VB-GIS 3D - A GIS specifically designed for Archaeology

Carlos Reynoso

Damián Castro

VB-GIS 3D is an experimental computer system for exploration, analysis, hypothesis testing and statistics in spatial or distributional archaeology. It is a working platform programmed in Microsoft Visual Basic for Windows, strictly based on an artificial intelligence prototype designed in PDC Prolog a couple of years ago. It also works on Windows NT environments. The underlying AI model is used to control the consistency of the procedural calculations, the sufficiency of the data for the statistical processing and the accuracy of the interface.

The main design goal is to provide a single open system with a state-of-the-art graphic interface and a strong mathematics engine. The program is aimed to interact with any other existing application, local or remote. Overall design and programming was carried out by the senior author, with additional routines by the junior author, a graduate student in Anthropology at the Universidad de Buenos Aires. Funds were granted by a CONICET program of research and development (PID 401/89).

The graphic interface

VB-GIS has a very rich, ergonomic and customisable interface with menus, iconic toolbars, command buttons and dialog boxes. No unknown file names and no formulas have to be typed or recalled. A complete help system associated with every command or option is available at any point. All menu items have keyboard acceleration traps allowing the interface to be controlled by keyboard or by voice pilot macros using Windows Sound System or any other similar sound kit. This means that complex sequences of commands can be controlled both by voice or by normal Windows macro procedures. The mouse is the natural choice, however, although a pen device or a digitising tablet may be more accurate for certain purposes.

To ensure ease of use the interface is highly adaptive: only logically applicable operations are available at each step. Some forms and objects grow and change as required. If the user needs an additional copy of a complex object, such as a

graphic or spreadsheet, they can create it at will. Several graphics and tabulations of data can thus be compared, edited, globally modified and inter-operated.

The design has carefully avoided the dead-ends and incongruous dialog arrangements so common in much archaeological software. Even the most complex procedures are both open and self-contained: if some operation is selectable, then it can be carried through completely using the available data. The user never has to remember a path or a file name and will never get lost in a sequence of cascading menus leading in the wrong direction. No procedure is initiated if the data is inappropriate.

VB-GIS also includes a systematic set of interpretive aids to deal with statistical procedures and results, including some basic statistics tables (such as chi square and typical values for t) for comparison with actual outcomes. Another primary design objective was to make the more complex statistics understandable by the non-specialist.

The program was also developed to be complementary to other software alternatives, no matter the degree of functional overlap. The system is compatible with most data formats: eg. dBase, Excel, Lotus, Quattro, Paradox, Btrieve, FoxPro, Oracle, SQL Server (through ODBC), Surfer, Arcospace, MapInfo, text files etc. No conversions are ever required and no redundancy occurs. The Open Data Base driver allows any database server to be used without recompiling the system. The data can be transferred to the internal matrices in several ways, from clipboard copy-and-paste to direct import.

Perhaps one of the program's most valuable features is the open-ended nature of its interaction with the user, who can choose between a wide range of options for data input and representation style. For example, text datasets can be automatically generated by an analogue pointing process, individual data items can be dragged and dropped on to a grid of squares or they can be generated by more or less constrained random algorithms. The data can be edited directly on one of the various spreadsheets provided, imported from data files, passed through the clipboard or abstracted from a complex OLE Automation object.

Graphic facilities

In the area of graphics, VB-GIS has a full featured bitmap toolbox which permits internal map design. A special graphics library includes icons, vectors and bitmaps for representing objects related to sites, non-sites, trenches and grids, including scale rules, cardinal pointers and archaeological symbols. The system accepts scanned pictures in PCX, BMP, RLE, GIF, TIF, TGA, JPEG and WMF formats. 3D vector images in AutoCAD DXF format can be imported and activated for further rotations and transforms. Additional vector and raster formats will be added in the next release. The graphic interface provides graphic tools for marking regions, with separate zooming, rotation and stretching. It is possible to work on several maps at the same time, to select and crop a region from a larger map or to tile partial maps or satellite views in a composite image.

Individual artefact types are normally represented by means of simple figures such as circles, stars, diamonds, squares, triangles and crosses of different colours. A configuration screen allows editing of a special table providing user-defined symbols and associated legends. Both nominal and numeric values can be used later in database queries.

