
First submitted uncorrected version (with author's formatting)

This version is available at: http://eprints.mdx.ac.uk/6588/

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author’s name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: http://eprints.mdx.ac.uk/policies.html#copy
Travails in the Third Dimension: a critical evaluation of 3D geographical visualization

Ifan D H Shepherd
Professor of GeoBusiness, Middlesex University, London

INTRODUCTION
Interactive 3D computer graphics are not new. Almost half a century ago, Ivan Sutherland developed the Sketchpad system (Sutherland, 1963b) which introduced the first interactive CAD-like toolkit that he felt would “generalize nicely to three dimensional drawing” (Sutherland, 1963a, p.138). Just a few years later, he was experimenting with a head-mounted 3D display device (Sutherland, 1969). Even the use of 3D computer graphics to visualise data is far from novel. In the 1980s, for example, a special report (McCormick et al., 1987) revealed a burgeoning use of the third dimension in what was then termed ‘scientific visualization’, and by the end of the decade significant progress had been made in the development of 3D GIS (Raper, 1989).

Interactive 3D computer graphics became more widely available in the 1990s due to the spread of relatively inexpensive graphics display technology, and domestic users of computer games were soon flying, running, driving and shooting their way around extensive, and increasingly realistic 3D worlds [1]. More recently still, augmented reality applications of 3D displays have begun to appear in applications as diverse as telemedicine, travel guides and geocacheing (Güven and Feiner, 2003). In view of this considerable history and widespread contemporary deployment, it might appear to be rather late in the day to reconsider the use of the third dimension for data visualization, and particularly to suggest that there may be troubles in this particular virtual paradise. However, there are several good reasons for taking stock at this point in time.

The first is that we may be in danger of repeating the ‘technology first’ mistake that occurred in the 1980s when colour displays began to displace monochrome displays (Hopkin, 1983). In some quarters, a ‘3D for 3D’s sake’ tendency appears to be repeating the ‘colour for colour’s sake’ trend of a couple of decades ago, among developers and users alike. It is perhaps timely, therefore, to reconsider the circumstances in which 3D is more appropriate than 2D, and when it is not. A second reason for undertaking a reappraisal is that some seasoned researchers -- and many college students -- who use current software to visualise their data do not fully appreciate the principles of effective data visualization, whether it be 1D, 2D, 3D or 4D. The growing popularity of 3D graphics in the media (from films and games to virtual web spaces and satnav gadgets), together with its appearance in applications software and operating systems (and perhaps most notably in the Vista release of
the Windows operating system), suggest that a new era of 3D popularity may be about to dawn. If data analysts are not to be swamped by inappropriate 3D technology, and are to know how to use it effectively, then they need to develop a better understanding of the principles, roles and limitations of 3D data visualization.

A third reason for undertaking a critical review at this time is that, somewhat paradoxically, 3D data visualization techniques have only recently been included in mainstream GIS and desktop mapping systems. Despite the depth of innovation in 3D data visualization over the past quarter of a century, most current practice in geographical data analysis is still rooted in the 2D era, supported by a technology that reveals a significant paper-based legacy. It might therefore be an appropriate time to consider the potential future role of 3D in mainstream GIS and mapping software, and how this might best support effective visual data analysis.

A final reason for the current appraisal lies in the danger that 3D might be seen as a final step in the linear evolution of data visualization technology. It will be argued that data visualization is only part of a broader framework of representational technologies, in which 2D representations continue to provide powerful insight into data. In addition, multi-sensory representation (or perceptualization) not only offers solutions to some of the limitations of 3D data visualization, but it also provides much-needed alternatives for those who are visually impaired.

One further introductory observation is in order, which concerns the scope of this chapter. A great deal of what is termed geovisualization concerns the creation of (photo-realistic) views of actual or proposed world features, usually based on digital spatial data. The current chapter, by contrast, explores ways in which suitable data of any kind, whether spatially referenced or not, may be visualised in 3D scenes. Such scenes may be based on real-world objects or locations (as when thematic layers are draped or otherwise superimposed over terrain models), but they may also consist of recognisable but artificial landscapes, or even more abstract spatial scenes. For this reason, the term geographical data visualization better indicates the concerns of this chapter.

Several broad questions will be posed about the role of the third dimension in data visualization. First, how far have we come in developing effective 3D displays for the analysis of spatial and other data? Second, when is it appropriate to use 3D techniques in visualising data, which 3D techniques are most appropriate for particular applications, and when might 2D approaches be more appropriate? (Indeed, is 3D always better than 2D?) Third, what can we learn from other communities in which 3D graphics and visualization technologies have been developed? And finally, what are the key R&D challenges in making effective use of the third dimension for visualising data across the spatial and related sciences? Answers to these questions will be based on several lines of evidence: the extensive literature on data and information visualization; visual perception research; computer games technology; and the author’s experiments with a prototype 3D data visualization system [2].

