Application of Surface Fitting Techniques for the Representation of Leaf Surfaces

B.I. Loch\textsuperscript{1}, J.A. Belward\textsuperscript{2} and J.S. Hanan\textsuperscript{3}

\textsuperscript{1}Department of Mathematics and Computing, University of Southern Queensland, Toowoomba, QLD 4350, Australia. E-Mail: lochb@usq.edu.au
\textsuperscript{2}Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia
\textsuperscript{3}ARC Centre of Excellence for Integrative Legume Research, ARC Centre for Complex Systems, University of Queensland, Brisbane, QLD 4072, Australia

Keywords: leaf surface representation, surface fitting, virtual plants, interpolation

ABSTRACT

Leaves play a vital role in the development of a plant, as they are major resource collectors. Adequate representations of leaves are therefore required for the modelling of plants. Such representations may be important to generate a realistic visualisation, or they may be used to study biological processes such as photosynthesis and canopy light environment. Accurate leaf surface representations are rarely found in the plant modelling community. This paper aims to show how detailed, accurate representations of leaf surfaces can be created from data; representations that may then be used as parts of virtual plants for applications in fields as diverse as the arts, agriculture or computer games. The techniques used here are mathematical methods of surface fitting applied to data that has been sampled from real leaves with a laser scanner (Polhemus FastSCAN 3D). These methods are interpolating finite element techniques, one using linear triangular elements, the other piecewise cubic triangles. The size of a laser scanned data set can be enormous and it may be important to represent the surface with significantly fewer points. An incremental algorithm is therefore used to identify significant points that result in a surface fit that approximates the entire data set to a pre-specified accuracy.

The algorithms are applied to two examples, a Frangipani leaf and a Flame Tree leaf. Figure 1 visualises results for the Flame Tree leaf. The images represent (a) a photo of this particular leaf, (b) the complete set of more than 5000 digitised data points, (c) positions of data points after application of the incremental algorithm for the piecewise cubic approach with an accuracy of 1%, (d) the same rotated to show the shape of the surface represented by these points, (e) the resulting triangulation and (f) the surface fit. From these point locations, guidelines are deduced describing where data points should be positioned, for example for measurement by lower resolution devices such as a sonic digitiser. The reduced point sets contain about $\frac{1}{10}$ of the number of points in the original data set.

The research presented in this paper is the first to model detailed and accurate leaf surfaces based on three-dimensional data points captured from real leaf surfaces. It provided a basis on which further research can be built. For example, detailed and accurate surface representations may be used in the simulation of pesticide deposition on leaf surfaces to determine the effectiveness of a treatment and help develop improved pesticide application techniques.

(a) \hspace{1cm} (b) \hspace{1cm} (c) \hspace{1cm} (d) \hspace{1cm} (e) \hspace{1cm} (f)

Figure 1. Representation of the Flame tree leaf surface. (a) A photo of the leaf, (b) the 5607 digitised data points, (c) the reduced point set, (d) the same set rotated, (e) the resulting triangulation and (f) the surface fit.
1 INTRODUCTION

Computer models of plants have been developed to study and to teach about the structure, function and development of plants as well as their interaction with the environment. The architecture of a plant plays a fundamental role in the acquisition and allocation of resources, tolerance to damage and competition [Bio85], hence many modelling approaches concentrate on plant structure and geometry. Virtual plants in particular are developmental plant models that incorporate topological and geometrical information, which can be used to produce a visualisation [RHP96, Pru98, God00]. They are tools for plant scientists and teachers in biology, agronomy, ecology and pest management [RHP96, Pru98].

Leaf models play an important role as part of a plant model. In the management of crop pests, for instance, the optimal timing of sprays to minimise the use of pesticides can be determined by simulation of the interaction of pesticide application to the surfaces of individual leaves and response in insect behaviour.

The leaf representations described in this paper are descriptions of the external architecture of a leaf. They may be used to generate a visually correct, realistic model that captures the shape for a visualisation.