Graphic processing capabilities are very extensive, including facilities such as: translucent maps and data; addition, abstraction or substraction of image elements; image editing optionally under the control of other specialised programs; combination of over 13 different drawing modes and 8 fill styles; analytical mapping from conceptual data; layer integration and segregation; perspective mapping and so on.

An important section of the program is a photographic laboratory utility with several options for controlling brightness, gamma and colour contrast, border smoothing and sharpening, blurring, masking, posterising and filtering of maps, plots and aerial snapshots. Colour depth is selectable at any time from 1 to 24 bits, specifying dithering and palette options. Palettes themselves may be borrowed from other images and animated. The laboratory section is as technically advanced as any photo editing program available on the market, and it is certainly also faster, even when working with more than 16 million colours. This module also allows conversion of any bitmap format to any other, including the highly compressed JPEG (Joint Photographic Experts

Group) standard. The degree of lossiness of JFIF-JPEG images can be varied by adjusting compression parameters. The internal photo lab uses a powerful Visual Basic library, Image Knife/VBX, from Media Architects of Portland, Oregon.

VB-GIS features second generation Object Linking and Embedding technology, both as OLE 2 object editing and OLE Automation. This means that data types incorporated in maps and data bases may include video clips, sound files, rasterised photos, vector data, spreadsheets, word processor documents, statistical graphs, flow charts and whatever can be handled by the installed Windows platform. Our program systematically brings together, and manages all the editing and maintenance power scattered among other installed applications.

VB-GIS also allows for automatic generation of papers and reports, with publication quality output of text, tables and graphics. Documents like this one with additional pictures inserted, for instance, may be edited using a standard word processor, passing all the graphic items on line from VB-GIS, without intermediate files. Conversely, a complex document can be a compound internal object.

Several special PostScript facilities for printing and plotting have been designed specifically for VB-GIS, including greyscales and texture support. The program should support any output device installed for Windows without special setup: all the capabilities of the Windows drivers regarding fonts, page size, orientation, resolution, dithering, etc. are automatically available.

Data editing

VB-GIS has a compound MDI (Multiple Document Interface) form specific for data editing and analysis. In fact, this MDI provides two fully configurable spreadsheets and two statistical graphs on which to plot. Graphs and spreadsheets can be copied at any time, even when they are empty: so the real scope of the interface is undefined, bounded only by the available memory and resources. The spreadsheets, which use the extraordinary Spread/VBX by Far Point totally embedded, hold up to 2 *billion* columns and the same number of rows each: a lot more than any conventional spreadsheet, including the forthcoming version of Excel.

Within the spreadsheets, the editing tools are correspondingly sophisticated: the user can drag and drop, make multiple selections, swap and transpose, sort, hide and display rows/columns, import and export any kind of data, pass a range to and from another program through the clipboard and plot several kinds of statistical graphics with simple commands from the menus, keyboard or mouse.

The spreadsheets also behave like complex documents, not necessarily in tabular form. Every cell of the spreadsheet is individually configurable in terms of font, colour, pattern, alignment, border and size. The user may import data in any format, with any number of fields, and adjust the records to the system's expected scheme later, using the available editing facilities.

Statistics and analytical functions

Once procedural specifications are clear, mapping from the mathematics into computer language is boring but easy. Maths is what computers do best. VB-GIS is able to deal with both intrasite and intersite data. When establishing the data, the archaeologist is free to choose between quadrats, regions, transects or exact coordinates, working with a maximum of 32,767 elements, types or layers (depending on the context). This particular version is more effective in intrasite analysis, but a set of intersite capabilities is in the pipeline. The main problem with intersite analysis is the relative lack of effective archaeological algorithms in published papers.

The goal of VB-GIS's analytical function is to set *all* the common archaeological approaches ever published, together with their documentation, critiques and shortcomings, in an environment suitable for cross-comparison. No single package contains so many archaeological methods, backed up by their original supporting documentation in the form of a dense hypertext help document. VB-GIS is thus a purely archaeological tool, not a more-or-less opportune number cruncher stolen from an unrelated field of research.