**WHAT IS GAINED BY GOING FROM 2D TO 3D?**

There are several good reasons for using the third dimension when visualising data. Some of these are briefly reviewed below.
Additional Display Space

However large the monitor, and however high the screen resolution, users will always reach limits to the amount of data that can be reasonably displayed in their data visualizations. Research in the field of information visualization (or infovis) indicates that the volume of data objects that can be comfortably displayed on screen is considerably larger when using a 3D representation than when using a 2D representation (Card et al., 1991). Where information exists in tabular form, then conventional 2D box displays may be converted to 3D solids to increase the information shown on screen; and where the information involved is hierarchically organised, such as the folders and files stored on a computer, or the documents available in an online repository, then conventional 2D trees may be upgraded to higher capacity 3D cone trees (Robertson et al., 1991). Finally, where data are organised as a network, then 2D connectivity graphs may be replaced by 3D node-and-link graphs (Hendley et al., 1995). A major advantage of this increase in information for the user is that it permits a larger amount of contextual information to be seen while focusing on particular objects of interest. A significant drawback is that 3D visualisations tend to impose often severe interaction demands on users (Shepherd, 2008b).

The conversion of existing 2D data visualization techniques into 3D equivalents is an attractive development strategy, and many conversions have proved successful, as is the case with 3D fish-eye distortion displays (Carpendale et al., 1997), 3D beamtrees (van Ham and van Wijk, 2003), and 3D distribution glyphs (Chlan and Rheingans, 2005). However, other kinds of 2D visualization techniques may not deliver equally significant benefits when converted into the third dimension. For example, although software has been developed to generate 3D versions of parallel coordinates and star glyphs (Fanea et al., 2005), the added complexity of the adapted versions means that the upgrade from 2D to 3D may not provide as much benefit as expected. The pseudo-3D versions of bar and pie graphs which appear increasingly in research publications introduce apparently unnoticed distortions into the messages conveyed to readers (Shepherd, 2008a).

Displaying Additional Data Variables

An immediate gift of the third display axis is that it enables at least one additional data variable to be mapped during visualization. At the simplest level, this means that the height axis may be used for showing the values of any interval-scaled variable. This technique has been used (arguably to mixed effect) in the prism map, which overcomes the inability of being able to apply the size visual variable to the surface area of polygons without distorting the area of real-world features. (2D cartograms solve this problem in another way, but at the expense of introducing spatial distortions to well-known maps that are often confusing to data analysts.) However, it should be noted that the third display axis is not restricted to displaying a single additional data variable, as in the prism map. One way of maximising the potential of the z axis is to map several data variables onto more complex 3D symbols or glyphs. Several examples have been discussed in the literature, and generally fall into two groups. The first group involves the construction of complex glyphs from multiple data variables, which are typically distributed across the x-y plane of the display space. These glyphs are most commonly used in information
and scientific visualization, a representative form being a tree whose branches are sized, angled and coloured to represent particular data variables. The analyst interprets the distribution of such glyphs in much the same way they would explore the distribution of any set of point symbols.

Where data are available for points on an equally spaced grid, an alternative approach has been adopted. In the iconographic display (Lefkovitz, 1991; Pickett and Grinstein, 1988), data values control the display of an array of articulated icons across the data grid, whose varying orientations, thicknesses and colours result in spatial textures that vary visually across the grid. The interpretation of these textures is undertaken differently from the more discrete display of 3D glyphs, because the iconographic display is designed to be perceived preattentively rather than cognitively. Although this display was first devised using 2D icons, later versions have been developed for both 3D and 4D visualizations. For example, the 3D approach has been used to display multiple environmental variables in North America by Healey (1998). As we will show later, however, even imaginative use of multivariate 3D glyphs and icons is insufficient for the visualization of datasets containing large numbers of variables.

Providing a Familiar View of the World

Two-dimensional representations of three-dimensional features in the real world have significant interpretive drawbacks. By flattening topography and using abstract symbols to represent surface features, the 2D paper map imposes a significant learning burden on occasional map users. (In a similar way, the traditional architectural drawing, with its attempt at capturing 3D reality in a multiple view -- plan, side elevation, and front elevation -- divides visual attention and makes spatial integration more difficult.) By contrast, it is often claimed that viewers find it easier to interpret data visualizations that are constructed as ‘natural’ or ‘familiar’ scenes (Gee et al., 1998; Robertson, 1990), which more closely represent the real world. One basis for this argument is that the vertical viewpoint used in the traditional printed map is unnatural for humans who spend most of their lives moving around a world that is seen largely in oblique view from ground level. Another is that the symbolic forms of representation widely adopted on paper maps to represent familiar surface features are too divorced from the everyday experience of buildings, roads or trees.

There are many recent examples where 2.5D and 3D data visualizations have been used to communicate environmental problems and issues to the general public more effectively than conventional 2D maps. The extent and impact of flooding, both in the case of actual floods (e.g. the impact of hurricane Katrina in New Orleans in 2005) and also predicted floods (e.g. the effect of global warming on central London) is often given a greater impact by being visualised using oblique views that show surface variations of the areas involved. However, some caution is needed here, because although Sweet and Ware (2004) suggest that an oblique viewpoint enables better judgements to be made about surface orientation, other interpretive benefits of adding perspective are questioned by Cockburn (2004).

