University of Otago

Te Whare Wananga O Otago
Dunedin, New Zealand

Interactive Visualisation Tools for Analysing NIR Data

Hayden Munro
Kevin Novins
George L. Benwell
Alistair Mowat

The Information Science Discussion Paper Series

Interactive Visualisation Tools for Analysing NIR Data

Hayden Munro *†   Kevin Novins†   George Benwell*   Alistair Mowat‡

July 18, 1996

Abstract

This paper describes a tool being developed to allow users to visualise the ripening characteristics of fruit. These characteristics, such as sugar, acid and moisture content, can be measured using non-destructive Near Infrared Reflectance (NIR) analysis techniques. The four dimensional nature of the NIR data introduces some interesting visualisation problems. The display device only provides two dimensions, making it necessary to design two dimensional methods for representing the data. In order to help the user fully understand the dataset, a graphical display system is created with an interface that provides flexible visualisation tools.

Keywords: NIR spectroscopy, Polhemus FasTrak^{TM}, interaction, interactive graphics, interfaces, visualisation, scientific visualisation.

1   Introduction

Scientific visualisation can be described in two ways; as a tool for discovering and understanding; and as a tool for communicating and teaching [2]. It is used to present information to the user in visual forms that appeal to his/her intuitive understanding. Thus, visualisation tools facilitate the extraction of knowledge from often very complex datasets.

This paper describes a tool that is being developed to allow users to visualise data about the ripening characteristics of fruit. These characteristics, such as sugar, acid and moisture content can be measured using non-destructive Near Infrared Reflectance (NIR) analysis techniques [6, 8, 9]. By comparing the information gained by NIR analysis with the physical ripeness of the fruit it will be possible to determine which NIR readings correspond to a good product. Once this has been done NIR analysis will be able to be used alone to measure fruit quality. The information gained using NIR analysis could then be used to tailor the fruit growing and storage processes to maximize market acceptance. It may also be possible to measure ripening characteristics for prediction of fruit quality before the fruit has reached maturity.

The multidimensional nature of the NIR data introduces some interesting visualisation problems. The data is four dimensional, whereas the display device only provides two dimensions. It is therefore necessary to discover two dimensional methods for representing the data [7]. Also, the users of the application do not know what they are looking for in the data, so it is important to create an application with a high degree of interaction to assist their exploration. In order to tackle these problems this implementation provides a suite of interactive visualisation tools rather than a single display mode. In addition the user will be given the power to modify the display for his/her own purposes. This should provide the maximum flexibility of the computer to assist the user's exploration of the dataset.

The work described here was undertaken as a joint research project by Otago University and Hort+Research.

---

*Department of Information Science, University of Otago, P.O.Box 56, Dunedin, NZ. (gbenwell@commerce.otago.ac.nz)

†Department of Computer Science, University of Otago, P.O.Box 56, Dunedin, NZ. ([hayden|novins]@cs.otago.ac.nz)

‡Hort+Research, Ruakura Agricultural Centre, Ruakura Road, Private Bag 3123, Hamilton, NZ. (amowat@hort.cri.nz)

2 Description of Data Collection

The data for this project was provided by Hort+Research. It consists of reflected intensities of light at discrete wavelengths in the near infrared spectrum, at several locations on the subsurface of the fruit. The spatial coordinates of these locations were also provided. The data was collected from Gala apples and kiwifruit.

The spatial coordinates were obtained using a Polhemus FasTrak^{TM} device [13]. Each fruit was placed inside a cylinder with the longer axis perpendicular to the cylinder's straight edge, as shown in Figure 1. A longitudinal line of nine holes had been drilled at 15 degree intervals around the cylinder's circumference. The stylus of the tracker was pushed through each of the holes to measure a position on the fruit surface. These point locations correspond to one eighth of the fruit. The fruit was then rotated 45 degrees to collect the next set of points and so on, until the whole fruit surface had been sampled. In total 72 points, distributed across eight points of latitude and nine points of longitude, were selected for spectral analysis. Because the spatial points were collected at angular intervals, a polar coordinate transform was used to reconstruct the coordinates in 3D.

