20

Expanding the Role of Computer Graphics in the Analysis of Survey Data

W. A. Boismier
Planning Department, Hampshire County Council, Winchester

Paul Reilly
IBM UK Scientific Centre, Winchester

20.1 Introduction

The Historic Landscape Project (HLP) is an innovative venture being performed by archaeologists in Hampshire County Council's (HCC) Planning Department. The principal aims of the project are twofold: first to obtain a representative inventory of the archaeological resources occurring on County Council-owned land, and second, to provide a description of both the negative and positive effects of current land-use practices on these resources. The basic units of information are derived from artefacts gathered along systematically spaced transects using a variety of collection techniques, including quadrat provenance and shovel-testing methods. As the area of study is extensive and resources are scarce, a sophisticated sampling strategy has had to be developed. The sampling design and the coordination of the field survey are the work of Bill Boismier and are described elsewhere (Boismier 1986, Boismier & Reilly 1986).

Activity areas are defined with the computer on the basis of mapped density characteristics as discrete high-density areas of artefacts. At the outset of the project, SYMAP contour maps were produced on the University of Southampton's mainframe to help evaluate the sampling design. These indicated that the collection strategies employed in the field survey, and the subsequent mapping of artefact densities per unit area, accurately locate high density areas under a number of varying surface conditions. Experimentation convinced Boismier that by changing the SYMAP contour-intervals it was possible to detect evidence of post-depositional patterning in artefact distributions caused, for instance, by cultivation and down-slope movement (Boismier 1986).

The Historic Landscape Project had been active for some considerable time when contact was made with IBM's UK Scientific Centre (UKSC). After some discussion it was agreed that the analysis of these distributions might benefit from the application of more sophisticated graphical display techniques. There were two aspects to this. Firstly, at a mundane level, the whole process could be made much more efficient if certain key phases could be automated. For example, the procedure for generating the final distribution maps for the HCC planners was labour intensive and relatively slow. The density plots showing activity areas were produced using SYMAP, but in order to produce distribution maps these data were transferred by hand on to a base map by a draftsperson. The base maps are not held on machine and therefore a new base map has to be redrafted by hand for every distribution plot. A straightforward solution is to digitise and store the base maps on computer and superimpose the distribution maps automatically.

Perhaps of greater scientific interest was the notion that that analysis of these data would be improved if they could be mapped onto three-dimensional terrain models. At the very least this could be expected to provide for the more rapid interpretation of these interpolated distributions. At a deeper level it was hoped that such an approach would facilitate the production of new and interesting types of questions about the data, principally in terms of how far land use patterns and topography actively influenced distribution patterns.

The graphics aspect of the work has only recently been put into operation, and developments are still taking place. This communication will therefore be restricted to a discussion of some the techniques being applied, and some of the problems being faced, rather than a presention any substantive analytical results. It is anticipated that this information will be available later in the year.

Before any graphical work could be attempted the data had to be transferred into a suitable computer environment. Essentially the rôle of the computer can be divided into six parts:

- data validation;
- database retrieval;
- graphics;
- digitisation;
- interpolation; and
- terrain modelling.

The last three items are actually a subset of the graphics problem.

20.2 Data collection, input, and validation

A consequence of the project being well-advanced when the possibility of three-dimensional graphics facilities became available was that the HLP recording system was already established. *Pro-forma* recording sheets had been designed and mass-produced, and large numbers of them were already filled in. This meant that the first task was to write a suite of data capture programs which would enable data collected on paper records in the field to be input into the UKSC's mainframes. The design of the data capture programs was tightly constrained by the existing paper recording format.

These programs provide basic error-trapping facilities as the data is entered, and an editor has been purpose-built to allow the project supervisor to browse through the data files and amend factual errors or update records at a later stage. It is only after the data have been passed through both validation procedures that the data will be entered into the actual database system.

20.3 Database retrieval

Given the extremely large volumes of data being generated, it is extremely laborious and time consuming to process the considerable amounts of data the project generates without the aid of a DBMS. The system being used is a UKSC relational database known as the Peterlee

Relational Test Vehicle (PRTV). As the research is principally concerned with the identification of particular types of activity area, a typical task would be the retrieval of all the burnt flint found within a specified property unit.