VB-GIS offers a collection of descriptive statistics, associational statistics, multivariate statistics, nearest-neighbour and spatial auto-correlation statistics, following the methodological ordering proposed by Hietala. Some of the 2D statistical facilities already implemented include:

Reynoso

1   Basic Calculations: Distance, Angle, Latitude-Longitude and Area. Both imperial and metric scaling are available;

2   Simple Overall Density;

3   Density by Artefact Type;

4   Spatial Scaling for determining the most effective grid partition;

5   Poisson Distribution and Chi Square for matching actual distributions against random or regular sets of points;

6   Carr's Polythetic Association;

7   Whallon's Unconstrained Clustering. This is in fact a series of calculations, plots and transforms;

8   Hodder and Okell's A Coefficient for assessing the association between distributions of points. This method is not scale dependent and provides significant information about the distributions;

9   Johnson's Local Density Analysis - used for cell frequency recording and analysis of artefact distributions;

10   Luton and Brown's Contiguity Analysis. A statistical method for assessing the form of arrangement that operates on grid-cell count data;

11   Point-to-item Distance (Pielou). This approach provides a single measure and test of the nature of the arrangement of items within a scatter, considering most of the scales of ordering of items;

12   Contingency tables (Bray, Driver and V Coefficients, plus 30 other indexes and coefficients). This applies to more than one type of artefact;

13   Nearest neighbour of Order N;

14   TNND or Total nearest-neighbour distance (Donnelly) with associated simulation to determine the variance and edge effect;

15   Standard Nearest Neighbour Analysis, with plenty of different authorial and theoretical flavours;

16   Linear Nearest Neighbour (Stark and Young). This statistical method facilitates decisions about whether points along a line are clustered, random or dispersed;

17   Morisita Index. This method uses the dimensional strategy of Dimensional Analysis of Variance, but not the Poisson distribution. The index constitutes a relative measure for comparing arrangement between study units;

18   Cluster Analysis: Cutoff Ratio Maps. Method for delimiting clusters of artefacts of one type using item point location;

19   Cluster Analysis: Kintigh and Ammerman's Heuristics. This procedure uses several alternative Artificial Intelligence techniques, including a Prolog Dynamic Link Library;

20   Cluster Analysis: Whallon's Overlapping Cluster approach. A method for defining the degree of co-arrangement of artefact types, using the point location of items;

21   Dimensional Analysis of Variance or Dimanova, also called variance-mean ratio;

22   Kriging (Zubrow and Harbaugh). This is a method of interpolation employing moving averages, which involves estimation of the values of a spatially distributed variable as well as assessment of the error associated with the estimates;

23   Sampling statistical suite. A set of different sampling calculations;

24   Regression Analysis;

25   MRPP (Multi-response permutation procedures) for analysis of artefact distributions from Berry, Mielke and Kvamme. This is analogous to one-way multivariate analysis of variance, but uses the average between-artefact distance among objects within classes as their primary analysis unit;