After the FasTrak^{TM} device had been removed from the guide holes, a Near Infrared Reflectance (NIR) probe was inserted into the guide holes to record the diffuse NIR spectrum at the fruit's subsurface at each of the 72 points. The diffuse NIR spectrum was sampled at 1100 wavelengths in the range 507.66nm to 1026.23nm. The sampled spectrum shows the amount of energy that is reflected at each different measured wavelength [5]. There is an error associated with each measurement because for each of the 1100 wavelengths sampled the photosensor collected light over a 2nm range, rather than at one discrete wavelength on the spectrum. A file was created containing the intensities of the 1100 sampled wavelengths for each point.
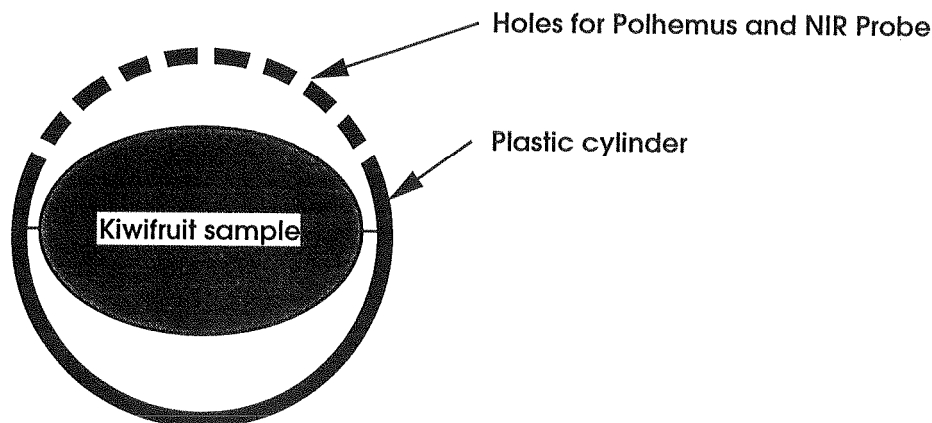
Figure 1: Cross section: selecting spatial points on a fruit.

3 Visualisation Methods

The data described in the previous section can be considered four dimensional: the points on the fruit's subsurface where wavelengths were measured must be described in 3D in order to maintain their spatial relationships; and the spectral wavelength information becomes a fourth dimension.

Since the data is four dimensional and the display is only two dimensional, no single visualisation method will be able to display all aspects of the dataset simultaneously. Also, the user cannot predetermine which view will be best. Fortunately the computer's flexibility allows a suite of visualisation tools presenting different conceptual *views* of the data to be supplied. By using the tools interactively and in concert the user can discover features in the data.

The tools currently provided are divided into two categories. The first is the *spectrum view* category. The spectrum view tool displays the intensity at all measured wavelengths for a single point on the fruit surface. The second category contains *spatial views*. Spatial views display 3D fruit geometry and intensity values for a single wavelength at all surface points. These views are described in detail in the following subsections.

The software environment is OpenGL[1] working with Tcl/Tk[2]. OpenGL is a standard graphics library designed for real-time applications [10]. It is designed to encourage the development of portable programs. The programs are written using the C language to call procedures from OpenGL. The OpenGL library makes it particularly easy to create and interact with the polygonal models used in the programs.

Tcl/Tk was also chosen for its flexibility. Tk provides a library of tools that are used for creating application interfaces; and Tcl is an interpreter that builds the program's interface from command scripts [11, 16]. Using scripts rather than library calls allows the interface to be modified while the application is running.

The software tools described here are available for many platforms; for the implementations described a Silicon Graphics Indigo was used.

