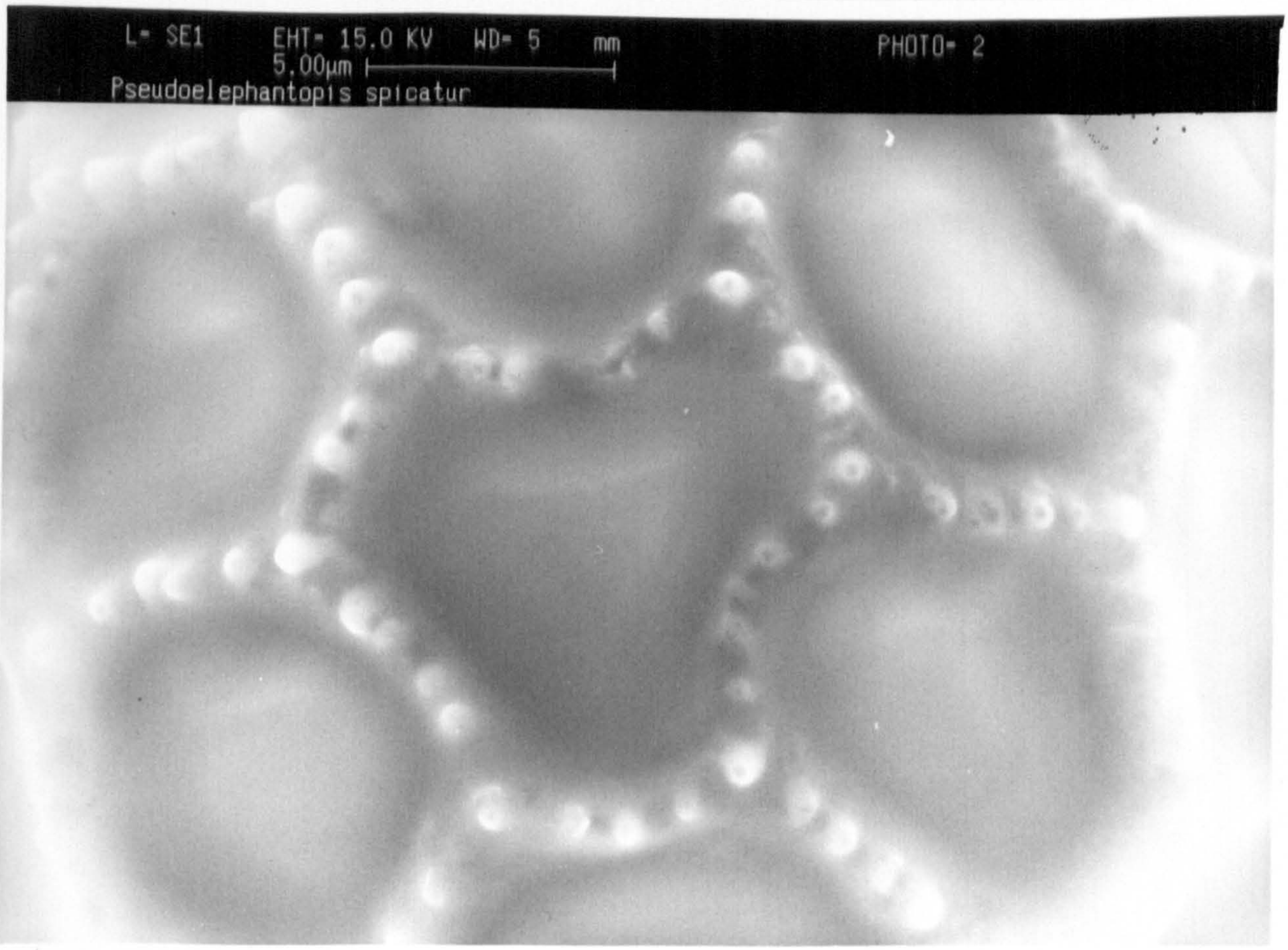


Plate 9(b) : Pseudoelephantopis spicatur (Texture)