Previous approaches to represent leaf surfaces include anti-aliasing disks by Smith [Smi84] and polygons and texture maps by Bloomenthal [Bio85]; bicubic patches (Bézier patches) were included into L-systems terminology as predefined surface objects, to which texture maps could be applied; Bézier patches were also implemented in other plant modelling environments such as AMAP [GGCC97] to represent leaf surfaces. Branching skeletons were used by Hammel et al. [HPW92] for compound leaves, and flat-bed scanner data and a boundary following algorithm were applied by Mündermann et al. [MMPP03]. Finally, Lintermann and Deussen [LD98, LD99] based their approach on splines and texture maps.

Note that none of the leaf modelling approaches mentioned above were based on detailed three-dimensional real world data; in most cases models were designed to fit visual observations. If data points were measured, then in general they were used to determine the position, orientation and size of a leaf, not to define its surface shape. Although many of the approaches mentioned above have resulted in visually pleasing leaf surface representations, no quantitative assessment has been made to compare hand-designed models to real leaves. In other words, accurate representations of leaf surfaces are rarely found. The object of this study is to close this gap by applying methods of surface fitting to sets of scattered data points sampled from leaf surfaces. Rather than modelling at the cell level, emphasis is placed on capturing visible surface details.

This paper is structured in the following way. First, digitising techniques are briefly presented and the digitised leaf species are described. Then two finite element surface fitting methods are explained, along with an incremental algorithm to reduce the size of the set of data points that needs to be collected. Results are presented and analysed in the form of data set sizes, point locations and visualisations of the fits. Future directions and applications are discussed in the conclusions.

2 TECHNIQUES

2.1 Digitising

2.1.1 Digitising techniques

Digitising is the process of sampling spatial coordinates of points from an object using a measuring device. Two digitising techniques suitable for measuring leaf surfaces are a laser scanner (e.g. Polhemus FastSCAN 3D) and a sonic digitiser (e.g. Freepoint 3D). The first is capable of returning very large data sets from small leaf surfaces, and can be classified as a multiple-point digitiser. The sonic digitiser on the other hand is a single-point device; only one data point can be digitised at a time. This means that the locations of data points on the leaf surface are very important, as usually only a small number of points are collected, which will have to suffice in describing the surface. More information on the various digitising techniques as well as issues arising from specific leaf properties can be found in [Loc04, HLM04].

A laser scanner was used to digitise leaf surface data for this study. Since a plant scientist may be more likely to have access to a sonic digitiser, information about the locations of significant points on a leaf surface are extracted from the laser scanner data in a later section. These guidelines can be used to help decide where a smaller set of data points should be measured, while still maintaining reasonable accuracy.
2.1.2 Digitised leaf types

Two different species are used as examples for which surface representations are generated. Leaves from a Frangipani and a Flame Tree have been digitised. These species are chosen because one has a simple shape while the other is palmately lobed. Results for two further leaf types can be found in [Loc04].

The Frangipani (*Plumeria japonica*) is a small deciduous tree. Its leaves have a simple, ovate (oval-shaped) boundary and possess a distinctive mid-vein. They taper to a point at the base where the leaf is attached to the petiole (the leaf stalk), and have a broader tip on the opposite end (Figure 2(a)). The Flame Tree (*Brachychiton acerifolius*) can grow up to 30 m tall. The blades of its leaves are simple ovate or palmately lobed, depending on the age of the tree (Figure 3(a) shows a leaf with a re-entrant boundary). Mid-veins are dominant in each lobe. Both plants grow in South-east Queensland.

![Frangipani leaf](image)

**Figure 2.** Frangipani leaf. (a) A photo of the measured leaf and (b) all 10473 data points.

![Flame tree leaf](image)

**Figure 3.** Flame tree leaf. (a) A photo of the measured leaf and (b) all 5706 data points.

These two photos show the actual leaves that have been digitised for this study. Both surfaces were sprayed with a Kaolin/water mix to change reflective properties so the laser scanner can capture surface information. For the Frangipani leaf, more than 10,000 points were collected, compared to about half this number for the Flame Tree leaf.

