SOME PROBLEMS IN ARCHAEOLOGICAL EXCAVATION 3D MODELLING

By Juan A. Barceló and Oriol Vicente
UNIVERSITAT AUTONOMA DE BARCELONA. Spain
jbarcelo@seneca.uab.es
oriol.Vicente@uab.es

ABSTRACT

Photogrammetry has been used in archaeology for the recording of complex structures. Therefore, the use of those techniques is more frequent in medieval and classic archaeologies. In those cases the presence of buildings and architectonics makes profitable going beyond the traditional paper and pencil approach. For prehistoric sites, traditional hand drawing is still the most profitable approach. Elements and structures to be drawn are geometrically quite simple, and the problems are more on the complexity of stratigraphic and sedimentary aspects, than in buildings, walls or floors. In this paper, we explain how to use digital photography in prehistoric excavations, how to modify those pictures to adequately represent the archaeological record, and how to build geometric models from photographs. Our final goal is to build a geometric and dynamic model of the site, in order to explain not its architectonic complexity, but taphonomy and the site formation process.

Archaeological Excavation and the definition of Archaeological Space

We can define archaeological space as a sequence of finite states of a temporal trajectory, where an entity (ground surface) is modified successively, by accumulating things on it, by deforming a previous accumulation (for instance, by spreading) or by direct physical modification (building, excavation). Archaeological sites should be considered as the result of successive and overlapping modification steps. (Barceló et al. 2003).

Natural and human process modify physical space, and as a result we are able to distinguish *phase* or modification steps, which can be used as *analytical units*. A *phase* is a homogenous region in space delimited by a well-defined discontinuity or boundary. A well defined boundary is an abrupt change in some spatial values. We may define an interfacial boundary or interface, when two phases are in mutual contact, that is, when two neighbouring regions in space have different probabilities for the same formation processes.

Spatial discontinuities have three main properties: geometry, topology and texture:

- a pattern of discontinuities in boundary orientation (curvature), that is, *shape*. It is defined as the information that is invariant under translations, rotations and isotropic rescaling, that is, those aspects of the data that remain after location and scale (size) information are discounted. It is then a quantitative property about spatial location and size. Phases are spatial units, and consequently they have *size* and location, whose relationship can be examined in terms of *shape*. Shape is a field for physical exploration: it has not only aesthetic qualities, nor is shape just a pattern of recognition. Shape also is determining the spatial and thus the material and the physical qualities of archaeological site components.

- a pattern of discontinuities between boundaries at different spatial positions, that is *topology.*
- a pattern of discontinuities in luminance variations in a scene with non-uniform reflectance, that is *texture*. It is the name we give to these variations, which seem to be usually caused as a result of the process that created the boundary discontinuity. Texture can be also seen as the definition of spatial attributes having either visual or actual variety, and defining the appearance of the observed area. Any point of the ground surface has variations in its local properties like albedo and colour variations, uniformity, density, coarseness, roughness, regularity, linearity, directionality, direction, frequency, hardness, brightness, bumpiness, specularity, reflectivity and transparency (Barceló et al. 2001, Adán et al. 2003). All these perceived qualities or attributes of spatial locations within archaeological space play an important role in describing the sources of irregularity and surface variation which are responsible of specific textures. Texture then, may be defined as the local variation of brightness from one pixel to the next or within a small region, where the brightness of a point is a function of the brightness and location of the light source combined with the orientation and nature of the surface being viewed. If the brightness is interpreted as elevation in a representation of the image as a surface, then the texture is a measure of the surface roughness.

A wall, a pit, a garbage accumulation are *phases* or distinct regions of the archaeological space, which can be defined not only in terms of their own properties, but also in terms of the differences with neighbouring phases. We need to distinguish where observed discontinuities or boundaries begin and end, that is what are the proper borders of an occupation floor, or the original shape of a pit, where pottery sherds are accumulated, or where an animal carcass has been broken into bones. Therefore, archaeological excavation cannot be reduced to the mere unearthing of artefacts and ruins, but an exhaustive documentation of an archaeological space in terms of a finite set of spatial variables. The purpose is to characterize observed discontinuities in terms of distinct components or relevant units with uniform value of shape, size, texture, composition. However, a phase cannot be defined only in terms of their boundaries (Barceló et al. 2003). They should be analyzed as the presence/absence of some qualitative spatial variable, that is a feature which has positive value if it is present, and negative value in case of absence. Observed discontinuities between phases can also be expressed in terms of quantitative variables. Quantitative variables exhibit a variation in value throughout spatial regions. Variables such as geomechanical properties, mineral grades, material accumulations, soil morphological features, or any other property of sedimentary/depositional units and archaeological contexts, can be sampled or measured in terms of real, numerical values.