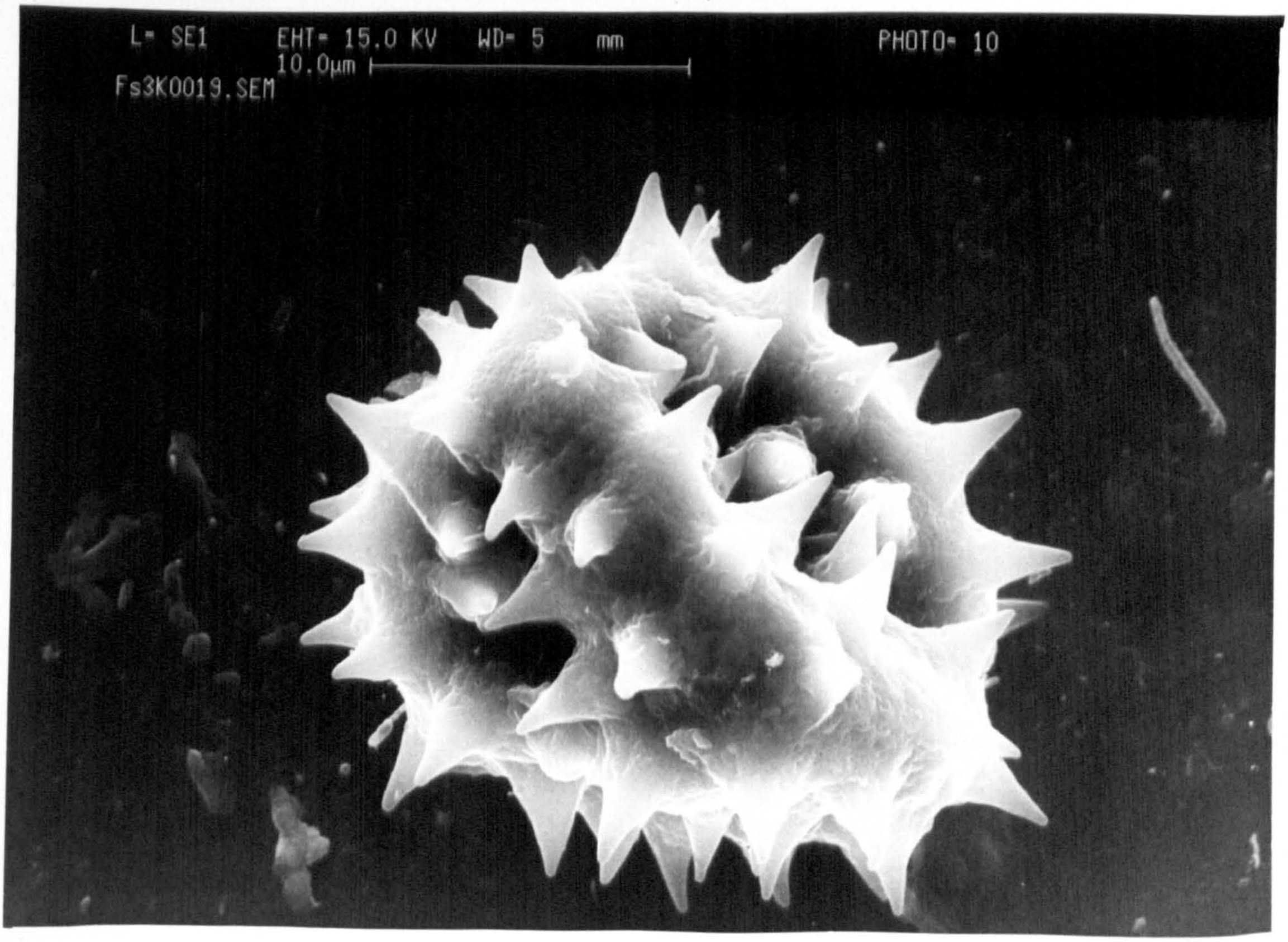


Plate 10 : Senecio stokesii

