COMMON-SENSE AUTOMATED CONTOURING / SOME GENERALIZATIONS

CHRISTOPHER M. GOLD, *Earth Sciences Division, Alberta Environment, Edmonton, Canada*

INTRODUCTION

MANY PAPERS and programs have been written on the automated contouring of arbitrarily distributed data points, starting perhaps with Dayhoff (1963). There is even a reasonably large list of reviews of contouring techniques, e.g., Walters (1969), Crain (1970), Peucker (1972), Rhind (1975), Gold (1979). While it is tempting to review the reviews, a more fruitful approach to evaluating alternative techniques appears to be that based on a breakdown of their methods and objectives into basic categories or stages (Gold, 1983) and an evaluation of their various properties. The author's geological activities form the background for this and previous studies. While many types of data are derived from that discipline, some of which are described in Gold (1980b), the primary problems coming to the author's attention concern 1 anisotropic data distributions (especially sampling along traverse lines of various types); 2 sparsity of data points by comparison with the features to be observed; and 3 the need to preserve the precision of these expensive observations. These last two concerns are particularly relevant in the processing of drill hole information from the oil and gas industry, where the cost per data point is extremely high.

DATA PROPERTIES AND MAPPING ASSUMPTIONS

The above concerns impose special conditions on the contouring or interpolation process that differ from those required for photogrammetry or computer-aided design. Photogrammetric data are usually fairly closely spaced (although traverse-type distributions are common) and possess fairly identifiable error components. Computer-aided design interpolation techniques assume that the surface is sufficiently well sampled to be modelled by the low-order polynomial patch being used. Methods intended for photogrammetric data would usually allow some form of generalization – smoothing, averaging or data reduction – that reduces the information content of the final map.

Methods intended for computer-aided design (CAD) will frequently introduce artifacts – surface features related to the polynomial patches rather than to the data. This occurs because some data distributions (for example one high value in

a grid of zeros) cannot adequately be modelled with low order polynomial patches. CAD methods assume that the placement (X Y) of individual points are 'surface specific' – that is, the location in itself contains information about the surface. This explains why the most successful applications of 'triangulated irregular networks' as digital terrain models have occurred when the data points and triangles have been hand picked to represent the terrain. Artifacts are, of course, not restricted to patch type techniques. It is common with grid-based techniques to observe 'string of beads'-type features representing a ridge or valley in the original surface. These beads may result from inadequate sampling of the feature in the first place, but quite often the fact that they are of comparable size to the grid spacing indicates that they are an artifact of the estimation-on-a-grid approach.

Three other issues concerning geological contouring should be touched on briefly. Firstly, slope information is occasionally available from dipmeter surveys, outcrop measurements, etc., and must be accepted as definitive along with the elevations when they occur. Secondly, the data have been recorded with some definable quantizing interval. If this interval is very coarse, and a precise interpolation technique is used, undesirable surface undulations will result from the incorrect assumption that the observations are 'precise'. The third issue, the problems of handling isopach maps or formation thickness maps has been dealt with elsewhere (Gold, 1980a).

The data we are concerned with are therefore, in summary:

Precise (including slopes on occasion)
Sparse
Anisotropically distributed.

The mapping techniques to be used must therefore generate surfaces that:

Honour data point values precisely
Honour slope values
Possess no artifacts (features caused by the mapping method itself – e.g., beads or cliffs).

CLASSIFICATION OF TECHNIQUES

We have only described one of many possible types of 'contouring problems'. As is so often true, solutions are frequently based on a breakdown of the problem into appropriate stages. The next step is, therefore, to look at the available tools and, if possible, to categorize them. One of the biggest historical problems in computer contouring is that systems design has often been driven by the availability of particular techniques rather than the needs of the data and the application.