3.1 Spectrum View

The *spectrum view* displays the intensity at all wavelengths for a single point on the fruit surface. An example is shown in Figure 2. The noise in the intensity data can be seen along the top edge of the spectrum intensities.
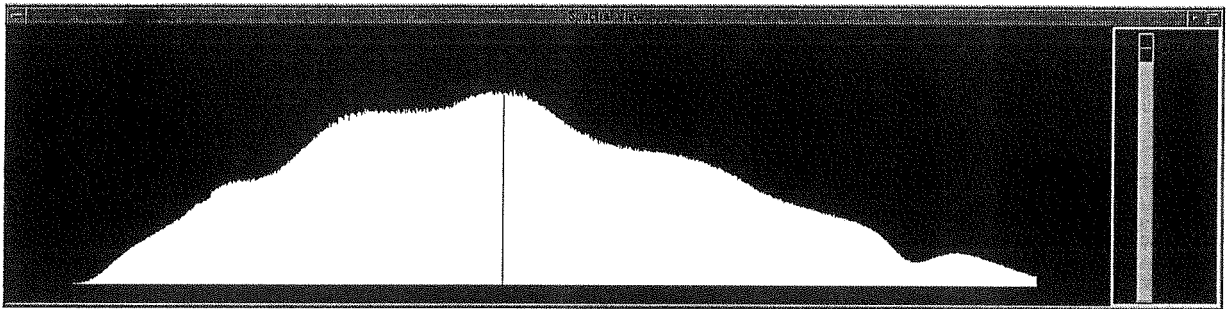

Figure 2: The *spectrum view*.

The spectrum view may be modified by what we call *smushing*. The smushed value for any given wavelength is an average of that wavelength's intensity and neighbouring intensity values. The primary reason for smushing is to reduce noise in the intensity data. Figure 3 illustrates the effect of applying the smush to average seven adjacent intensity values. The interface allows the user to adjust how many neighbouring intensity values are combined to create each averaged value.

3.2 The Spatial Views

The following three views have a common trait that sets them aside from the spectrum view: each spatial view shows multiple points on the fruit surface, but intensity at only one wavelength. This enables comparison of intensities between points, and provides an opportunity to observe their spatial relationships. It is necessary to introduce a slider to select the wavelength that is to be displayed.

3.2.1 3D Model View

The *3D model view* displays 3D fruit geometry coloured according to the intensity at a single wavelength. Using the spatial coordinate data a 3D model can be created by plotting the

---

[1] OpenGL is a trademark of Silicon Graphics, Inc.

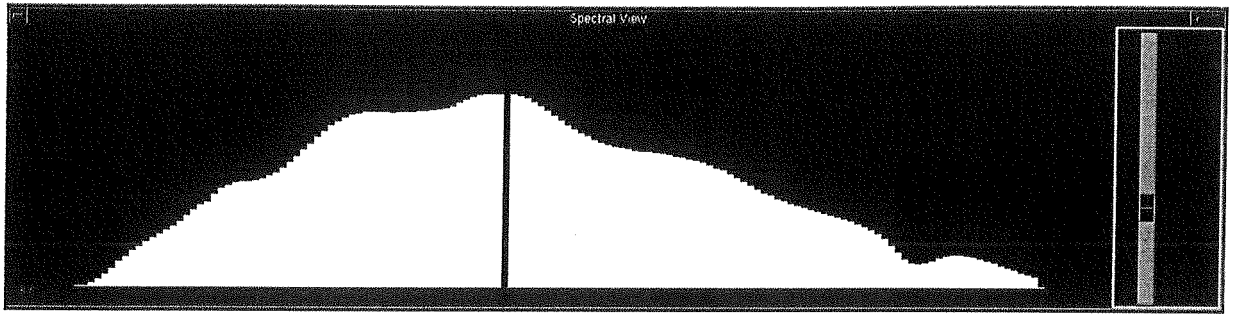[2] Tcl/Tk stands for Tool Command Language/Toolkit.

Figure 3: The effect of *smush*, averaging every seven intensities.

points in a virtual 3D space, as shown in Figure 4. A perspective projection transforms the 3D model into a 2D picture for display [3]. The user can interact with the scene using a virtual trackball [4]. The mouse manipulates the viewing position so the model can be seen from any angle, giving the impression of a three dimensional object.
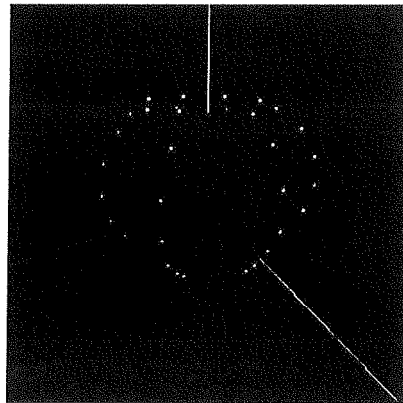

Figure 4: The 3D spatial points.