Even where data are spatially referenced, it may be more appropriate to construct metaphoric spatial scenes to visualise such data rather than reproduce the actual 3D environments from which these data are derived. Over twenty years ago, for example, the rooms metaphor was proposed for a general computer interface.
(Henderson and Card, 1986), and the BBC Domesday system adopted a museum metaphor for the display of images stored on videodisc (Rhind et al., 1988). More recently, e-commerce web sites have introduced the virtual shopping mall as a context for browsing purchasable products, collaborative learning websites have adopted 3D virtual worlds (e.g. Börner et al., 2003; Penumarthy and Börner, 2006), while some 3D social networking web sites have adopted familiar 3D environments such as landscapes (e.g. Second Life) or hotels (e.g. Hotel Habbro) for their avatars to inhabit.

In experiments undertaken to visualize telecommunications network data, Abel et al. (2000) suggest that specific tasks may be better supported with 3D scenes built using visual metaphors such as buildings, cities and the solar system, rather than on the inherent geographical structure of the network infrastructure. Their research suggests that an analysis should be undertaken of task requirements in order to determine the utility of adopting natural rather than unconventional spatial representations. Similar considerations also apply to the choice between 3D interfaces that mimic natural human actions and those that overturn conventional principles. Pierce (2000), for example, argues for the need to break existing assumptions about how people interact in the real world when designing interaction methods for virtual 3D worlds, suggesting that “3D worlds are governed by more complex laws than 2D worlds, and in a virtual world we can define those laws as we see fit”.

![Figure 1a. A 2D vertical view of family classes in the East End of London in the late 1880s (Source: Author).](image-url)
Resolving the Hidden-Symbol Problem
One of the largely unrecognised problems of displaying large numbers of point symbols on a conventional 2D map is that they often obscure one another. Where this results from overlaps in the extent of proportional or graduated symbols, then various techniques may be used to resolve the problem, including transparency and symbol cut-outs (e.g. Rase, 1987). However, where symbols obscure one another because the objects they represent share identical locations, an alternative solution is required. An example of this problem is illustrated in Figure 1a, in which the social classes of individual families living in part of the East End of London in the late 1880s are displayed on a conventional 2D map, using data from the manuscript notebooks of Charles Booth’s poverty enquiry (Shepherd, 2002). Unfortunately, because of the widespread incidence of dwelling multi-occupancy in this area, many of the point symbols visible on this map conceal others beneath them. A solution to this problem (Shepherd, 2007b) is to use the third display dimension in a 3D visualization to display stacks of point symbols at each location, as shown in Figure 1b. Because of the widespread occurrence of locationally coincident phenomena, 2D point symbol maps should be used extremely carefully, and 3D stacked symbol visualizations should be used wherever appropriate.

Figure 1b. A 3D oblique view of family classes in part of the East End of London in the late 1880s (Source: Author).

SOME PROBLEMS WITH 3D VIEWS
Although 3D data visualization has undoubted benefits when compared with 2D approaches, a considerable body of user experience and experimental research
reveals significant problems in using the third dimension to create effective data visualizations. A number of these are explored below. However, space prevents detailed consideration of the highly significant problem of user interaction within 3D virtual worlds, which are the subject of a separate study (Shepherd, 2008b).

**Scale Variation Across 3D Scenes**
A perspective view tends to be adopted for most 2.5D and 3D visualizations of spatial data. Unfortunately, because of the foreshortening effect in such views, which increases with distance from the observer, it is difficult to make accurate visual comparisons of objects within a 3D scene. This problem was recognised early on in the history of 3D digital mapping in an evaluation of the problems of interpreting graphical output from the SYMVU software (Phillips and Noyes, 1978). More recently, evidence from perception research reveals that not only is depth consistently underestimated (Swan at al., 2007), but that the human visual system perceives relationships in each of the three directions separately, and that the relationship between physical and perceived space is non-Euclidean (Todd et al., 1995). A number of solutions have been proposed to assist users in making effective distance and measurement estimates in 3D worlds, four of which are briefly reviewed here. It should be noted, however, that no single technique is entirely effective across the range of tasks that need to be performed in 3D scenes.

**Reference Frames**
Most 3D histograms and bar charts provide generated from non-spatial data a set of 3D axes as an integral part of the display, and this assists across-scene position and size estimation by the viewer. However, when spatially referenced data are visualised in 3D, a standard reference framework of this kind is usually absent. In such cases, a user-defined bounding box may be drawn around some or all of the objects within the scene in order to provide some sense of scale. This may be a simple wireframe box, or a more elaborate set of calibrated axes, and the geographical extent of the reference framework may be chosen by the software and/or the user. (Several examples may be seen in the images created using the OpenDX software, which are available at www.opendx.org/highlights.php.) Where large parts of the globe are shown in a 3D scene, the bounding box may needs to be more complex, and where thematic layers are stacked above the globe, each layer may also require its own local reference frame. Clearly, the number of variables being displayed places limits on the effectiveness of this approach, which will be considered again later. For the moment, it is worth noting that perceptual research has revealed a possible distortion effect in the use of an enclosing frame, which may flatten 3D scenes, and reduce perception of depth within them (Eby and Braunstein, 1995).