One possible way out of this situation is to attempt to formalize the stages in producing a contour map (or other surface representation) and categorize all techniques in order to assist in the selection of the most appropriate one for each stage. One tentative list of stages is:

1 Data point entry and retrieval

2 Sample site selection for surface estimation
3 Neighbouring point selection
4 Surface estimation procedures
5 Display methods

STAGE 1: DATA ENTRY

When data points are read at the beginning of the program they must be stored in some suitable fashion so that they may be easily recovered when needed. For small data sets the simplest solution – save them in memory in the order they are read in – is probably the best. Operational mapping requirements need to handle data points by the thousand, however, and here more efficient searching of the data is required. Since points are requested based on their X Y coordinates, some form of geographical data base is required. A simple approach would be a set of 'pigeon holes' covering the map area so that only a few percent of the data points would need to be scanned to find if they were sufficiently near to some of point of interest to be used. Of course, for any pigeon hole configuration, a set of data could be designed to defeat its purpose.

Another approach would be to generate a triangular or polygonal network with data points at the vertices (e.g., Gold, 1977). This eliminates the problems of Cartesian coordinate-based storage systems as networks are adaptive to the actual data distribution. Additionally, they form an implied binary tree structure for neighbourhood searches. (Gold and Maydell, 1978). However, the storage cost and complexity are high for small data sets.

STAGE 2: SAMPLE SITE SELECTION

One place where 'common sense' is rarely practised in contouring is in deciding at which locations surface elevation estimates are to be made prior to stringing contours between them. Starting with one of the first contouring packages (Dayhoff, 1963), various methods have been devised for threading contours through values placed on a square grid. As discussions of the 'saddle point problem' have shown, even this is ambiguous, and a sampling scheme based on an equilateral triangular grid would be preferable. Nevertheless, the starting point for most mapping packages is "We can produce contours from a square grid. How do we obtain the grid?"

While the square grid is fundamental to many mathematical transformations desired by some end-users, it has serious limitations. The grid spacing tends to be either too coarse to pick up the desired features or so fine that storage and computation costs escalate, making the grid a valuable commodity that must subsequently be preserved and treasured. Again, data distributions can be generated that will make any grid appear ridiculous. In addition, it is clear as a matter of principle that no surface generated in this fashion can be guaranteed to pass precisely through all data points (perhaps with the correct slope) unless major assumptions are made about the data distribution, e.g., – one point per cell.

If it is required that the surface pass precisely through each data point, however distributed, the sample sites (used as an intermediate step between the raw data and the contour generation) may not be selected on a regular basis, but

each data point must correspond to a sample site, with additional sample sites somehow arranged in between.

One procedure that comes to mind is the subdivision of the map area into triangles with data points at vertices, as mentioned previously. Each triangle is then subdivided into N-squared sub-triangles where N, the number of equal parts into which each triangle edge is subdivided, is a resolution or smoothness factor. This generates a local regular triangular mesh, and the surface is estimated at the nodes of this mesh.

In this or any other adaptive sampling scheme, the actual contour-threading procedure would have to be modified. Nevertheless, this is not an especially difficult procedure.

STAGE 3: NEIGHBOURING POINT SELECTION

If we exclude global methods, for example polynomial regression techniques, where perturbation of one data point affects the estimates throughout the map, all local techniques require the selection of a subset of the data points that are deemed to influence the surface at the sampling locations. Kriging techniques, where a semi-variogram is generated prior to actual surface estimation, provides a mechanism for generating an appropriate cutoff circle or ellipse. Other approaches require the user to enter an all-too-arbitrary rejection radius. However derived, these functions are global rather than adaptive – that is, they assume similar data distribution properties across the map area. The addition of extra data – either within or adjacent to the previous point set – may invalidate these global assumptions.

Another problem concerns the potential anisotropy of the data distribution. Some systems permit the specification of an elliptical search radius, but this is still not adaptive to local data variations. Many programs specify or permit an expanding search-radius feature terminated by the discovery of an 'adequate' number of points within the whole circle or within each quadrant or octant of it. Again, for any particular strategy a data set may be prepared that defeats it. In particular, traverse data (ships' soundings, redigitized contours, etc.) may be so anisotropic that many programs cannot handle it: too many points fall in some directions before any are located in the others.