Once retrieved, the spatial component of the information can be represented graphically as a point-provenance distribution map. Different colours or colour scales can be used to represent variables such as weight or frequency. This is facilitated by a bridging program (written in-house in PL1), between the database system and a suite of 3-D graphics facilities called the Winchester Graphics Sub-System (WGSS). Both WGSS and PRTV are subsets of the Winchester Graphics System (WGS). As the WGS system has been described a number of times, there is no need to go into detail here (Colley & Todd 1985, Heywood *et al.* 1984). The particular point of interest to be brought out here is the bridging program between the database and the graphics systems. In principle, the construction of comparable bridging programs should be possible between many commercially available DBMS and graphics packages. Within UKSC there are several such bridging programs, including one between Prolog to GDDM (Graphical Data Display Manager). Indeed, one or two archaeological database systems with 2D graphics capabilities have already appeared (e.g., Alvey & Moffett 1986).

20.4 Digitisation

In order to ascertain the relationship of the local terrain to the distribution of finds in the area it is necessary for us to build a model of the terrain. The contour information is digitised using a table digitiser and stylus, and this information is then interpolated onto a regular grid as the basis of a terrain model. Terrain models may be displayed as vector or raster models. The vector or wire-framed models have the advantage of being susceptible to real-time viewing transformations such as rotations, pans, or zooms. Interesting activity areas can be isolated rapidly in this mode. Having isolated some interesting aspect a high resolution solid model of the terrain and distribution pattern can be computed and displayed on a raster unit.

20.5 Interpolation

As the survey data collected only constituted a sample it is also necessary to interpolate these data in order to obtain the required generalised density contour maps. Other programs are involved here. The SYMAP program continues to be of value, but we are continually assessing the relative merits of various interpolation algorithms. Cheaper and more efficient hardware mean that it is practicable to progress beyond the two-dimensional isoline contour maps. These remain a good alternative for line-printer technology, and, moreover, current publishing methods favour two-dimensional display methods. However, the potential of three-dimensional display is a much more exciting prospect allowing far more information to be represented on a single display. There is still a lot of scope for progress in this step in the exercise, and improvements continue to be implemented. The first interpolation algorithm we used was not judged to be completely reliable, especially in areas where a paucity of datum points happened to be adjacent to a area where dense concentrations of readings occurred. An implementation of the Akema algorithm, in which the data are first triangulated and fitted with a Quintic polynomial, was used unsuccessfully as an interpolant. An apparently much more robust algorithm, from our point of view, is that given by Shepherd. Shepherd's method is much simpler than Akema's, and is based on the use of weighted averages. The weightings are a function of the distances to the reference points considered. A minimum number of points to be referenced and the effective weighting range is specifiable by the user.

20.6 Activity areas in a landscape context?

One of the main points of this experiment is to try to assess the affect of topology on the distribution of artefacts in the plough-zone. Three-dimensional graphics systems offer the real possibility of mapping the interpolated distributions onto 3-D terrain models. In this way one can begin to take account of the effect of local topography for instance. The most successful method of achieving this from our efforts so far is by interpolating both the topographic and the sampling data onto an identical grid of points. A bi-quadratic b-spline is then used to produce the final terrain model. The wires represent the terrain and the colours or shades indicate densities.

One can establish whether artefact distributions appear to have been skewed owing to changes in slope for instance. We can best illustrate this process with reference to a real data set. Consider the colour picture, Fig. 20.1, which shows the distribution of modern pottery over a parcel of Hampshire County Council-owned land in Fair Oak Hampshire. The height readings, taken from the base-map have been accentuated to bring out the changes in the slope. One immediately notices a concentration of material lying at the top of the slope spilling down to a lesser concentration of material at the bottom of the gradient. Rather than inferring the existence of two activity areas, this might prompt us to ask whether material has been carried down the slope. One might hypothesise, for instance, that ploughing or soil-creep may be responsible for the apparent down slope movement. However, the reasons why this patterning should occur are obviously matters of interpretation which cannot be gone into here. The main point to be brought out is that displaying sampled survey data in this way prompts the investigator to look at their dispositions in a way which would otherwise be impossible. This can only contribute to the generation of richer theories and explanations of the patterns observed. In many cases these can be tested using other data collected from the same sampling area. On might wish to contrast the patterns caused by different classes of artefacts with regard to period, or perhaps size or material. Since the data are held on a database system the user is in a position to explore theories rapidly.