Our objective is then to analyse how qualitative and quantitative variables “vary significantly from one location to another”. Formation process and depositation effects appear in some locations and not in other because of their position relative to some other location for another process or a reproduction of the same process. A visual model then pretends to examine if the characteristics in one location (for instance a wall, or an activity area) have anything to do with characteristics in a neighbour location (for instance an accumulation of pottery or lithics, or bones) through the definition of a general model of spatial dependence. In other words, the main objective of visual model is the spatial analysis of phase correlation: how distinct formation process have influence over spatio-temporal discontinuities observed through the site. What we are looking is whether what happens in one location (temporal or spatial) is the *cause* of what happens in neighbouring locations. One possible effect of spatial causality is the similarity of values in neighbouring locations, but this is not the only effect. Obviously, we should not limit ourselves to the analysis of “spatial similarity” relationships, but all effects *probabilistically* related to the spatial or temporal location of the cause.

Archaeological Information Sources: working with variables and coordinates

Archaeological information sources can be reduced to three basic data types: variables, characteristics and coordinates. A regionalized variable exhibits variation in value throughout a (theoretically) indeterminate region. They are properties of the subsurface that can be samples and measured in terms of real, numerical values. In contrast, characteristics are observable qualities of the archaeological space that have a finite number of possible descriptive values, and uniform value within finite, irregular volumes. Characteristic values are associated with discrete archaeological areas with distinct boundaries (a wall, an occupation floor, a pit, etc.). Their importance to subsurface characterization lies in the fact that they frequently influence the spatial variation of regionalized variables (Houlding 2000:16ff).

The common feature of the archaeological information sources is that every variable or characteristic value is associated with a location and an extent (point, line, area, surface, volume) which in turn are defined by an implicit data geometry. We use scalar fields to represent this geometric structure.

A scalar value is a single component that can assume one of a range of values. Example of this are texture (roughness, porosity, etc.) and composition (frequency of artefacts). A *scalar field* is a name we give to a function who take in points in a two or three dimensional space ($\mathbf{R}^2$ or $\mathbf{R}^3$) and outputs real numbers. A scalar field is an arrangement of scalar values distributed in a space. Archaeological spatial components can be characterized in terms of scalar fields. The scalar field is a concept spawn from the natural and physical sciences since they often deal with a region of physical space with a function attached to it. For example, the function that gives the temperature of any point in the room you are sitting is a scalar field. In an archaeological case, the function that gives the quantity of rabbit bones at any point of the site is a scalar field. However, a function doesn't need to expressed and defined as a mathematical formula for it to be an explicit function. Just the input-output correspondence. So particular scalar field may be specified by a mathematical expression, or it may be a function whose value at any point could be obtained by physical measurement (during excavations).

As scalar field data, the archaeological site should be specified by a multidimensional array of points instead of a set of delimited objects (walls, floors, pits, stones, etc.). The underlying mathematical definition of such a model is a set of scalar fields that define the geometrical and physical properties of every point *p* in three-dimensional space. Each point in field data has the following:

- The location in 1D, 2D, 3D or 4D (space plus time) where this point is located
- One or more data values (variables and/or characteristics) associated with this point

For example, at a specific location with spatial coordinates (x_i , y_i , z_i), we have measured the values of two quantitative variables (quantity, density and some archaeological material) and the presence or absence of some archaeological characteristic (a wall of mud bricks). Quantity, density, and type-identification are the information which we have to collect at that specific point in 3D space.

Therefore, scalar field archaeological data have the following four dimensional generic format:

$$W(t,x,y,z)$$

Note that the model being suggested here is not a standard shape model. That is, the spatial variable to be analysed is not the height of the ground, but how different four-dimensionally located points have different properties. In this example, there are four dimensions (*x, y, z,* time). The first 3 dimensions are spatial: rows, columns, and levels (or latitude, longitude, and height). The 4th dimension is time. *W* represents possibly many functions w_1 , w_2 ,…w_n . Each *w* corresponds to a dependent variable, and can be used for quantitative variables (sediment hardness, porosity, degree of consolidation, density, porosity, cohesion, strength, and elasticity) or for qualitative characteristics (presence/absence of specific built structures). Then we consider a series of different W_i values at a position in the array defined by time and the three standard spatial coordinates.