The problems are twofold: firstly, program arrays may not be large enough to handle apparent pathological (although not necessarily infrequent) situations; and secondly, a surface estimate at any sampling location based on many tens of points may become virtually meaningless – uncontrolled surface smoothing or averaging takes place. At the risk of repetition, an adaptive system is required – e.g., the triangular mesh as long as the criterion for a 'good' triangulation is reasonable – for example, the Delaunay triangulation. In this case neighbouring points are defined as being close within the network, which need not imply being close in Cartesian space.

STAGE 4: SURFACE ESTIMATION PROCEDURES

Once the raw data have been saved, the surface sampling scheme defined and the appropriate neighbouring points selected, a suitable local procedure is used to

estimate the surface elevations. Many different estimation procedures have been devised, often in an attempt to eliminate problems in some of the other stages described here. Most, but not all, procedures can be considered to be either of patch type or weighted average type.

Patch procedures usually involve selecting appropriate domains within the map area and then fitting some mathematical function, usually a polynomial in x and y, to the data points within that domain. With suitable overlapping of adjacent domains and appropriate mathematics, the patches can be made to fit or merge to some extent.

Two concerns arise with patches. If the domains cover many data points the resulting surface will probably not pass precisely through these data points or agree with their slopes. More serious, however, is the problem that the degree of continuity across patch boundaries will usually be considerably less than that at the centre of the patch. For example, the patches for two adjacent domains may be cubic polynomial functions. The best mathematically-proven fit between these patches will probably be only linear; the surface itself is continuous between the patches, and so is the slope. The quadratic and cubic continuity within one patch will not extend to the other. This means that for many data configurations the selection of patch domains and boundaries will affect the final surface form.

If we take triangular domains as our example (e.g., Gold 1977) these issues will be well illustrated. Once a basic triangulation has been defined it is easy to treat each triangle as a domain and proceed from there. If our patch is linear, flat triangular plates can easily be generated that fit the three data points at the domain boundaries. Clearly the resulting surface is continuous, and equally clearly the slope may change abruptly at the boundary between two plates. If we specify a cubic patch in x and y for each triangle there are ten coefficients to be determined. Other information is needed besides the three elevations – usually the slopes at the vertices and some constraint on slope behaviour along the edges. Since the slopes at the vertices are not usually measured in the field they have to be estimated from nearby points (see Stage 3). After all this mathematics (which is computationally expensive) we have a function with cubic continuity within the triangle and only linear continuity across the edges. The final surface therefore has breaks in slope change along triangle edges. While this behaviour is often valuable in computer-aided design applications, these artifacts of the program can cause considerable confusion to those familiar with extracting surface behaviour changes from contour maps. The problem is less serious in applications where data point locations (and triangle edges) are selected to show breaks in slope, but for arbitrarily-located data points the artifacts introduced by the method may render the resulting maps unacceptable.

The other main class of procedures, weighted average methods, is derived on a different basis. As the name implies, the underlying (and often unrecognized) assumptions are statistical. It is presumed that each neighbouring point will provide an estimate of the surface elevation at the desired sampling location. This estimate will have a significance (weighting) depending (among other things) on how far that data point is from the sampling location. The result is a tabulation of

elevation estimates and associated weightings. The simplest (and most obvious) statistic to calculate is the weighted average or mean which will be used as the surface estimate, but, in so far as weighted statistics are valid in the first place, a weighted variance or standard deviation is just as legitimate. This permits the usually unanswerable question "What is the error at any point on this map?" to be addressed on some workable basis, assuming the values to be correct at the data points.

So far we have assumed the presence of any workable weighting procedure for any data point with respect to the sampling location, and also an appropriate surface estimation. The simplest surface estimate that can be produced from a data point is its elevation: zero slope is assumed between the neighbouring data point and the sampling location. Another approach would be to fit a simple surface through the data point and its neighbours. This would provide a more sophisticated estimate at the sampling location. In fact, it is becoming more common to fit a plane through each data point (if slope data were not available as input) than used to be the case, and results are generally considered to be better. It should be noted, however, that this permits the extreme values of the surface (local maxima and minima) to be located at other than data point locations. For arbitrarily located data points this is often reasonable and, among other things, permits accurate modelling of planar sloping surfaces. If zero slopes are assumed at data points then each data point must be a local maximum, minimum or inflection point. For some kinds of map (e.g., bathymetry) the intended function of the map precludes interpolated extrema, and the zero slope facility must be used.