Occasionally the observed patterning will warrant a revisit to the site and a more detailed gridded collection strategy will be employed in order to improve the resolution of the data. The same procedure may show up distinct activity zones such as discrete chipped stone scatters within larger activity zones. These surveys can also be used to cross-check whether the interpolated activity is accurate.

Displaying the gridded data presents it own special problems; we have tried and abandoned several approaches as worthless. For instance, it was found difficult to understand open colour markers, representing counts or size classes for instance, laid over the grid. Comprehension was improved when area-filled panels were used. This is done by generating a mesh over the panel with vectors or with solid patches on a raster system. Colour patches are only of use if there is a fine colour blending or some readily understood colour-coding such as the spectral sequence. Contrasting colours are useful for showing complimentary distributions but not local densities.

Although work using the system just described has barely begun, it is suggested that the use of this type of *investigative loop*, using powerful database and graphics systems, is a potentially powerful tool for field archaeologists trying to process the results of large and rapid surveys. The basic idea of being able to interact with the data, display various distributions and formulate

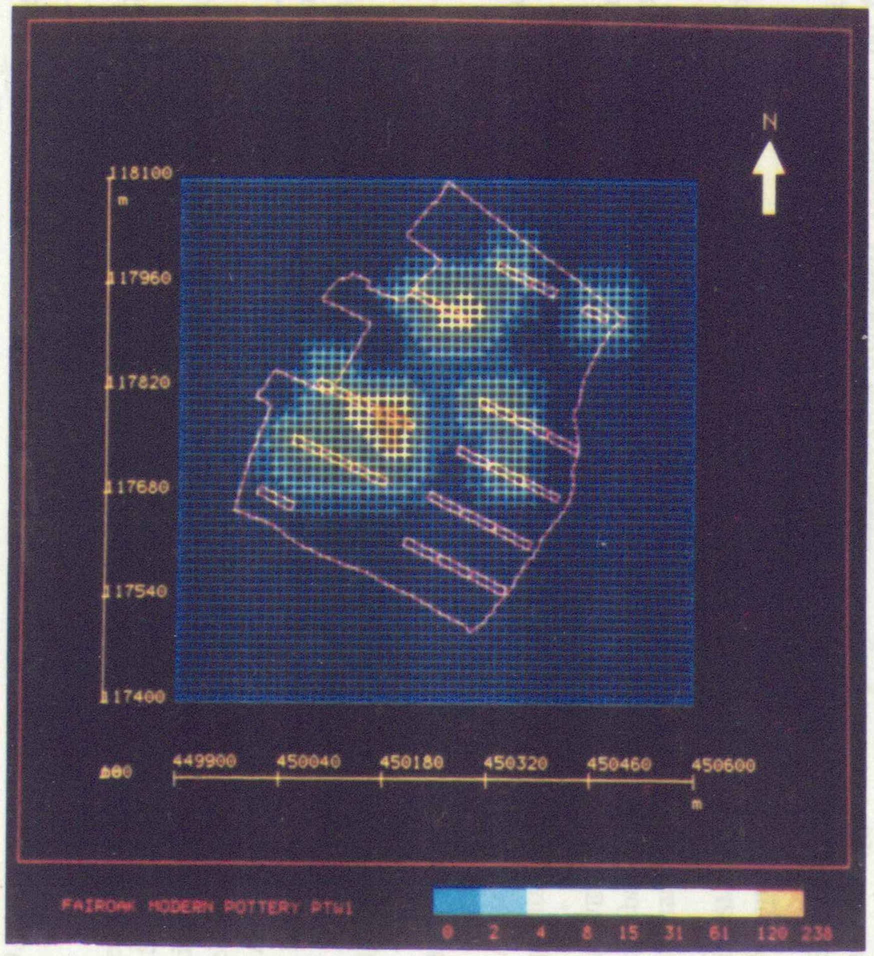

Figure 1.