Plotting the points in 3D is not sufficient; it is necessary to find a way to show the spectral data. This can be achieved by displaying the intensities at a single wavelength as colour information. By mapping the largest intensity to white and the smallest to black, these and the intermediate values can be shown as levels of gray on the fruit model. A green to red colour scale is constructed using a similar method and provided as a menu option. A slider can be used to select the individual wavelength to display.

The points chosen for each fruit represent a rather sparse sample. In order for the fruit to be displayed as a "skin" the intensity information is interpolated between points. The method used to interpolate this colour information is known as Gouraud shading [3]. Gouraud shading is a popular technique because it is simple and can be computed using graphics hardware. It is useful to extend these points to form a skin for two reasons. Firstly the shape of the model is clearer if it is represented as a skin. And secondly the point colour is interpolated across an area, making it easier to see. The skin is made from polygonal surfaces constructed using adjacency information extracted from the latitudes and longitudes of each point.

Representing intensity with grayscale has a drawback. If the model is shown on a 24-bit monitor, there is a restriction to only 256 grey values. This means that it is only possible to distinguish between 256 intensity values in the dataset. By using the red, green, and blue channels independently it is possible to increase this number by a factor of six. Ware [15] suggested that the lack of colour resolution is a minor problem when compared with systematic errors that can arise from interpretive effects of the human visual system, such as simultaneous contrast. With this in mind it is more important to reduce these perceptual effects than
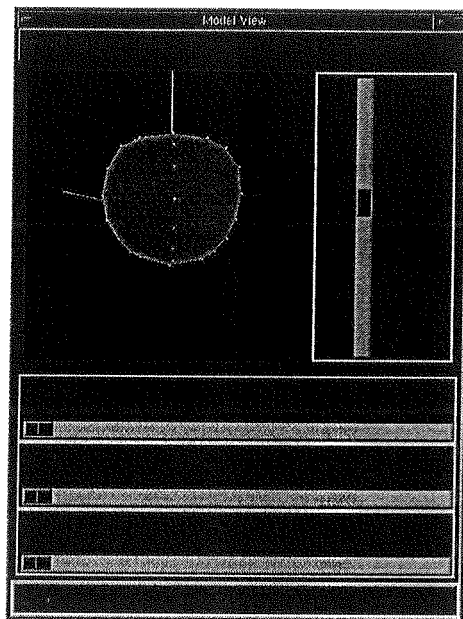
4

Figure 5: The *model view*.

it is to reduce quantisation of the displayed data. In accordance with this philosophy an approximation to the visual spectrum is made available as a colour map for the wavelength data.

3.2.2 The Map and Height Field Views

The 3D tool is useful for viewing the fruit in a virtual environment but it is restrictive since at any time the "back" part of the fruit is occluded, leaving the data only partially visible. It is possible to "unwrap" the model view so that the full surface of the fruit is shown in one picture. This is shown in the *map view*. An example map view is shown in Figure 6. The data collection method provides a natural latitude-longitude coordinate system for the data. In principle, this partitioning allows us to use any of the techniques developed by cartographers for mapping the earth. In order to "flatten" the fruit surface our current tool simply plots longitude as the x-axis and latitude as the y-axis in what is known as an equirectangular projection [14]. However all 2D maps of curved objects will contain distortions, and while accuracy of the equirectangular plot is high at the fruit's equator it decreases rapidly as the poles are approached.

The map view reduces visual complexity and allows us to use the 3D graphics hardware for another purpose. In the *height field view* intensities are interpreted as altitudes, and the data is displayed as a 3D relief map. This artificial terrain helps to clarify the relative distance between intensity values, which is difficult to determine when only colour information is used. The height field can be viewed from any angle by manipulating a virtual trackball. The heights also respond interactively to movement of the wavelength slider. It is important to incorporate same stationary reference for this view; otherwise the surface appears as if it is floating in space. The stationary reference is created by vertical lines below the height field in Figure 7.