**Reference and Slicing Planes**
In this solution, a plane is drawn at a specific level within the view, thus acting as a visual plane of reference for the analyst. The plane may be opaque or transparent, it may have grid lines drawn across it, and it may be drawn on any of the three axes, though in most geographical visualizations it will be positioned along the z axis. Cutting planes are commonplace in ‘slice and dice’ voxel medical models, which have become more widely known through the well-known ‘Virtual Body’ and ‘Visible Human’ projects. User-adjustable cutting planes are also widely used for geophysical exploration (e.g. Fröhlich et al., 1999).
A simpler example is illustrated in Figure 2, in which a semi-transparent reference plane is used to enable users to visually identify locations in part of north London where more than five businesses are located at a single address. However, there are design problems with this kind of visualization technique. First, it is only effective where the stacked symbols are drawn with equal heights. (This is an example of the kind of conflict that frequently occurs between visual variables and other visual effects in 3D data visualizations.) Secondly, where the individual stack symbols are sorted by business attributes, as in this example, then the kinds of businesses visible above the reference plane will not necessarily be representative of the range of businesses within the entire stack. Despite these and other limitations, however, reference planes afford numerous insights into value relationships across 3D scenes, such as which symbols lie above or below a particular threshold value, or which land is lower/higher than other land.
Divided Symbol Stacks

In 3D data visualizations which display stacks of locationally coincident point symbols, there are usually no visual cues to indicate the vertical extent of each symbol. This prevents the analyst from making comparative visual estimates of the number of items of particular colours, forms or sizes in each stack within the view. Where stacks contain several symbols using the same set of visual variables (e.g. size, type, colour), it becomes even more difficult to make comparative estimates of the number of symbols in stacks across the entire scene. One way of providing an additional visual cue is to insert small gaps between the symbols in each stack, automatically adjusted by the software to remain proportional to the standard symbol height. An example is provided in Figure 3, which uses stacks of coloured cylindrical solid symbols to indicate the genders of all individuals living in residential properties in part of the East End of London, using data from the 1881 census of population. Not only do the inter-symbol gaps enable the analyst to make rough visual estimates of the number of individuals resident at each location, but the software also assists the making of broader comparative judgements by automatically sorting the symbols into common value groups (males versus females in this example) within each symbol stack.

Non-Perspective Projections

Although non-perspective projections have been widely used in architecture and engineering, and the isometric grid became popular with computer games in the 1990s (e.g. SimCity, Civilisation, A-Train and Theme Park), the perspective projection is almost universally used for 3D data visualization. Wyeld (2005a, 2005b) suggests that this is part of a visual tradition that has dominated art and other visual
media since the renaissance, and that a deliberate effort is needed to wean people away from this expected form of representation. For some data interpretation tasks which involve the comparison of objects across a scene, a case can be made for setting aside the perspective view, and using instead a non-perspective projection. In Figure 4, for example, the view shown in Figure 3 is redisplayed using an orthogonal projection, which makes it easier to make cross-scene comparisons of the sizes of symbol stacks. (With this software, the viewer is able to rapidly toggle between the two projected views with a single keystroke to assist their interpretation of the data.)

![Figure 4. Orthogonal projection of the view shown in Figure 3. (Source: Author).](image)

**Symbol Occlusion within 3D Scenes**

Although 3D data visualization may be used to resolve the problem of hidden symbols in 2D maps, as discussed above, most 3D visualizations still suffer from symbol occlusion, due to the alignment of objects within a scene in relation to the user's viewpoint. Some of the more common solutions to this problem are discussed below. The first four involve often significant modifications of the scene contents to address the problem, while the last three involve less radical surgery. As with solutions to the problem of scale variation, none of the proposed solutions is entirely satisfactory.

**Object Culling**

This is perhaps the most obvious, but also the most draconian, solution to the occlusion problem. The chief drawback is that the removal of occluding objects...
reduces the ability of the analyst to view objects of interest in their natural context. Because much of the power of both 2D and 3D data visualization derives from the user's ability to simultaneously appreciate both detail and context in a single view, object culling considerably reduces the benefits of being able to see contextual data during data exploration. Various approaches have been taken to resolve this problem. For example, in some experimental visualizations of 3D node-and-link graphs (e.g. Hypergraph, 2007), the analyst is able to select the focal point of interest interactively, and selectively expand and collapse sub-trees. A more rigorous solution is proposed by Shen et al. (2006), who suggest the use of semantic and structural abstraction to declutter selected nodes and links from such graphs.