The weighting function itself has not yet been discussed, in part because so many alternatives have been attempted in the past. Its fundamental property is that it decreases with increasing separation between the data point and the sampling location, although the presence of nearby data points can be expected to reduce the overall influence of the data point being considered. Clearly, if a sample location coincides with a data point the weighting of that data point should be unity and all others zero if the surface is to fit all data points precisely; this is less important if some uncontrolled smoothing of the data is acceptable. In addition, as a data point approaches a sampling location, the influence of the data point should approach unity asymptotically if continuity of slope is to be preserved.

This is not all. A more subtle problem involves the outer limit of weighting contributions. If a point is dropped from the list of relevant neighbours before its weighting has decreased to zero the potential exists for a surface discontinuity; a minor perturbation of the sampling locations will include or exclude a data point contributing significantly to the surface estimate. While in many cases the discrepancy, even if not zero, will be trivially small this cannot be guaranteed to be the case particularly when the data points are ill-distributed around the sampling point. In this case, care needs to be taken that a cluster of points to one side of the sampling location does not entirely overwhelm the influence of a single point, either near or far, on the other side; or, conversely, that nearby

points entirely shadowed by even nearer points do not acquire an influence not rightly theirs.

On a more cheerful note, some but not all of the above problems can be dealt with by ensuring that the neighbour selection technique discussed previously selects at least all those points with non-zero weighting, perhaps permitting some small positive weightings to be considered trivial. In a similar vein, small mathematical discrepancies will not be noticeable in some applications. Nevertheless it is valid to ask, for any particular package, whether sophisticated weighting techniques may not be defeated by overly simplistic neighbour selection.

While polynomial weighting functions of the 'inverse squared distance' variety are the most common, it should be remembered that Kriging techniques fall into the same weighted average category, and can suffer from similar problems.

STAGE 5: DISPLAY

This is the final stage of tools to be evaluated prior to selecting or producing a contouring system. While in fact the least important component, in many cases it is the driving force behind the design of a system, as previously mentioned. It is necessary to define a surface interrogation technique that fits the job specifications and only then consider the display procedure – whether to generate contours or other products. Working the other way round is a case of the tail wagging the dog, although there is no denying that the details of contour suppression, labelling, etc., may form a major part of the initial programming effort.

SYSTEM DESIGN DECISIONS

The key point of this paper is that once the different stages have been defined it is possible to select the appropriate programming tools on a systematic basis. It is normally assumed that the design decisions for one part of the job will be appropriate for several others, since particular approaches often overlap several of the stages described earlier. This may not be true, especially as no one technique is appropriate for all types of mapping problems. If the job specifications can be defined in terms of the stages outlined in this paper, or some other appropriate set, the chances of nasty surprises later on can be reduced. For example, let us look at the problem that initiated this project – the requirement to contour large numbers of arbitrarily distributed data points with precise values but with no guarantee that all features on the surface of interest had been well sampled.

For Stage 1, data point entry and retrieval, it is noted that large numbers of data points are anticipated. Hence some kind of internal data structure for storage and retrieval of information would be appropriate – for example, pigeon holes or a triangular mesh. Of some concern here would be how badly the data could be distributed, as this could affect retrieval efficiency. The requirement for precise agreement between the generated surface and each data point would point very strongly towards a triangulation-based sampling location scheme

(Stage 2), unless the data could be considered near-uniform in distribution. If the data are expected to be very irregularly distributed a simple (fixed or expanding) search radius for neighbouring point selection (Stage 3) would be inadequate and a system must be used that is capable of adapting to varying point-densities in different directions. This is absolutely critical if the data are likely to be of the traverse type – very dense along arbitrary paths but with wide spacing between paths.