The usual way to define dependent variables in a four dimensional model is by using observable characteristics. For instance, consider a texture classification or material identifier, with the following values:

Limestone, quartz, granite, pottery, charcoal, clay,…

Or

Wall, occupation floor, brown sediment, red sediment, black sediment,…

Or

Wall1, wall 2, wall 3, Floor 1, Floor 2, Floor 3, Pit 1, Pit 2,…

Although archaeology deals necessary with *time*, this is one of the less tractable independent variables. Even in the case of C14 dates, we do not have scalar values, but irregular intervals, which constitute mathematically complex structures. The easy way to solve this problem is by using ordinal values:

Period 1, Period 2, Period 3

Note that we are not equating stratigraphic *depth* with *time*. In some cases, the temporal evolution of the site can be correlated with stratigraphic ordination, but not necessary.

In addition to the data itself, there are a number of attributes needed to describe a five dimensional scalar field: the sizes of the five dimensions (number of rows, columns, levels, time steps, and variables), geographic position and orientation of the data (map projection), the names of the variables, the actual times and dates associated with each time step, etc.

Scalar field visualization is the graphical expression of relationships between scalar values distributed in space. To adequately visualize these complex data structures we must consider a semi-infinite continuum made up of discrete, irregular, discontinuous volumes which in turn control the spatial variation of archaeological features. Therefore, an archaeological site should be described in terms of a volumetric information, that is, a group of data that describe a solid object from a three-dimensional space. Volumetric data occupies a volume of space. For example, when data is collected by excavation, there are data points spanning the height, width, and thickness of the archaeological element and a data value representing the type of material, sediment, structure, bone accumulation, etc. at each point.

Volumetric Data Acquisition

With scalar field data, the actual location in space of each point in the field must be supplied. This is made during fieldwork: coordinates are taken using topographic equipment, but there is not an easy way to acquire simultaneously quantitative variables, archaeological characteristics and coordinates.

The usual way is to acquire separately coordinates, and after excavation, characteristic regions are built by joining characteristic points with lines, areas or surfaces. The archaeologist has an external database with a qualitative description of information related to each point, and then is able to create geometric shapes and structure by linking already existing points. by using polygons connecting points, by interpolating parametric surfaces or volumetric primitives. For instance, an occupation floor is acquired by calculating the 3D coordinates of some characteristic points along the contour, and then fitting a polygon to those coordinates.

The problem with this approach is that it is a modelization process, and not a real sampling procedure. Archaeologist is projecting what he/she things to know about the nature of a finite and very reduced series of coordinates, and focusing only on specific characteristics. Walls, and built structures can be easily built using standard geometric fitting tools, but then we are forgetting all information about interfacial boundaries. A wall is a characteristic of the archaeological space, but the sediment covering the wall, or the accumulation of stones around it, are also important characteristics which should be sampled.

We suggest to use a different approach to sample archaeological data. We need coordinates and characteristics to be acquired simultaneously, and we need a huge quantity of data points to approximate to scalar fields. The only way is through photographs.

A photograph is a spatial pattern of different luminance values. It is not a surrogate for reality, but a device for capturing some initial input (luminance perception) which should be translated into observed data. Given that spatial discontinuities are the building blocks of archaeological discontinuities, and they can be analysed in terms of texture variation, photographs can be used as a model of texture discontinuities.

Of course, we need more than a single photograph. 3D data sampling presupposes that a series of cross-sectional images, representing some volume which was regularly sampled at some constant interval, exists in digital form. An *image stack* is a display of multiple spatially or temporally related images in a single window. The images that make up a stack are called slices (Fig. 1). Stacks can be viewed from different perspectives, treating the layers of the stack as another spatial dimension.