Also a new slider is introduced for the height field view. It controls the height of the maximum value that is in the dataset, acting as a scaling constant for the height of the height field. This is provided so the user has the flexibility to exaggerate the disparity between similar intensity values that are displayed.

It is important to note that applying heights to the intensity values in the model view would not be beneficial. The 3D object is too complex to act as an adequate reference for the changing amplitudes. It would be difficult to distinguish between bumps that represent fruit
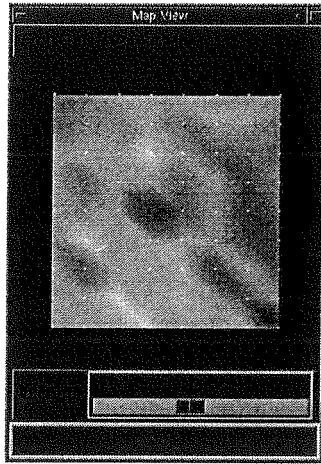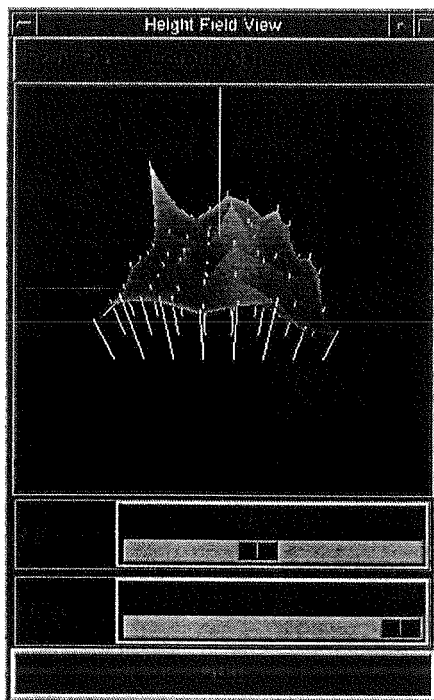
5

Figure 6: The *map view*.

Figure 7: The *height field view*.

geometry and bumps that represent intensity data.

4 View Integration

The spectrum and map views described in the previous section are combined to create a single application. These views present different viewing concepts so they occupy separate windows. To support exploration using multiple windows, the parameters from each view are linked. Integration of the model, height field and volume views is described in section 5.

Each point shown on the map view can be selected by clicking or dragging the mouse in the display window. The closest point to the mouse is highlighted and the spectrum data relating to that point is used to create a spectrum view. As the mouse is dragged between points the spectrum display changes interactively. The cohesion between the views provides a fast and effective way to view the data using the tools that have been created.

Moving the wavelength slider in the map view automatically updates the spectrum view so that the same wavelength is shown consistently in both views. It is important to maintain this single environment to avoid confusing the user.

5 Conclusion and Future Work

The tools created provide a great level of flexibility for exploring the dataset. They are considered promising and the project will continue until the aims presented in this future work section have been fulfilled. Four interactive visualisation views have been created to present the data in a variety of different ways. Each is successful at presenting different methods of visualising the dataset, and each is designed to be very flexible. Proposed ways to enhance these tools follow.

A new view is proposed called the *volume view*. It would represent the existing data as a stack of 2D maps, superimposed one in front of another, like a deck of cards. Each 2D map would correspond to one sampled wavelength. The stack of maps would effectively create a third "wavelength" dimension. It would still only be possible to see one map at a time, but the volume created by the stack would give a stronger sense of the current position in the NIR Spectrum data. [12, 17].

Integration of the map and spectrum views greatly simplified navigation within the dataset. The views will continue to be integrated until all views are collected into the one application. The 2D map view, the height field view and the volume view all present the same information but in slightly modified format. These will be combined to exist in the same window frame. A menu will be provided to switch between the views. The model and spectrum views each present different viewing concepts so they will occupy windows of their own. To support simultaneous exploration using multiple views, the parameters from all views will be linked.