Object Minimisation
An alternative approach is to temporarily redraw inessential objects so as to reduce their visual interference with the point(s) of interest. For example, in experiments with 3D point symbols on a planar base map, Chuah et al. (1995) reduced the sizes of unimportant, occluded symbols, leaving the sizes of the focal symbols unchanged. Although reductions to both the heights and widths of occluding symbols is possible, the study found that width reduction was more useful where the analyst needed to compare the heights of selected symbols with all other symbols within a scene.

Object Displacement
By moving selected objects away from one another in a cluttered scene, it is possible to reveal objects of interest without entirely removing or rescaling the selected objects as in the previous two methods. Three approaches have been proposed. In the first, a subset of objects of interest, together with a reference plane, is raised above the full set of objects in the original visualization. This technique is adopted in experimental data visualisation software by Chuah et al. (1995) and Schmidt et al. (2004). The second method involves the differential movement of objects, or parts of objects, away from one another within a scene. A 2D example is provided by the point symbol dispersion routine provided with the MapInfo desktop mapping system. A variation of this method is more widely used in the 3D display of anatomical and other medical illustrations, creating what an ‘exploded view’ (e.g. Bruckner and Gröller, 2006). However, the use of exploded views in statistical data visualization, as in the case of 3D exploded pie graphs, introduces interpretive errors into the display (Shepherd, 2008a). The third object displacement method involves the repositioning and spatial distortion of contiguous (usually zonal) objects in geographical space, as in 2D area cartograms. A danger with this family of techniques is that they render the user less likely to be able to interpret the exact spatial relationships between objects that have been moved (often by some arbitrary amount) from one another.

View Distortion
By differentially distorting the overall geometry of a scene, objects near the user viewpoint may be more clearly seen. Perhaps the best-known technique is the fish-eye view (Furnas, 1991), which selectively distorts a scene so as to enlarge objects near the point of interest. One of the problems of this technique is that users may be unable to obtain a proper sense of the spatial relationships between objects, though this is more likely to affect displays of inherently spatial data than displays of purely statistical data, and it should be noted that many geographers will be used to the spatial deformations used in map projections and cartograms.
**Rotation or Viewer Movement**
One of the biggest advantages of interactive 3D data visualization over printed 3D images is that users have the ability to resolve some of the occlusion problems interactively, either by rotating objects within the scene or by moving their viewpoint with respect to the scene. This induces the kinetic depth effect, which not only reduces symbol occlusion, but also enhances the viewer's appreciation of the depth relations between objects in a scene (Ware and Franck, 1996). However, this may be accompanied by unpleasant user side effects, especially in an immersive 3D environment and may also required adeptness in navigating within 3D scenes which many analysts will not possess.

**Symbol Transparency**
By displaying selected symbols in reduced opacity, occluded symbols may be seen through foreground symbols. Several options are available: symbols in a scene may be drawn with equally reduced opacity, or the transparency of selected symbols near the focus of the user’s attention may be increased, leaving those further away at full or increasing opacity. A somewhat different role for symbol transparency is to enable the comparison of two datasets in a single scene by applying transparency to one set of symbols and rendering the other set opaquely. An example is illustrated in Figure 5, which shows the comparative distribution of cigar makers in the East End of London from two datasets: the population census of 1881 (the green transparent symbols) and Booth’s poverty survey of the late 1880s (the red opaque symbols) (Shepherd, 2000).

![Figure 5. Use of opaque and semi-transparent symbols to compare the distribution of cigar makers in the East End of London in 1881 (transparent) and 1887 (opaque) (Source: Author).](image-url)
Symbol Shadows
When objects displayed in 3D space are viewed from a particular direction, symbol self-occlusion often makes it difficult to perceive their spatial distribution within the three axes of that space. One way of providing information about the spatial distribution is to project object shadows onto one or more planes of a bounding box surrounding the objects. (Some of the key design issues are discussed in a medical illustration context by Ritter et al., 2003.) An example of this technique is illustrated in Figure 6, which shows the distribution of earthquakes below the Big Island of Hawaii, as viewed from a medium-oblique angle. The grey base plane displayed below the scene shows the shadows of the 3D earthquake symbols projected from above, and helps to clarify their distribution in the x-y plane. Some studies show, however, that not all 3D tasks are equally enhanced by including object shadows (Hubona et al., 2000).

Figure 6. Use of projected symbol shadows to reveal the distribution of Hawaiian earthquakes in the x-y plane of a 3D visualization (Source: Author).

Multiple linked views
The occlusion of objects caused by scene complexity sometimes prevents users from comprehending the entire visualized scene from a single viewpoint. One potential solution to this problem is to display the virtual world simultaneously as seen from two or more alternative viewpoints in linked views. In one experiment with
a multi-view 3D visualization, Plumlee and Ware (2003) found that the provision of a view proxy (i.e. a triangular symbol in a vertical view indicating the field-of-view in an oblique view) and view coupling (i.e. keeping vertical and oblique views oriented in the same direction) were both beneficial for undertaking certain tasks in 3D worlds. In an evaluation of alternative methods of presenting route instructions on mobile devices, Kray et al. (2003) revealed that in order to find their current location on a 2D map, users often found that a 3D view was useful in identifying their location in the real world. This study also confirmed a finding of several other studies, which is that although the 3D display did not improve task performance, it was found to be ‘fun’ to use.