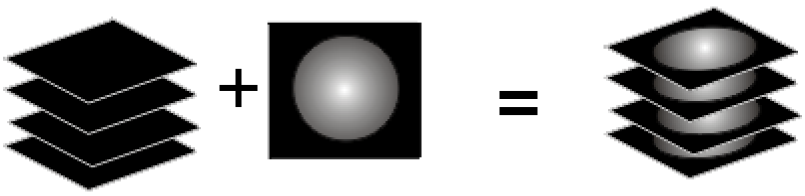


Fig. 1. Sampling a 3D reality using photography-slices

This works only if certain conditions are satisfied. First, the individual files must be in the same format and bit depth. Second, they must have similar structure, i.e., same numeric type, number and sizes of dimensions, etc (Fig. 2).

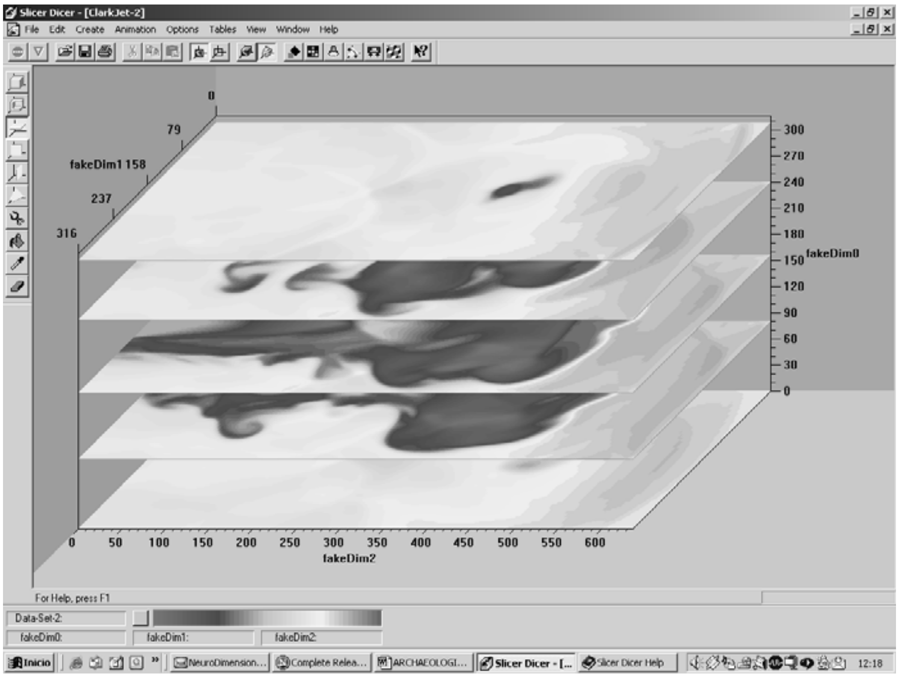


Fig.2 A stack of images sampled at different intervals. Simulated data using the SlicerDicer software from Pixotec,Inc.

Each image or *slice* in a given dataset is made up of a number of picture elements or pixels. The distance between any two consecutive pixel centres in any slice within a dataset represents a real world distance referred to as the interpixel distance. Similarly, the distance between any two consecutive slices represents some constant real world depth with which the volume was sampled. This constant depth is referred to as the interpixel distance.

A series of cross-sectional digital images of this type is referred to as a *volumetric dataset* or simply as a dataset. Such a data set is represented by a series of photographs, each containing a similar n-dimensional data array. Collectively, these files are interpreted as a single array of n+1 dimensions. For instance, we have 5 image files, each containing a 100 x 200 x 200 data array. By opening them all at once, a computer can read these data as if they come from a single 5 x 100 x 200 x 200 array, where 5 is the number of slices.

Processing a volumetric dataset begins by stacking the slices of a given dataset in computer memory according to the *interpixel* and *interslice distances* so that the data exists in a "virtual" coordinate space which accurately reflects the real world dimensions of the originally sampled volume. The next step is to create additional slices to be inserted between the dataset's actual slices so that the entire volume, as it exists in computer memory, is represented as *one solid block of data*. The number of slices needed to fill in the blanks is based on the dataset's interpixel and interslice spacing and the slices needed are created through interpolation.

In this case, spatial values should be defined on regular, rectangular grids. Such data take the form of n-dimensional Cartesian arrays. The illustration below depicts a regular, rectangular grid (Fig. 3) .