The spectrum view would be more useful if it was modified to act as a toolbox for manipulating the contents of the other views. The existing smush slider could be combined with a wavelength slider and bounding lines. The bounding lines could select the range of intensities the user is interested in investigating. This would provide a tool for controlling the levels of quantisation in the display. The bounding values would update the colour values used in the spatial displays interactively. The wavelength slider could be presented in the spectrum view so that all the tools for manipulating the views are in the same location.

It is currently possible to click in the map view to select a point for displaying in the spectrum view. If the views were integrated it would be useful to be able to select points in each of the spatial displays, and have each of the displays update interactively. This would require rotating the object in the model view for the highlighted point to appear at the "front".

The fruit model used in the model view uses a sparse sample of points to reconstruct a fruit shape. It would be possible to use Magnetic Resonance Image (MRI) data taken from a single fruit sample to represent an ideal fruit shape [1]. A more accurate reconstruction could be created using this information, providing a display model that looks realistic.

It is possible to measure a fruit at different times in its life cycle. By collecting this time dependent data, a fifth dimension could be added to the visualisation problem. It would be interesting to investigate this data to gain some insight into the evolution of ripening characteristics in the fruit.

The equirectangular map projection distorts the fruit geometry to a greater degree than is desirable. Implementation of different map projections will produce maps with less distortion [14].

When mapping wavelength intensities to colour values it is important to consider what is being compared. If comparison is made between fruit then the colour map should be consistent, and should therefore be computed using the maximum and minimum values from all fruits. However, this may inhibit the ability to distinguish different intensities on a particular fruit; if the user wants to examine a particular fruit's intensities, the colour map should be based on the maximum and minimum values for that fruit alone. As the choice of strategy depends on the user's immediate goal, the interface should supports both approaches.

References

[1] http://wwwcivm.mc.duke.edu/default.html. The Center for In Vivo Microscopy.

[2] T. A. DeFanti, M. D. Brown, and B. H. McCormick. Visualisation, expanding scientific and engineering research opportunities. In Gregory M. Nielson, Bruce Shriver, and Lawrence J. Rosenblum, editors, *Visualisation in Scientific Computing*, pages 32–47. IEEE Computer Society Press, 1990.

[3] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics, Principles and Practice*. Addison Wesley, 2nd edition, 1990.

[4] A. S. Glassner. *Graphics Gems*, chapter 8, pages 462–463. Academic Press, 1990.

[5] R. Hall. *Illumination and Color in Computer Generated Imagery*. Springer-Verlag, 1988.

[6] S. Kuwano, T. Fujiwara, and M. Iwamoto. Nondestructive determination of sugar content in satsuma mandarin using near infrared (NIR) transmittance. February 1993.

[7] B. H. McCormick. Computer graphics: Special issue on visualisation in scientific computing. Technical Report 21(6), ACM SIGGRAPH, November 1987.

[8] M. Murakami, J. Himoto, and K. Itoh. Analysis of apple quality by near infrared reflectance spectroscopy. September 1993.

[9] M. Murakami, J. Himoto, M. Natsuga, and K. Itoh. Analysis of pumpkin quality by near-infrared reflectance spectroscopy. June 1992.

[10] J. Neider, T. Davis, and M. Woo. *OpenGL Programming Guide*. Addison Wesley, 1993.

[11] J. K . Ousterhout. *Tcl and the Tk Toolkit*. Addison Wesley, 1994.

[12] L. J. Rosenblum. Visualisation of experimental data at the naval research laboratory. In Gregory M. Nielson, Bruce Shriver, and Lawrence J. Rosenblum, editors, *Visualisation in Scientific Computing*, pages 245–253. IEEE Computer Society Press, 1990.

[13] G. S. Smith and J. P. Curtis. A fast and effective method of measuring tree structure in three dimensions. Technical report, The Horticulture and Food Research Institute of New Zealand Ltd, 1995.

[14] J. P. Snyder. *Flattening the Earth*. University of Chicago Press, 1993.

[15] C. Ware. Color sequences for univariate maps: Theory, experiments, and principles. *IEEE Computer Graphics and Applications*, 8(5):41–49, September 1988.

[16] B. Welch. Practical programming in tcl and tk. Draft Manuscript, 1995.

[17] G. Wyvill, 1996. Informal Communication.