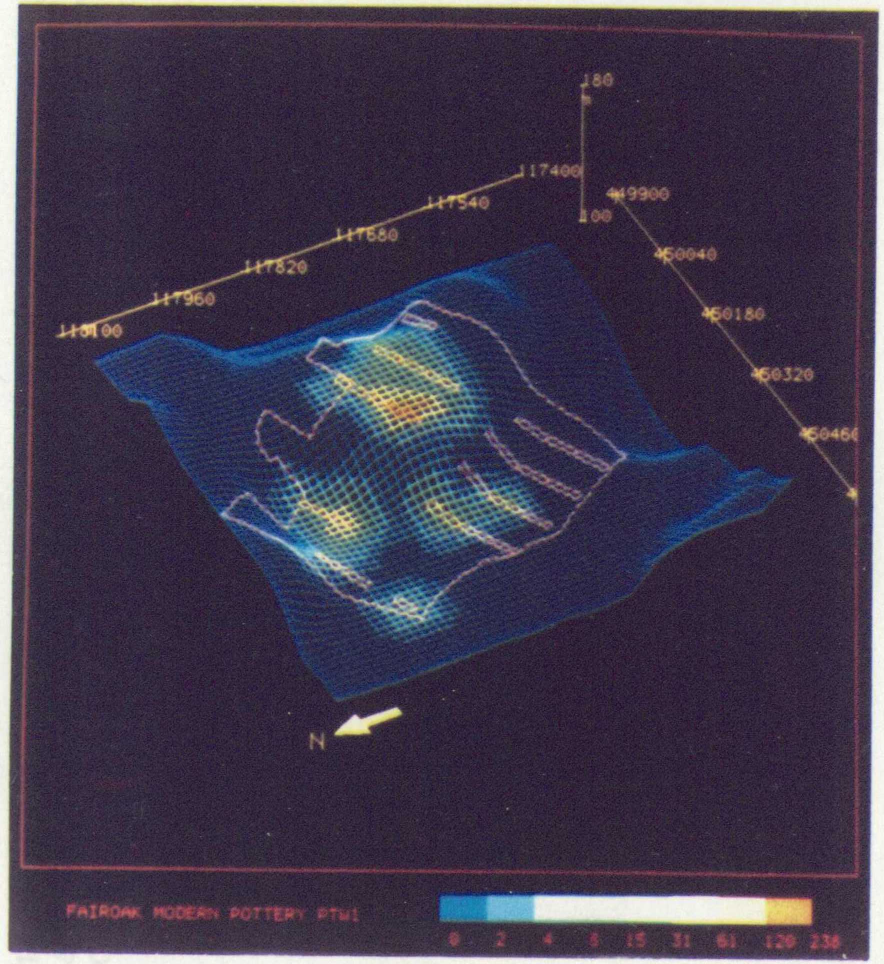

Figure 2.

theories to explain them, means that hypotheses can be followed up and tested quickly.

Since one of the main purposes of this type of survey is to gather data from large areas as quickly as possible and formulate cultural resource management policies the approach could well have a major impact in the future.

As a final rider, as optical disc technology continues to reduce storage costs, the archiving of survey data in processed image form and as a raw database is a viable option. As the costs of storage and processing of three-dimensional information continue to plummet this sort of approach will soon be within the reach of many archaeologists.

References

AKIMA, H. 1978. 'Bivariate interpolation and smooth surface fitting for irregularly distributed points', *ACM Transactions on Mathematical Software*, 4, pp. 160–4.

ALVEY, B. & J. MOFFETT 1986. 'Single-context planning and the computer: the plan database', *in Computer Applications in Archaeology 1986*, pp. 59–72, University of Birmingham.

BOISMIER, W. A. 1986. *'The Historic Landscape Project'*, pp. 9–12, Archaeology Section, Hampshire Planning Department.

BOISMIER, W. A. & P. REILLY 1986. 'Using computer graphics and DBMS in non-site archaeology: work in progress', *Archaeological Computing Newsletter*, 7, pp. 6–10.

COLLEY, S. & S. TODD 1985. 'Archaeological computing at IBM UK Scientific Centre', *Archaeological Computing Newsletter*, 4, pp. 3–8.

HEYWOOD, T. R., N. B. GALTON, J. GILLET, A. J. MORFFEW, P. QUARENDON, S. J. P. TODD, & W. V. WRIGHT 1984. 'The Winchester Graphics System: a technical overview', *Computer Graphics Forum*, 3, pp. 61–70.