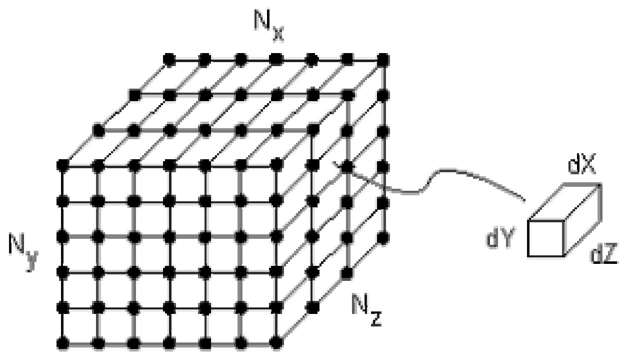


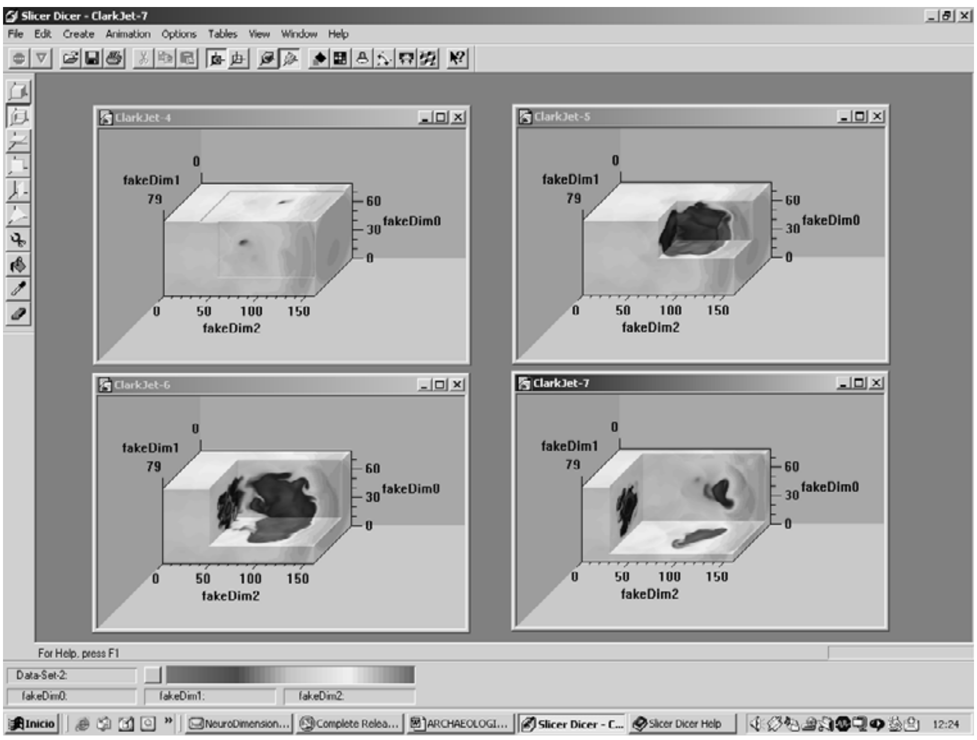
Fig. 3. A 3Dimensional Coordinate Grid, and elements for voxel definition

The numbers, *Nx*, *Ny*, and *Nz*, defining the size of the grid, are arbitrary, as are the dimensions, *dx*, *dy*, and *dz*, of each grid element. Note that although a grid element can have any shape (*dx*, *dy*, and *dz* can be unequal), grid regularity requires that all elements within the grid be identical in size, i.e., *dx* is the same for all elements, *dy* is the same for all elements, etc.

Once a dataset exists in computer memory as a solid block of data, the pixels in each slice take on an additional dimension. In effect, the pixels become volume pixels or *voxels.* Once loaded into memory, a volume can be translated and rotated and a *rendering* of the dataset can be obtained.

Rather than visualize a single data set, we want to go a step farther and explore the interrelationships between two or more scalar fields. This can be done, at the expense of visual complexity, by tagging data values with multiple independent attributes.

Fig 4. A rendered volumetric data set. Simulated data using the SlicerDicer software from Pixotec,Inc.



Note that archaeological data are always four dimensions (t, x, y, z), but the view coordinate system can be at most three-dimensional. (We live in a three-dimensional world, when it comes to physical phenomena and our visualizations of them.) So, for visualization purposes, we need to decide which dimensions will be treated "spatially," as coordinates, and which will be treated parametrically.