Symbol Viewpoint Dependencies
A normal feature of navigating through 3D scenes is that the perceived dimensions and/or shapes of objects in the scene vary according to their orientation with respect to the viewer. This becomes problematic in those data visualizations where the appearance of the 3D symbols is meant to visualise data variables. An example occurs where the dimensions of 3D bar symbols (i.e. height, width, and thickness or depth) are used to encode data variables. The problem in doing this is that the perceived bar widths will not only vary according to the assigned data values, but will also vary with the observer’s viewing angle. (An example occurs in a study of retail businesses in Toronto (Hernandez, 2007), in which 3D bar symbols representing businesses in Toronto are aligned with the streets on which they are located.) In extreme circumstances, where bars symbols with minimal thickness are viewed from the side, they may all but disappear. This problem was identified by the authors of a study involving the data visualization of ocean-bed characteristics by Schmidt et al. (2004) and was resolved by using spherical 3D glyphs whose perceived sizes were independent of their orientation.

The viewpoint dependency problem not only affects the ability of symbol dimensions to carry data information, but it may also undermine the use of symbol variations to represent data. For example, where symbol shapes are varied to reflect nominal scale data, then the analyst’s viewpoint may make it difficult to perceive these shape differences across the scene (Lind et al., 2003). There are several other reasons why the 3D equivalents of 2D symbols may not work effectively in 3D scenes. As Kraak (1988, 1989) has indicated, the surface shadowing applied to 3D point symbols to enhance their realism may conflict with the visual variations in the lightness and texture applied to their surfaces. Krisp (2006) illustrates a similar problem with 3D density surfaces, in which the colours used to encode height are locally modulated by the hill-shading used to enhance viewing realism. The general conclusion seems to be that although it is relatively trivial to convert the standard 2D geometrical symbols used in conventional thematic mapping into 3D equivalents for data visualization, in which a square becomes a cube, a circle becomes a sphere, a line becomes a ribbon or wall, and a region becomes a prism, variations in the size and shape of these symbols may not be accurately perceived in a 3D data visualization because of viewpoint dependencies.

Stereo 3D: pretty useful, or just pretty?
Viewing 3D data visualizations in full stereo is assumed to provide the benefits of binocular viewing enjoyed by human primates in their natural environments. Although stereo images possess a certain ‘wow’ factor, as indicated by the
entertainment value of the iMAX cinematic experience, effective stereo data visualizations are not always easy to create, and require that close attention be paid to known perceptual principles. In Figure 7, for example, which shows the distribution of earthquakes below the Big Island of Hawaii in 2003, use is made of the principle that variations in the lightness visual variable are better able to heighten the viewer’s appreciation of depth than variations in hue (Ware, 2004).

Figure 7. Stereo view of earthquakes below Big Island, Hawaii, using luminescence/lightness variations to enhance the depth effect (Source: Author; data from CNSS).

An even more important decision-making factor concerns the evidence from visual perception research which suggests that stereo is less effective than several other techniques for indicating depth and scene layout to the viewer when undertaking tasks in 3D visualizations. Indeed, stereo is only one of nine major visual depth cues and, in many circumstances, is not the most important (Cutting and Vishton, 1995). A question worth asking is whether this visualization provides users with a better indication of the three-dimensional distribution of the earthquakes than, say, a monographic view that includes symbol shadows (as in Figure 6). One of the more significant alternatives to stereo is the kinetic depth effect, in which an awareness of the depth relations among objects in a scene is induced by the relative motion of foreground and background objects, either as the viewer moves or as the objects are rotated. In experiments using a head-mounted device to induce 3D perception, Ware and Franck (1996) found not only that motion parallax alone is better than stereo
alone for tasks involving full 3D awareness, but also that motion parallax in combination with stereo provides the best depth cues for the analyst. Stereo should not therefore be considered a ‘must have’ facility but, as Ware and Franck (1996) suggest, individual data visualizations should use particular combinations of 3D-inducing effects. Few hard and fast rules are available, because the relevant combinations need to be task specific, and further research is needed to evaluate which 3D effects are best suited to particular data visualization tasks.

Z-axis Contention

With the increasing availability of spatial tracking data, several attempts have been made to automate the production of the space-time cube (Oculus, 2006; Kraak, 2007), first introduced in the 1960s (Hagerstrand, 1970). In the resulting visualizations, the third dimension is used to display time, with the x and y axes retaining their traditional role of displaying ground surface features. Unfortunately, when spatial data are visualised in a space-time cube, the z-axis is not really available as a spare dimension to be used exclusively to show time, because the movement of people and objects also takes place within the vertical spatial dimension. Although it is tempting to ‘overload’ the z-axis by using it to display both the three-dimensional landscape surface and the space-time lines above it, doing so can lead to difficulties in visual interpretation. This is because the space-time lines attached to different points on the surface are no longer visually comparable, because they are offset in the vertical plane by different amounts depending on the height of the landscape surface to which they are attached.