The two digitised data sets used in this study are displayed in Figures 2(b) and 3(b), and rotated to show the shapes of the surfaces in Figures 8(f) and 9(f) respectively.

2.2 Surface fitting

The mathematical methods of surface fitting applied to the scattered leaf surface data are interpolating finite element methods [Lv86]. Finite element methods are based on the concept that the domain on which data points are provided is divided into subdomains (patches) on which a function is defined. For scattered data points, triangular elements are commonly used; this will also be the case here. For each triangular segment, a surface function (usually a polynomial) is then determined that interpolates the vertices and, if required, derivative information. This function is only defined for that particular element. Derivatives may need to be estimated if they are not supplied with the data. The derivatives required depend on the finite element used. By joining many polynomial patches (in a possibly smooth way), the complete surface is generated. See Lancaster and Šalkauskas [Lv86] for a detailed introduction into this area.

Two different triangular finite elements are applied in this paper, the linear triangle (LIN) and the piecewise cubic Clough-Tocher triangle (CUB). The linear triangle [Lv86] requires values at the three vertices of the triangle to determine the three coefficients of the bivariate linear polynomial uniquely on each triangle. For the interpolation approach used here, data points are the vertices. This results in a planar surface on each triangle and in a planar faceted surface overall. The surface over the whole domain is therefore continuous, in other words there are no jumps (a C⁰ surface), and normal derivatives of the surface function are in general discontinuous at inter-element edges. This discontinuity can be regarded as a serious disadvantage of the piecewise linear method. If the original surface that provided the data points was smooth, then the interpolant should preferably be smooth. A smooth interpolant delivers a visually more pleasing surface and the visual aspect of an interpolant may be very important [Fra82]. However, piecewise linear polynomials can deliver a satisfactory fit if sufficiently many data points are interpolated, and if smoothness is not the first priority. The piecewise linear approach is widely used because of its simplicity: no derivatives need to be estimated and only three nodes are required for each triangular element.

The piecewise cubic Clough-Tocher triangle [CT65], on the other hand, is an element resulting in a smooth (C¹) surface over the whole domain. It is a seamed element approach, treating a triangle as a macro-element by subdividing it into subtriangles, the micro-elements. An interpolating cubic polynomial is
found on each subtriangle, and the bivariate piecewise cubic interpolant over the entire triangle can be constructed as continuously differentiable from these three polynomials. Only 12 degrees of freedom are required: the function values and gradient at each vertex, as well as normal derivatives at the mid-point of each side (see [Lv86] for further details). The Clough-Tocher approach has the advantage that it results in a smooth surface over the whole domain.

As is the case for leaf surface data, derivative information is usually not provided with the data and needs to be estimated. The normal derivatives at the side mid-points are often estimated as the mean of the normal derivatives at the vertices at each end of that side. This is based on the assumption that the normal slope along the sides of the triangle changes linearly. All quadratic but only some cubic polynomials can be reproduced when the interpolation scheme is reduced to a nine node scheme in this way [Lv86]. In this context, directional derivatives at the vertices are estimated following an approach introduced in [Bre01], which is also mentioned in Appendix B in [Loc04].

In order to use these techniques of surface fitting a reference plane is needed. This may be obtained by making a least squares fit of a plane to the data points. When this is done we refer to the distance \( z \) of the data points from the reference plane as function \( z(x, y) \) of the position \((x, y)\) of the data points in the reference plane, thus \( z = f(x, y) \). An interpolation method then enables construction of an approximant \( \Phi(x, y) \) such that \( z_i = \Phi(x_i, y_i) \) at each data point.