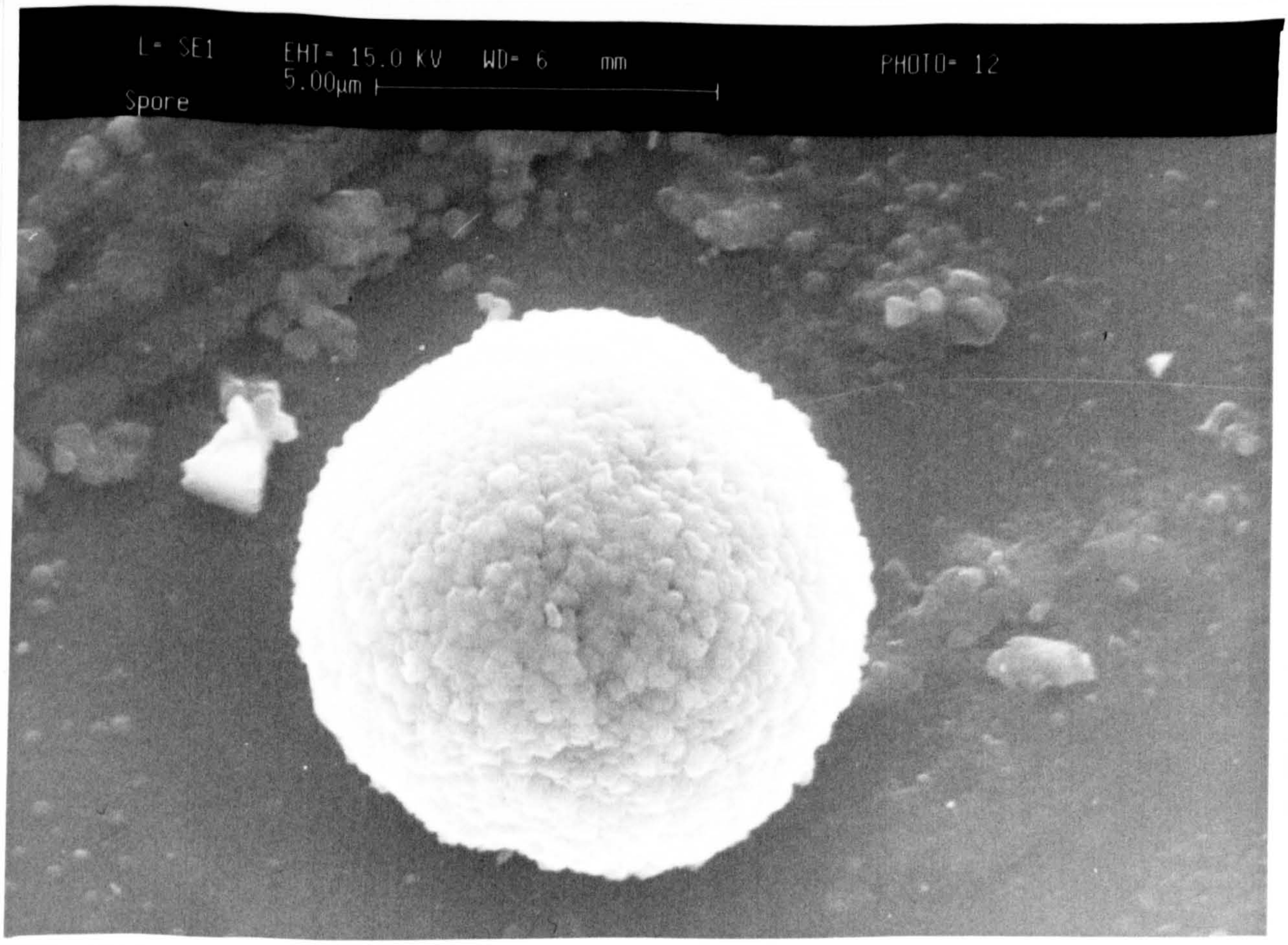


Plate 11(a) : Spore, Fungal