In some low-lying study areas, as in parts of the Netherlands (Kraak 2003), this problem may be largely ignored. In other areas, where it is known that topography has little influence on space-time patterns, this problem may be solved by adopting the age-old fiction of the traditional 2D map: a flat world. However, in many hillier and mountainous areas, where it is important to understand how the landscape surface impacts on people’s movement patterns, terrain height must also be shown in the cube. In such cases, z-axis contention becomes a serious problem that threatens to undermine the benefits of this particular 3D visualization technique.

One possible solution may be to attach the bases of all space-time lines to an arbitrary plane above the highest point of the landscape surface, though this may make it difficult for the interpreter to relate the two sets of information. As previously mentioned, this stacking technique is used in several visualizations of spatial point-located data (e.g. Chuah et al., 1995; Schmidt et al., 2004), and also in numerous 3D meteorological visualizations where a single surface layer is displayed some way above a globe. In one example (Aoyama et al., 2007), the suspended isosurface showing land-surface temperatures is visually related to the plane base map of the USA below it by having the outlines of regions of interest in the former projected down onto the latter. However, where the z-axis is used to display data for several variables, resulting in a stack of multiple 3D thematic layers displayed above a reference surface. Even more complex techniques are needed to relate the locations of objects on one layer to those on other layers. In oil reservoir visualizations (e.g. Calomeni et al., 2006), the inclusion of vertical lines representing boreholes partly reduces this problem. In general, however, and despite the undeniable artistry of some of the complex visual models, one begins to wonder whether one has reached the limits of 3D data visualization as an exploratory and interpretive tool. Rather than
overload the z axis, it might be more effective for the analyst to resort to map overlay
analysis or, as discussed in the next section, to adopt some form of dimension
reduction techniques in order to simplify the data before or during the visualization
process.

The ‘Dimensionality Curse’
The challenge of what Robertson et al. (1991, p.189) refer to as “intellectually large
data collections” is a problem that is only partly solved by moving from 2D to 3D
visualizations. Although it is possible, as previously discussed, to display very large
numbers of objects in 3D scenes with powerful hardware and clever algorithms,
visualization limits are rapidly reached as the number of dimensions (i.e. variables,
fields or attributes) increases above a relatively small number. As Bertin (1967,
1977) fully appreciated, the graphical sign system cannot be used to display more
than a handful of variables in a single scene, and this remains the case even after it
has been augmented with additional visual variables (e.g. transparency, blur, and
specular highlights), or extended into the third dimension through complex
articated glyphs and icons.

Even when 3D symbols are used to encode up to half a dozen data variables, this is
still a long way away from being able to satisfy the needs of many analysts, whose
datasets commonly include scores or even hundreds of variables. In the innovative
3D seafloor visualizations created by Schmidt et al. (2004), for example, only 5
variables were simultaneously displayed, and in the iconographic visualizations
created by Gee et al. (1998) and Healey (1998), only a similar number of data
variables were encoded by means of icon geometry and colour. Claims (e.g. by
Wright, 1997) that certain 3D data visualization techniques are capable of displaying
hundreds of variables are therefore little more than marketing hype.

Given this problem, it becomes necessary to think the unthinkable, and to consider
setting aside the representational spatial framework in which much of our data is
gathered. Indeed, relevant 2D and 3D visualization techniques that are stripped of a
geographical frame of reference may provide powerful insights into complex data in
ways that cannot be provided by conventional spatial visualizations.

Three broad strategies have been adopted. The first is to retain the geographical
semantics of the real world in the 3D display space, but to switch between visual
representations of selected (small) subsets of variables. Most of the discussion in
this chapter so far has focused on this strategy. A second strategy is to abandon
geographical semantics in the 3D display space, and to use visual data mining or
visual analytics techniques (de Olivera and Levkowitz, 2003) to display information for
multiple variables in 2D displays. Such visualizations include parallel coordinates
(Inselberg and Dimsdale, 1989), pixel-based displays (Keim, 2000), ‘dimensional
anchors’ (Hoffman, Grinstein and Pinckney, 2000), and an increasingly large number
of other multivariate data visualization techniques (Wong and Bergeron, 1997).
Where the number of variables or dimensions is extremely large, 2D projections may
be made of multidimensional data into 2D spaces, or components may be derived
from the original data by various data reduction techniques, and displayed in 2D
display space (e.g. Jing et al., 2007; Yang et al., 2004; Yang et al., 2007). By coupling
data reduction techniques to data visualization, visualization continues its tradition of
helping the analyst to steer the data mining in potentially fruitful directions (Keim and
Kriegel, 1994; Keim, 2002), but operating within statistical space rather than geographical space.

A third strategy is to combine the spatial and non-spatial approaches in a multiple linked views environment. For example, parallel coordinates displays have been incorporated into several software systems designed for the analysis of spatial data, along with 2D maps and other forms of tabular and graphical display (e.g. Stolte et al., 2002; Andrienko and Andrienko, 2003; Guo, 2003; Guo et al., 2006; Marsh et al., 2006). (This multiple linked views approach has also been developed in other research fields (e.g. Gresh et al., 2000).) A notable feature of these hybrid systems is that 3D spatial representations are notable by their absence; almost without exception, they only incorporate 2D maps and 2D statistical graphics.