### 2.3 Point set reduction

The finite element surface fitting methods described above interpolate all available data points. Since it is possible to measure tens of thousands of data points situated on the surface of a small leaf, data sets acquired with a laser scanner can be enormous in size. Fitting an interpolating surface to a very large set of data points results in an elaborate data structure, so a more compact representation of the surface may be desirable. Aside from the computational aspect, a representation based on all available data points may not be required if emphasis lies on just a realistic visualisation of the surface, or if a less-detailed representation which includes major surface features is sufficient. Ideally, the number of interpolated points is reduced to a minimum. It is usually not practical to determine a minimum set due to the computational complexity of this problem. The aim is instead to determine sets of significant points for particular leaf surface types. Note that a representation that is based on fewer data points will approximate the original data, and may be a less accurate representation of the real surface. A balance needs to be found between the computational effort, storage and memory requirements and the resulting surface quality.

A reduction of the size of the data structure can be achieved using an incremental method. An initial surface is fitted to a small subset of the set of data points. Then, step by step, data points selected from the total set of data points are added to improve the surface fit. Such a point may be inserted into an existing triangulation in a computationally efficient way by identifying the triangle to which it belongs, and then locally retriangulating the region to restore the Delaunay property [Law77]. This adaptive method can be applied until the discrepancy between each data point and its approximation on the fitted surface has fallen below a threshold, in other words until the fit has reached a certain closeness to the data. See [FFP85, PK95] for previous results with this approach.

To quantify a fit based on a subset of the set of all data points, the accuracy of an approximation is measured in terms of a maximum error associated with the particular surface fit in relation to the maximum variation pointwise in \( z \). The maximum error is measured over all remaining points in \( P \), in other words all points that are not interpolated. It gives an indication of how well the fit behaves at its worst point. More formally, let \( f \) be a surface function interpolating the subset of data points \( P_f \subset P \), with \( P \) all available data points. The maximum error associated with \( f \) will then be defined as

\[
e(f) = \max_{p \in P \setminus P_f} e(p),
\]

where \( e(p) \) denotes the error associated with a data point \( p = (x, y, z) \) that is not a vertex, and

\[
e(p) = |z - f(x, y)|,
\]

the vertical distance between point \( p \) and its approximation on the surface. Note that these measurements do not give an exact indication of how good the fit is compared to the original surface from which data has been collected, since the data points are sampled.

For a surface with a specific boundary like a leaf surface, it is reasonable to begin an incremental point addition method with data points situated on the boundary of the surface. While there are many sensible approaches to choose the data point to be added, one that aims to reduce the maximum error includes the data point that carries the largest error. This approach was used for piecewise linear fits in [FFP85], while in [PK95], \( k \geq 1 \) points were added simultaneously for piecewise cubic Clough-Tocher surfaces. In experiments with leaf data sets, the maximum error criterion with parameters \( k > 1 \).
generally led to larger point selections than \( k = 1 \). For this reason, the case \( k > 1 \) is not considered further.

An interesting observation made in the two aforementioned papers as well as here is that the method adding the maximum error point does not monotonically reduce the resulting new maximum error. It is in fact possible that the addition of the maximum error point may lead to an increase in the maximum error for the fit. This situation often occurs when important features such as a midrib with a strong curvature are represented by too few points, and also when points situated on the midvein are not connected by triangle edges that are oriented along the vein. It was observed that the maximum error point is very often located close to the centres of triangle edges.

For the application of this incremental method to leaf surface data, the error tolerance limit is the accuracy, or in other terms the percentage of the maximum variation pointwise in \( z \) over all data points. After the accuracy has reached a certain percentage of the total range of surface elevation the algorithm is stopped.

### 2.4 Results

#### 2.4.1 Point set sizes

Techniques described in the previous section are applied to data sets of the two species. Boundary points that are chosen as initial configuration for the incremental algorithm are marked with circles in Figures 2 and 3. These points were selected to represent the boundary in an adequate way, that is to capture any specific features without leaving large gaps. Two tolerance limits are applied: the algorithm is stopped when the surface fit has reached 5% or 1% accuracy compared to the total set of available data points. Both facetted (LIN) and smooth (CUB) surfaces are generated.