26   Simulation: Random Walk;

27   Simulation: Gravity Model;

28   Contouring, Smoothing and Trend Surfaces.

A host of standard simple statistics (mean, average, maximum, minimum, standard deviation, best fit, etc.) are always accessible from the graphics and spreadsheet windows. Furthermore, almost any statistical operation can be developed step by step using formulas in the spreadsheets. The search for new procedures and algorithms for inclusion in VB-GIS is still open. More than 50 additional statistical routines are planned for implementation in Release 1.0.
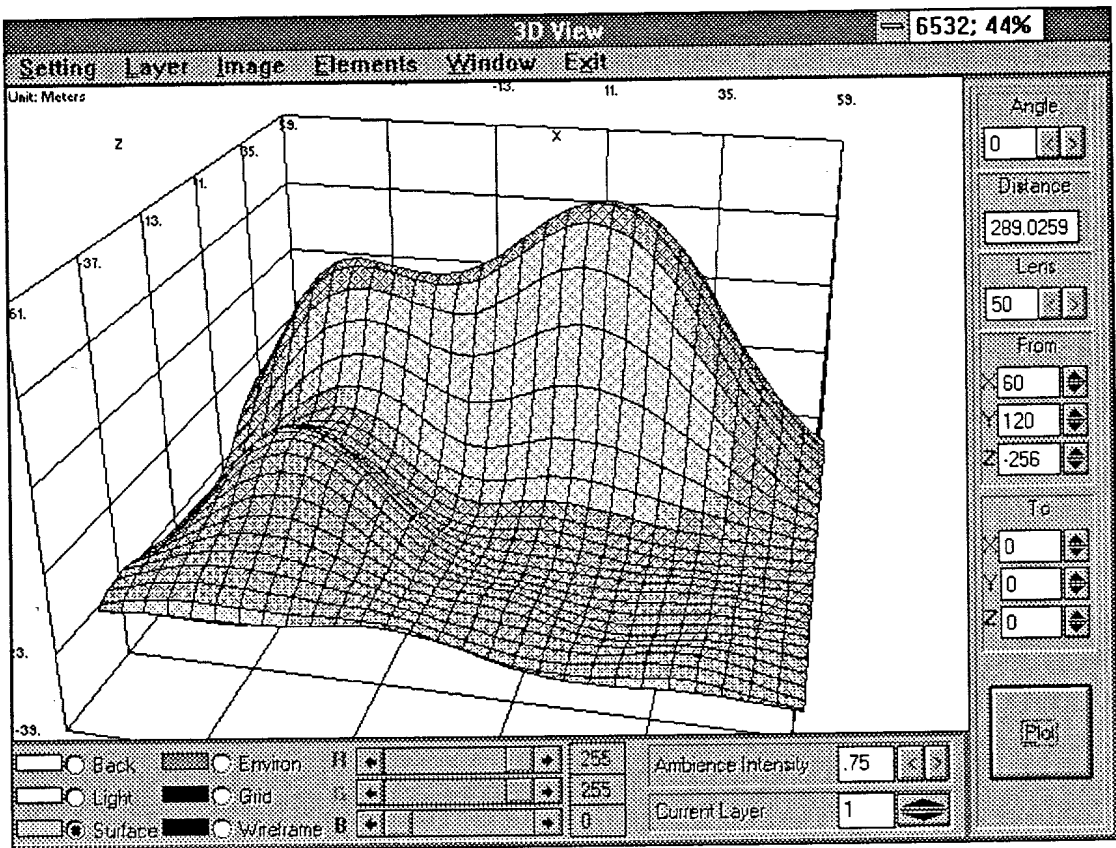
Fig. 2: 3D screen displaying a simple spline surface.

a default setting for mapping colours onto altitude ranges.

VB-GIS is able to control four different three-dimensional surfaces at a time (hundreds of overlapping surfaces may be plotted when running on Windows NT). A surface can denote anything: a roof, a topographical terrain, artefact diversity or all of these things together. Any mathematical function which renders a 3D surface can be plotted in one step. A set of 3D basic figures (spheres, cylinders, cones, cubes, toroids) is also at hand. The graphical plotting of 3D objects supports up to 16 million colours with 32 bit boards.

Given a three-dimensional object, VB-GIS uses the 'camera and lens' metaphor to situate the viewpoint. There are separate controls for object orientation, scale and placement, with options between parallel and perspective projection. Hidden surface removal is automatic (it's possible to turn it off, however), and the transparency index can be changed to allow both wireframe or solid surfaces.

One of the program's most evolved capabilities concerns lighting: light can be of any colour and intensity, and several (different) light sources can be placed at any point in space. One can simulate light hazing, darkening distant objects, or reproduce specific light conditions, e.g. an underwater scenario or sunset. Shading (flat or Phong) is also selectable. Light, surface colour, shades and refraction values interact with texture settings, such as rock, wood and bumps to give realistic landscapes. Additional textures may be used to identify a bitmapped image file which will be projected onto the facets of a given object.

VB-GIS has a repertory of interpolation methods: Bspline, Hermite, Catmullrom, Bezier and Beta surfaces are selectable, with further partitioning and subdivision. Smoothing algorithms and plotting procedures are rather fast compared with common three-dimensional plotting. Particularly impressive images can be obtained using Zbuffer procedures - the calculation involved is fairly slow, but the image quality is excellent by any standard.

Databases and information systems

Last but not least, VB-GIS has a complete database and information system embedded, distinct from the tabular data itself. This system operates in relation to both artefacts and coordinate locations, and the information corresponding with each item may be of any type: photos, presentations, descriptive data, documents, matrices, video, sound or movie clips, vectors, compound objects and even applications with specifically loaded data. Internally, all complex data is represented as a reference to an object which may link with other objects.