This suggests two interesting possibilities. The first is that 2D display techniques may be more effective than 3D techniques for routine data interpretation purposes, even for spatial data. The second is that analysts looking for comprehensive data visualization software should not expect to find them solely among the offerings of current GIS and desktop mapping software vendors, whose focus is mainly on spatial data management and analysis. Most of the innovative data visualization software of the past two decades has emerged from the non-spatial sciences, and especially the information visualisation community. By interfacing effective modular software from these sources to standard 2D mapping tools developed within geography, it may be possible to acquire the most effective toolkit for both spatial and non-spatial data analysis. This has been the motivation for the author’s own experimental 3D data visualization software, which imports data from MapInfo and ArcView, in the spirit of his earlier advocacy of DIY GIS (Shepherd, 1990) and software federations (Shepherd, 1991). However, thorough evaluation of the usefulness of such software federations for routine analytical tasks is necessary if we are to determine whether they outperform currently available GIS and mapping software, and whether they render eye-catching 3D visualizations an unnecessary luxury.

CONCLUSIONS
Over the past quarter of a century, a wealth of experimentation in 3D data visualization has taken place, so that just about anything one wants to see in 3D can now be produced almost automatically. However, our review suggests that 3D is not always as useful for data visualization as it has sometimes been made to appear. Each of the advantages and benefits claimed for 3D have their caveats, and not all of the known problems with 3D have completely satisfactory solutions. As graphics hardware gets increasingly powerful, and visualization software gets ever more sophisticated, it becomes increasingly important to step back from the compelling visual image on the screen and ask some relatively simple questions: is the visualized scene free from distortion, bias or other visual error?; are the display methods used appropriate for the task in hand?; would any patterns hidden in the data be more evident if 2D visualization methods were used?; and: are the visualization techniques being used best suited to the current user?

Just because it can be done does not meant that it should be done; some 3D effects are of questionable analytical value, and 3D is not always better than 2D for visualizing data. Indeed, Lind et al. (2003) suggest that because of the distortions
introduced by human space perception, “the general usefulness of a 3D visual representation may be limited”, particularly in situations where analysts are meant to discover relations based on Euclidean distances or shapes. They suggest that the primary role for 3D may be in providing users with a general overview of object relationships in a scene, and especially for spatial data. Others, (e.g. Kray et al., 2003) have suggested that a large part of the appeal of 3D displays for users undertaking particular spatial tasks lies in their entertainment or ‘fun’ value. For his part, the guru of web usability has thrown the following provocative claim into the ring: “3D is for demos. 2D is for work.” (Nielsen, 2006; see also Nielsen, 1998).

A great deal of evaluation remains to be undertaken to identify which, when and how currently available 3D data visualization tools and techniques should be used. Despite the considerable progress made in recent years, the technology of data representation is still in its formative stages, and developers, researchers, educators and users alike have major contributions to play in improving the technology, and its effective use. Developers need to bridge the gap between what is currently available and what is desirable; researchers need to undertake rigorous evaluations of alternative approaches to visualization and interaction, in order to identify the fitness for purpose of existing and emerging technologies; educators and trainers have a responsibility to help users understand the principles and limitations of 3D visualization, as well as teaching them how to make effective choices in harnessing the power of available tools for their needs; and individual users face a continuous learning challenge in making effective use of the many dimensions available to them in making sense of their data. We may all have been born into a 3D world, but we need to be continually aware that our virtual 3D worlds are sometimes more challenging than the real thing.

NOTES
1. It is generally agreed that Wolfenstein 3D, a first-person shooter (FPS) game, which was released by id Software in 1992, triggered the initial mass appeal of 3D computer games. This success was followed in 1993 by Doom, and in 1996 by Quake, with its full 3D engine (id Software, 2007).

2. The experimental 3D data visualization software used to generate many of the illustrations that accompany this chapter was devised by the author, based on data visualization principles and the published results of experimental research in visual perception and human-computer interaction. The software was programmed by his son Iestyn Bleasdale-Shepherd, who is a software engineer specialising in real-time computer graphics at Valve Software in Seattle.

ACKNOWLEDGEMENTS
It is a pleasure to thank my son, Iestyn, for the many fruitful discussions on 3D computer graphics we have had over the years, for helping me to understand the potential contributions that modern computer games technology can make to geographical data visualization, and also for cutting the code. Many of the ideas in this chapter could not have been developed without his continuing partnership in my research.
REFERENCES


ACM International Conference on Information and Knowledge Management, pp.9-16.


Kraak, M-J. (2007) Geovisualization and time - new opportunities for the space-time cube, this volume.


Wyeld, T. G. (2005b) The pedagogical benefits of stepping outside the perspective paradigm: challenging the ubiquity of Western visual culture, ETCC’05, Taiwan.