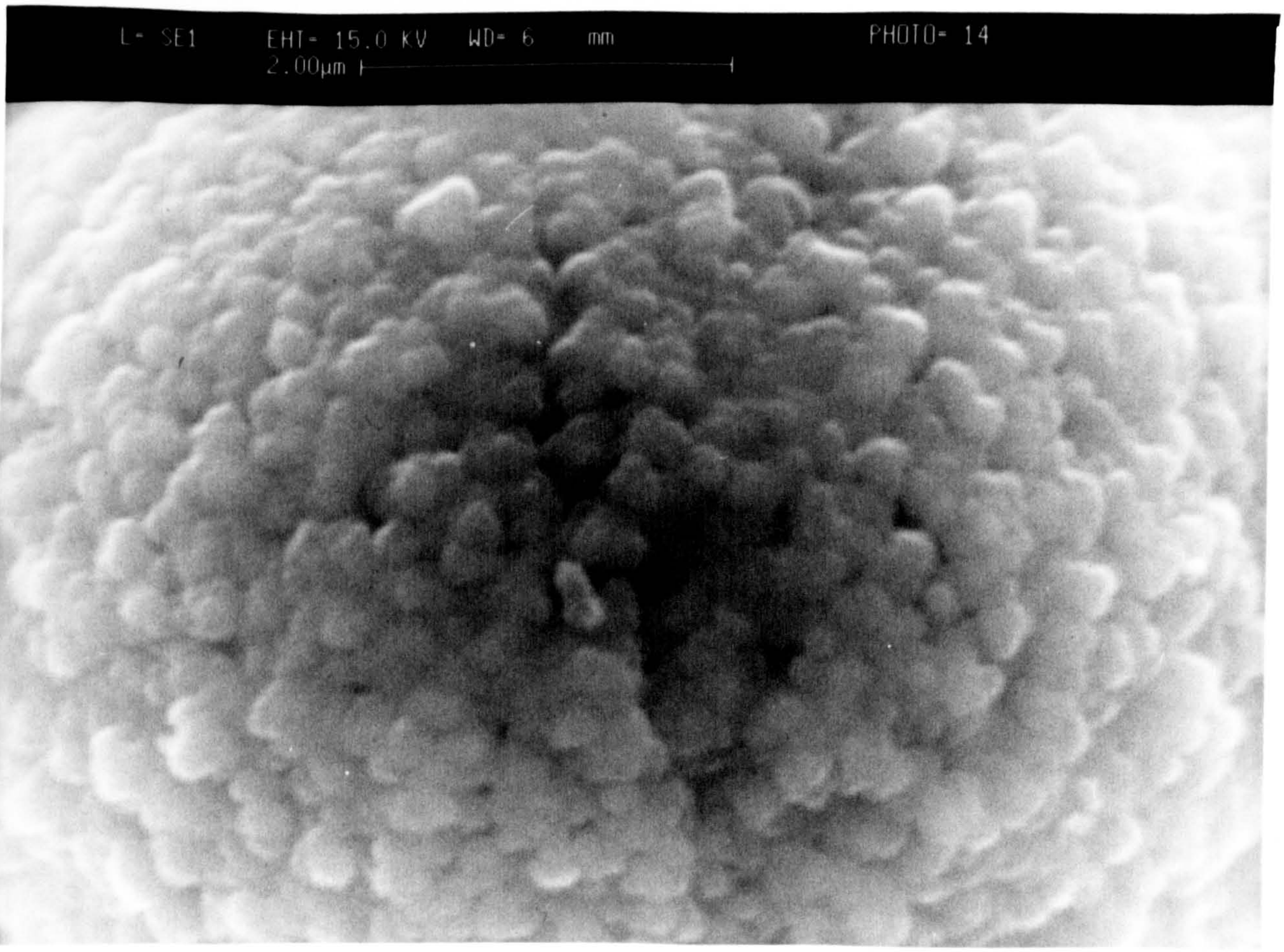


Plate 11(b) : Spore, Fungal (Texture)