The resulting interpolation data set sizes are listed in Table 1. Outcomes are similar for both surface fitting methods, although point positions generally vary.

**Table 1.** Resulting sizes of point sets for the reduction approach, for both leaf types.

<table>
<thead>
<tr>
<th></th>
<th>Frangipani</th>
<th>Flame tree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LIN</td>
<td>CUB</td>
</tr>
<tr>
<td>total points</td>
<td>10473</td>
<td>10473</td>
</tr>
<tr>
<td>initial points</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>for accuracy 5%</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>for accuracy 1%</td>
<td>323</td>
<td>327</td>
</tr>
</tbody>
</table>

For the frangipani leaf, the original data set of more than 10,000 points can be reduced by several orders of magnitude if an accuracy of 5% is sufficiently precise. An accuracy of 1% requires only \( \frac{1}{30} \) of the original size.

For the flame tree leaf the reduction is less, though still significant. More data points need to be interpolated for this species to capture the surface shape of each of the lobes accurately. This is also the reason why the size of the initial (boundary) data set is larger.

Numerous data sets were acquired for each leaf, varying greatly in size. Application of the incremental algorithm to each of these leads to similar results in the order of magnitude of resulting point sets. Table 2 displays the point set sizes for three different data sets of the same frangipani leaf.

**Table 2.** Comparison of results for three different scans of the same frangipani leaf.

<table>
<thead>
<tr>
<th></th>
<th>LIN</th>
<th>CUB</th>
<th>LIN</th>
<th>CUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>data set 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>10473</td>
<td>10473</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for accuracy 5%</td>
<td>55</td>
<td>62</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>for accuracy 1%</td>
<td>323</td>
<td>327</td>
<td>327</td>
<td>327</td>
</tr>
<tr>
<td>data set 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>9208</td>
<td>9208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for accuracy 5%</td>
<td>71</td>
<td>87</td>
<td>71</td>
<td>87</td>
</tr>
<tr>
<td>for accuracy 1%</td>
<td>405</td>
<td>431</td>
<td>405</td>
<td>431</td>
</tr>
<tr>
<td>data set 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>3402</td>
<td>3402</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for accuracy 5%</td>
<td>61</td>
<td>58</td>
<td>61</td>
<td>58</td>
</tr>
<tr>
<td>for accuracy 1%</td>
<td>390</td>
<td>367</td>
<td>390</td>
<td>367</td>
</tr>
</tbody>
</table>

#### 2.4.2 Guidelines for the location of data points

Apart from generating representations of leaf surfaces from available data, the aim of this study is to decide where data points need to be positioned to capture important surface features. This information can then be used to give advice to plant scientists with single-point digitising devices such as sonic digitisers, to help decide where surface data needs to be measured. Locations of data points for the two species are displayed in Figures 4-7; these correspond to results given in Table 1. Comparing these sets for each leaf type leads to the following observations about the structure of the data sets.

For a leaf with an undulate boundary such as the Frangipani leaf (Figure 5), the surface needs to
Interpolate a sufficiently large number of data points along the boundary. Figures 4(left) and 5(left) show that the initial set of 17 points has been complemented by many more points along the boundary for an accuracy of 1%. Furthermore, points along the mid-vein have been included. This can be observed even for the 5% sets in Figures 4(right) and 5(right).

Figure 4. Frangipani leaf: point sets. (a) The reduced set for LIN with 1% accuracy, (b) for LIN with 5% accuracy, (c) for CUB with 1% accuracy and (d) for CUB with 5% accuracy. Initial boundary points are marked in bold.

Figure 5. Frangipani leaf: point sets. Data sets from Figure 4 rotated to show the surface shape.