In this example, we can view the data in the 3D volume W(3, x, y, z) (i.e., at t = 3). Or, we could view W(t, 64, y, z), or W(3, 64, y, z). The latter represents a 2D subvolume of the data array. In general, two or three of the available dimensions must be selected as view coordinates. The remaining dimensions are treated as parameters and must be assigned specific values prior to visualization.

Technical considerations before building an image stack

If we need to overly photographs to *slice* a 3D reality (the archaeological site), it is necessary that all slices be spatially related. Of course, sharing attributes on sample points only works if all scalar fields were sampled at the same locations. To make comparable all photographs in the stack, we have to geo-referentiate them. This step can be done with photogrammetrical methods, where a image is modified introducing the real coordinates, and deforming it to be adjusted to a real scale.

> *Photogrammetry is the art, science, and technology of obtaining reliable information about physical objects and the environment through the processes of recording, measuring, and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena.*
> *From the Manual of Photogrammetry, Fourth Edition, published in 1980 by the ASPRS.*

Georeferenciation implies that discontinuities observed in the photographs should be positioned in space.

The *x ,y* and *z* axes of the representation are aligned with the east, north and azimuth axes of the projection being used. Each pixel in the photo and each geometric primitive in the vectorial representation is geo-coded with its associated coordinate in the map projection. We introduce in the model a new variable: location. Some control points should be measured independently and then transferred to a database with a reference to the photo. Those control points will be used to *register* the image and substitute pixels using specific geo-registration algorithms.

Also, the georeferenciation implies the rectification of perceptual errors, focal angle or focal distortion produced by the position of the archaeologist. In the case of a photograph a preliminary rectification transforms the pixels of the photo by interpolation of luminance wave length according to *x* and *y* axes.

However, even after geo-rectification photographs are always a misleading representation of reality. Scale variations are caused by the natural point-to-point variations in the elevation of the terrain being photographed. Scale variations are also caused by the varying distances of objects out from the principle point of the camera, as it is a perspective projection. Photographs are a 2D contrast map of luminance reflections on a 3D real surface. In other words, a 2D photograph is a deformed representation of reality. If it is deformed, a sequence of images within a stack would not be a right volumetric data set, because the spatial distribution of variables and archaeological characteristics sampled at one layer has nothing to do with data sampled at other layers.

Consequently, slices within an image stack cannot be considered a true input of volumetric data if we do not remove scale variations from any image. Once these variations in scale are removed from a photo, the photo becomes a true image map of the ground, where “map” is defined as a constant scale representation of a portion of the Earth’s surface.

Limitations of this approach

However, in most archaeological cases we cannot use this procedure. It would be correct if archaeological scalar fields be described in terms of a three-dimensional function

$$W(x,y,z)$$

Where *z* is sampled at specific points using interslice distance, and each slice contains a sampling of x, y, w_i.

There is a temptation to consider the process of building an image stack as a suggestion for excavating imposing artificial layers. Nevertheless, modern archaeology abhors the use of artificial layer slicing, because it makes impossible the correct reconstruction of stratigraphic sequence. We always follow observable discontinuities, that is, what we often call *natural layers*. As a result, slices in an archaeological volumetric data set are not planes, but complex data arrays.

We should remember that archaeological characteristics and all variables describing the archaeological space are intrinsically four dimensional. Stratigraphic sequence is, in fact, a measure of temporal modification, which should be always taken into account.

Only when considering a stack of 4D data arrays we can represent properly archaeological of interfacial phases (contact surfaces between spatial discontinuities) at different time steps. For instance, Fig.5 is a 4D representation of an archaeological time step. Here colour (grey level) is used to represent different *W* values at different x,y,z coordinates. Let W' be the value at a position in the array defined by t = t', x=x', y = y',and z = z'. This datum will be rendered in the data view as a coloured pixel. The colour is defined by a data-to-colour mapping, or colour table, and the position of the pixel in the window is defined by a data-to-view coordinate mapping.