There is no limit to the number or size of data bases associated with an archaeological set. Once a database has been defined, its content can be selectively mapped by means of SQL (Structured Query Language) filtering criteria. These criteria, in turn, may be quantitative (e.g. artefacts heavier than 1 kg), qualitative (e.g. objects coloured red) or mixed. Old data need not be re-entered: any existing database can be linked as a VB-GIS information set, provided that coordinate data are included or added to the data.

The future of VB-GIS

Generally speaking, there are three main categories of GIS programs available today:

1   *Information systems*, usually related to more or less proprietary (i.e. closed) database engines: MapInfo, the best seller of this class, lacks any 3D plotting option and its statistical aptitude is quite rudimentary. It started out as a commercial program for executives and, in spite of an internal 'Map Basic' language implementation, it shows;

2   *2D and 3D plotting utilities*: The best examples are Idrisi & Surfer. These also lack statistical sophistication, offer poor intrasite analysis capabilities and no database integration. They are not (yet) Windows applications. For plotting with a new printing device in Idrisi the user has to capture the screen, losing the palette information. Strictly speaking, Idrisi & Surfer are not GIS programs, but terrain modelling systems. And in this class there are far better choices, such as the AVS Network Editor from Advanced Visual Systems;

3   *Analytical systems*, like ArcoSpace, well oriented towards archaeology but with poor integration and no concurrent information database.

In contrast to these non-integrated tools, VB-GIS is an integration tool and an open system specifically designed for archaeology. However, in its current incarnation our program should be considered only as a prototype. Even Windows itself is now evolving into the so called Chicago, a thoroughly object-oriented operating system, and Visual Basic is following the same path.

An operating release of VB-GIS will be ready by August 1994. It will not be expensive (less than US$200). Further revisions of VB-GIS will include features and capabilities currently under development. Some of them are worth mentioning in this context:

- Automatic support for extended video modes (800x600, 1024x768, 1280x1024). Current version works only at 640x480 pixel resolution for speed in design mode;

- New special mapping controls (Map Control and Map Server) designed by GeoSoft Limited (Leeds, U.K.). These allow import, display, manipulation, query and hardcopy of colour maps;

- Fractal geometry for texture design and 3D randomisation;

- Bitmap pattern recognition tools and edge detection algorithms. Sobel transforms (also called convolution), smoothed or Hough transforms, Laplace transforms, isotropic and stochastic transforms. Low contrast photos will be transformed to more or less accurate line maps;

- Improved 3D graphic tools from Mark Owen;

- 32 bit computing for extremely large arrays. Large matrices should be handled easily;

- Rushmore algorithm for querying databases. This is the fastest retrieving method so far, and it will be included with Access 2.0 database engine (embedded in Visual Basic 4).

- Thiessen polygons and Delaunay triangles for area estimation;

- NeuroWindows neural network and fuzzy logic library for pattern and cluster recognition from Ward Systems Group (Maryland, USA);

Reynoso

- Embedded scanner management through TWAIN standard interface.

Most of these improvements arise from the natural development of the libraries already included.

Some archaeologists have already suggested interfaces, filters, corrections and new procedures. Foss Leach (Archaeozoology Laboratory, Museum of New Zealand at Wellington) asked for a specific 3D interface in terms of Euler angles and requested compatibility with DTE and DTM files. Ian Johnson (University of Sydney) provided detailed information to improve programming of Local Density Analysis. At the Blue Mountains Conference, Zoran Stancic (University of Ljubljana) gave us helpful hints for the inclusion of imports from Idrisi formats. We still have to talk with Sebastian Rahtz (University of Southampton) and other members of his team who are programming SyGraf, also in Visual Basic. Additional inspiration will be elicited at the next CAA meeting in Glasgow. I would particularly like to thank Ian Johnson & Zoran Stancic for their valuable corrections, comments and ideas.

VB-GIS is planned as an open system without a definite ending. It is only about 15% of the maximum recommended program size. There is enough room to include better libraries, to make the system tighter and to obtain more optimal results. Meanwhile, any ideas and feedback will be most welcome.