Figure 6. Flame tree leaf: point sets. (a) The reduced set for LIN with 1% accuracy, (b) for LIN with 5% accuracy, (c) for CUB with 1% accuracy and (d) for CUB with 5% accuracy. Initial boundary points are marked in bold.

Figure 7. Flame tree leaf: point sets. Point sets from Figure 6 rotated to show the surface shape.

Point locations for the lobed Flame Tree leaf are shown in Figures 6 and 7. The basal part of the leaf can be divided into regions that are dominated by the main veins leading into the lobes. At the interfaces between those regions, below the “re-entrant”, the surface forms a crest. Points are selected along all main veins and also along the interfaces leading towards the re-entrant points. For a 1% accuracy, this structure is more pronounced for LIN than for CUB, where the points are scattered along the veins.

An interesting observation for both species for LIN and 1% accuracy is that data points are located along what appear to be double lines following dominant veins (top left of Figures 4 and 6). Comparison of the 5% to the 1% selections of both leaves reveals that in fact a second line of points has not been added next to those on the original vein, but the new points are scattered around the existing ones along the veins. This results in a refined triangulation along the veins, and therefore in an improved representation of the width of the veins for LIN.

2.4.3 Visualisation - triangulations and surfaces

The triangulations corresponding to the reduced data sets are shown in Figures 8 and 9 (top four plots). Colour represents the average surface height of the three vertices of each triangle. By visual comparison with the original data set (bottom right), the 1% accuracy surfaces, and even the 5% accuracy fits are good representations.

A visualisation of the surfaces for LIN can be found in the two plots on the top left in the two figures. The surface fit for CUB with an accuracy of 1% is shown on the bottom left. Interpolated shading and a green scale are applied here.
3 CONCLUSION

Accurate leaf surface models are rarely found in the plant modelling community. The research presented in this paper describes an approach allowing the user to model detailed and accurate leaf surfaces based on three dimensional data points. Mathematical surface fitting techniques are applied in combination with an incremental algorithm to reduce the set of interpolated data points to a more compact size. Resulting surface fits are of satisfactory quality. After an analysis of the locations of selected data points, guidelines are given on where data points should be digitised when a single-point device is used.

The work described here delivers a basis on which further research can be built. This may be in applications of detailed models, or into extensions of the surface representation. Such representations may be used not only to generate realistic and accurate images of leaves, but also for applications that go further than visualisation. For example, in the simulation of pesticide application to plant surfaces [RCBF98, HRY03] it may be important not only to model splash, but also to find the path that a droplet of a pesticide sprayed onto the plant follows along a leaf surface, before it leaves the surface or comes to a standstill. Knowledge of this behaviour may be used to determine the effectiveness of a treatment, and eventually to help evaluate differing pesticide formulations. A realistic and detailed surface and boundary shape is a prerequisite for such a simulation. See Appendix A in [Loc04] for a prototype simulation.

Another application would be the simulation of light interception in canopies. España et al. [EBA+99] found that when ray tracing techniques are applied to compute the radiative transfer in a canopy of the maize plant, the level of detail with which the leaves are represented, expressed by the number of triangles used, does not influence the estimate. This remains to be confirmed for other leaf types. More detailed models may lead to more accurate results than simple prototypes like those used in [Lan73, STMK98].

The presented surface models contain structural but no functional information. It should be possible to use these descriptive surface models as a first step to generating a functional model. From the digitised data sets and the fitted surface, information about the location of veins could be extracted. This could be used, for instance, to visualise a model of the transport of sugars produced from carbohydrates that are generated during photosynthesis from leaves to other parts of the plant. Another application could be the simulation of insect-plant interaction, for instance an insect attack on a leaf, and the consequences to sections of the leaf surface when veins are severed.
An advantage of the leaf models described in this paper is that they are not restricted to a particular software package; they may be used in different plant modelling environments such as AMAP [GGCC97, AM], xfrog [LD98, LD99] or L-Studio [PKMH00].

4 REFERENCES