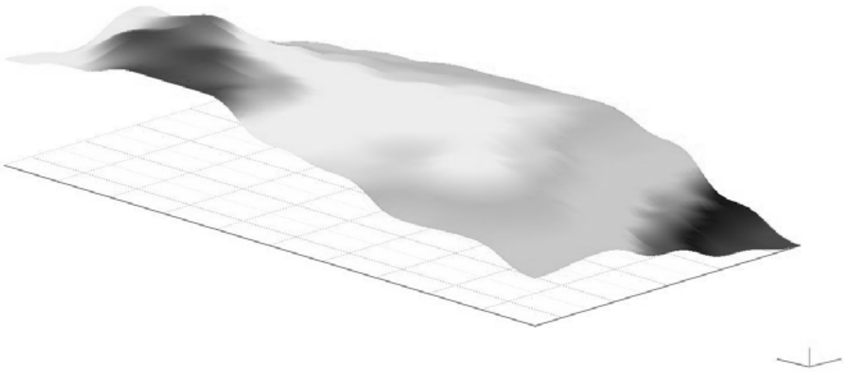


Fig. 5. 4D representation of an archaeological time step

If we had 3 files, each containing, for instance a 10 x 100 x 200 x 200 data array, we can integrate all data into a single volumetric set with 3 x 10 x 100 x 200 x 200 array, where 3 is the number of temporal steps (slices), 10 the number of values of the *W* characteristic, and 100 x 200 x 300 the dimensionality of the 3D grid where spatial values vary .

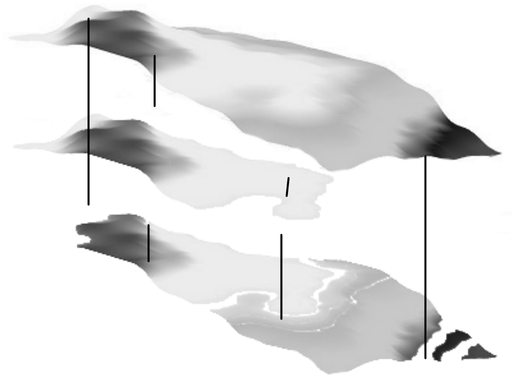


Fig. 6. A 5Dimensional representation of an archaeological site

If we use *z* to represent height of ground level, then image rectification as explained previously does not work. We have here scale variations far greater than when considering only camera placement and focal angle. The only solution is to create a 3D model with textures, were textures can be used to build discontinuities, and then to locate characteristic values according to real topography. In this case, we can excavate following natural layers, and building 4D models (spatial coordinates plus textures) for each time step, identified also in terms of texture or qualitative discontinuity.

Conclusions

When visualizing archaeological spatial data, the exact values of the data are not as important as the relationship between values. Data visualization is used to gain insight into the data set, and expose relationships between values that might not be apparent in the raw data. As a result, intuitive, but less exact, representations of data values are often used.

Regularly gridded data are not very easy to take during field work (see an alternative method in Barceló et al. 2003). The problem is that spatial values defined on unstructured grids, which often take the form of tables with *X-Y-Z-Value* columns, cannot be directly visualized with common volumetric data visualizing software. It is often necessary to resample them onto a regular grid. One way of obtaining this rectangular grid is by creating images stacks.

However many real-world data sets are irregularly sampled. There may be a strong temptation to resample the data into a regular grid. This approach can be problematic in many cases, since the data may not change linearly between grid points. False relationships can be created in the resampled data.

This is the case in archaeological site modelling.

REFERENCES:

ADAN, M., BARCELO, J.A., PIJOAN-LOPEZ, J., PIQUE, R. and TOSELLI, A., 2003. Spatial Statistics in Archaeological Texture Analysis. In Doerr, M. and Sarris, A. (eds.), *The Digital Heritage of Archaeology - Computer Applications and Quantitative methods in Archaeology,* Hellenic Ministry of Culture, Archive of Monuments and Publications.

BARCELO, J.A., PIJOAN-LOPEZ, J. and VICENTE, O., 2001. Image Quantification as Archaeological Description. In Stancic, Z. and Veljanovski, T., *Computing Archaeology for Understanding the Past,* British Archaeological Reports S934, ArcheoPress, Oxford:69-78.

BARCELO, J.A., DE CASTRO, O., TRAVET, D. and VICENTE,O., 2003. A 3D Model of an archaeological excavation. In Doerr, M. and Sarris, A., *The Digital Heritage of Archaeology, Computer Applications and Quantitative methods in Archaeology,* Hellenic Ministry of Culture, Archive of Monuments and Publications.

HOULDING, S.W., 2000. Practical GeoStatistics - Modelling and Spatial Analysis. Springer Verlag, New York.