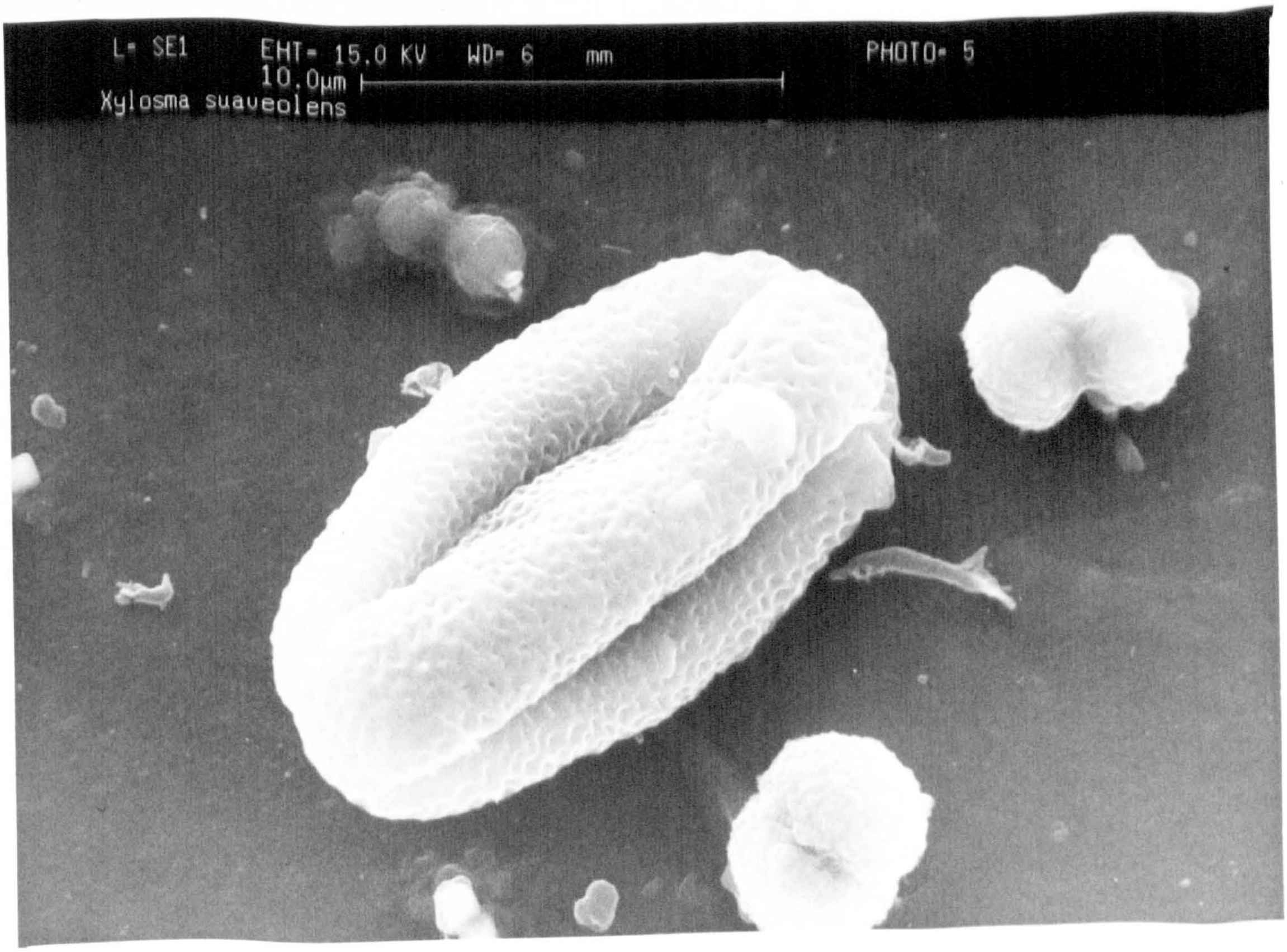


Plate 12(a) : Xylosma suaveolens