Since there are no *a priori* grounds to assume that all features of a desired scale were adequately sampled by the data points provided, a surface patch approach is probably inappropriate for Stage 4. The choice of an appropriate weighted average function is still something of a research topic, although some systems exist (mostly proprietary) that are thought to be adequate. Since smoothing of the surface is not desirable the weighting must reach unity when a sampling location coincides with a data point. Beyond this the primary concern is that the neighbourhood search retrieve all those data points with significantly non-zero weightings. The display methods (Stage 5) should be appropriate to the sample site selections of Stage 2.

If, to take another example, a smoothed map was an appropriate product the previous decisions must be re-evaluated. First of all we must not only decide how much smoothing is desirable, but also what is understood to be the smoothing process. A common approach is to select a grid size sufficiently coarse to preclude the retention of small-scale features; unfortunately this does not necessarily prevent small-scale features that happened to be sampled by the grid-node scheme from being preserved as larger features – the size of the grid squares. Thus many grid-based maps have grid-cell sized features that are really artifacts of the method. A second approach is to allow the weighting of a data point to be less than unity when it coincides with a sampling location, so that a data point is modified by some average of the neighbouring points – a rough form of convolution. Lastly, smoothing may be used in a more rigorous sense to mean frequency filtering, so that all high frequency short wavelength components of the surface are removed by convolving the data with an appropriate filter function. Since the mathematics for this are only known for regular grids this approach must either be ignored, however desirable, or preceded by a sufficiently fine gridding process. Depending upon the chosen process, Stages 2 to 5 of the previous example may need to be revised.

CONCLUSIONS

Much of the confusion about the desirable properties of contour maps and how to achieve them may not be necessary. The foregoing discussion introduces no new techniques, but may assist in the definition of the mapping problem in the first place and the selection of the best available techniques in the second. With an understanding of the available tools, their implications for each stage of a computer mapping project can be specified. The limitations of each technique should immediately raise questions about the properties of the proposed target data set and the suitability of that technique for the job, since there is clearly no one approach that is appropriate for all types of data. What for one data set may

be trivial problems can easily be prohibitive for the next, and rarely are the assumptions on which a particular program was based laid out in the documentation. With this understanding it should be possible for the mapping requirements to dictate the choice of tools or procedures, rather than limitations of individual techniques dictating how well or poorly any particular data set can be mapped.

REFERENCES

CRAIN, I.K., 1970. Computer interpolation and contouring of two-dimensional data: a review. *Geoexploration*, vol. 8, pp. 71–86.

DAYHOFF, M.O., 1963. A contour map program for X-ray crystallography. *Communications of the Association for Computing Machinery*, vol. 6, pp. 620–622.

GOLD, C.M., 1977. The practical generation and use of geographic triangular element data structures. First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, *Proceedings*. Laboratory for Computer Graphics and Spatial Analysis, Harvard University, vol. 5, 18 p.

——— 1979. Triangulation-based terrain modelling – where are we now? *Proceedings*, AUTO-CARTO IV, Washington D.C., vol. 2, pp. 104–111.

——— 1980a. Drill-hole data validation for subsurface stratigraphic modelling. *Computer Mapping of Natural Resources and the Environment*, Harvard Library of Computer Graphics, pp. 52–58.

——— 1980b. Computer mapping in geology. Chapter 8 in D.R.F. Taylor (ed.), *Progress in contemporary cartography*, John Wiley and Son.

——— 1983. Some problems in the mapping of geological surfaces (Abstract). *Proceedings*, AUTO CARTO VI, Ottawa, October 1983, vol. 2, p. 414.

GOLD, C.M. and MAYDELL, U.M., 1978. Triangulation and spatial ordering in computer cartography. *Proceedings*, Canadian Cartographic Association third annual meeting, Vancouver, 1979, pp. 69–81.

PEUCKER, T.K., 1972. Computer cartography: Commission on College Geography Resource Paper no. 17, Association of American Geographers, Washington, D.C.

——— 1978. The triangulated irregular network. *Proceedings*, Digital Terrain Model Symposium, American Society of Photogrammetry, St. Louis, May, 1978.

WALTERS, R.F., 1969. Contouring by machine: a users' guide. American Association of Petroleum Geologists, *Bulletin*, vol. 53, pp. 2324–2340.