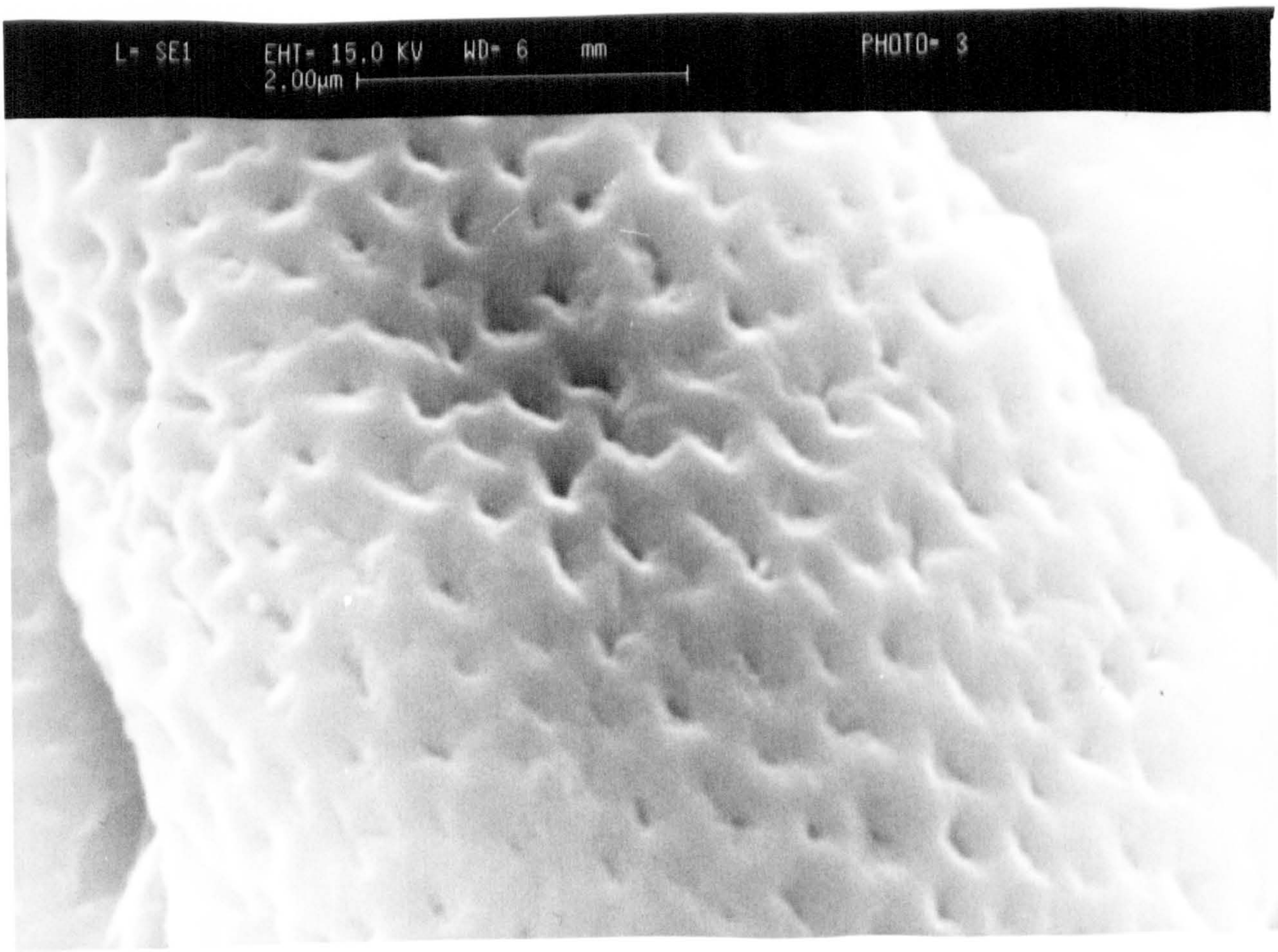


Plate 12(b) : Xylosma suaveolens (Texture)