

# Visibility graphs and landscape visibility analysis

DAVID O'SULLIVAN and ALASDAIR TURNER

Centre for Advanced Spatial Analysis, University College London,  
1–19 Torrington Place, London WC1E 6BT, England, UK;

e-mail: {david.osullivan, alasdair.turner}@ucl.ac.uk

## Abstract

Visibility analysis based on viewsheds is one of the most frequently used GIS analysis tools. In this paper we present an approach to visibility analysis based on the *visibility graph*. A visibility graph records the pattern of mutual visibility relations in a landscape, and provides a convenient way of storing and further analysing the results of multiple viewshed analyses for a particular landscape region. We describe how a visibility graph may be calculated for a landscape. We then give examples, which include the interactive exploration of a landscape, and the calculation of new measures of a landscape's visual properties based on graph metrics—in particular, neighbourhood clustering coefficient and path length analysis. These analyses suggest that measures derived from the visibility graph may be of particular relevance to the growing interest in quantifying the perceptual characteristics of landscapes.

## 1. Introduction

Viewshed analysis is one of the most commonly used analysis tools provided by geographical information systems (GIS) (Davidson *et al.* 1993). Burrough (1986) provides a useful introduction. Development of the technique continues apace. Areas which have attracted particular attention are the development of improved algorithms for the basic technique (Wang *et al.* 1996), and the effects of errors and uncertainty in underlying data on the viewshed areas which are generated (Fisher 1991, 1995, 1996). The generation of viewsheds on triangulated surfaces is also of interest (De Floriani and Magillo 1994). The methodology has widespread relevance to the planning of facility locations, and to the management of environmental resources. The location of line-of-sight communications transceivers is dependent on the intervisibility properties of a region (Goodchild and Lee 1989). In rural areas, planners may be concerned about the intrusion of new facilities on current residents (Davidson *et al.* 1993). Visibility is also important in preventing visual intrusion on the environment of endangered species (Camp *et al.* 1997), and in providing the best recreational amenity for hikers and others (Lee and Stucky 1998). Others have attempted to use the approach in modelling micro-scale variations in local property markets (Lake *et al.* 1998). The methodology also has an obvious interest in military contexts.

Viewshed analysis has also been used in archaeological studies, in attempts to

understand the ways in which ancient cultures may themselves have understood and used their physical environment, and in turn how their understanding may have influenced the siting and location of monuments and settlements (Wheatley 1995, Llobera 1996). Attention has also begun to focus on the experience of landscape in geographical information science, and the potential for deriving measures of landscape properties from viewshed and other analyses of the digital elevation models (DEMs) or triangulated irregular networks (TINs) representing landscape (Baldwin *et al.* 1996). Viewshed analysis also has a long heritage in landscape architecture (Tandy 1967, Amidon and Elsner 1968, Lynch 1976). In that tradition, under the guise of 'isovist' analysis Benedikt and Burnham (1981) explicitly relate viewsheds to concepts from Gibson's (1979) ecological theories of spatial perception. This suggests that approaches based primarily on visibility analysis may be of particular relevance to psychological, cognitive, and perceptual studies.

A *visibility graph* is formed by linking locations in an environment according to whether each is visible from the other. The concept first appears in the geographical information science (GISci) literature in the context of determining preferred routings through a set of line-of-sight communications transceivers in a landscape (De Floriani *et al.* 1994). Visibility graphs are also used in the field of robot navigation in complex two- and three-dimensional environments (de Berg *et al.* 1997). In the two-dimensional case of navigation by a robot amongst a set of polygonal obstacles the visibility graph is based on the set of polygon vertices and their mutual visibility. Such a graph contains all the possible shortest paths through such an environment and is therefore useful. In three-dimensional cases further complexity is introduced because shortest paths may be tangential to polyhedral obstacle edges, so that these must also be considered as vertices in the graph (Jiang *et al.* 1996, 1999), but the principle is the same.

The character of a visibility graph is dependent on complex metric and geometric properties of the environment for which it is derived, and it is this which makes it an interesting object of enquiry. A generic visibility graph connects any set of points, and we can use various analysis techniques to construct point measures for these locations. The method we use is based on the visibility graph analysis of architectural systems (Turner *et al.* 2001). In landscape settings a highly connected graph typically results, which can be used both as a convenient data structure to explore visibility characteristics of the landscape, and as a tool to provide further analyses not calculable directly from viewsheds. We show how the technique can be generalised to an analysis of any environmental space, how we can implement the method using a GIS, and how it can be used to explore and analyse landscapes.

In §2 we introduce necessary concepts, definitions, and symbols from graph theory. In §3 we describe the process we have used to construct landscape visibility graphs. In §4, we describe possibilities for progress in visibility analysis which are provided by visibility graphs. A discussion of the methodology and suggestions for further work follow in §5. Tentative conclusions are offered in §6.

## 2. Graphs and graph terminology

For an introduction to the mathematical theory of graphs, and relevant terminology refer to Wilson (1996). A graph  $G$  consists of two sets  $V$  and  $E$ , and is often written  $G(V, E)$ .  $V(G)$  is a set of vertices (or nodes)  $\{v_1, v_2, \dots, v_n\}$ , where  $n$  is the size of the graph.  $E(G)$  is a set of *edges* (or links) between the vertices, where each edge  $e$  is an unordered pair of vertices  $v_i$  and  $v_j$ , and may be written  $v_i v_j$ , or  $e_{ij}$  for brevity.

The sub-graphs of a graph are an obvious and useful concept. A sub-graph  $H$  in  $G$  is defined such that  $V(H) \subseteq V(G)$  and the edge set of  $H$ ,  $E(H)$  is any subset of  $E(G)$  with the obvious constraint that edges in  $E(H)$  may only be a subset of those edges in  $E(G)$  which link members of  $V(H)$ . A sub-graph consisting of some subset of  $V(G)$ ,  $X$  and all the members of  $E(G)$  which link members of  $X$  is said to be *induced* by  $X$ .

The *neighbourhood* of a vertex  $v_i$  in  $V$  is the set of vertices  $N(v_i)$  which are joined to  $v_i$  by edges in  $E$ :

$$N(v_i) = \{v_j : v_i v_j \in E(G)\} \quad (1)$$

which we may abbreviate as  $N_i$ . The size (or *degree*) of  $N_i$  is often denoted by  $k_i$ . Note that  $v_i$  itself is not normally considered a member of  $N_i$ —since  $v_i v_i$  is not a member of  $E(G)$ . In some cases it may be useful to think of the neighbourhood of  $v_i$  as the sub-graph induced by  $N_i$ . We will denote this understanding of the neighbourhood by  $N_i^*$ . This neighbourhood sub-graph consists of the pair of sets given by

$$N_i^* = \langle N_i, \{v_i v_j : v_j \in N_i \wedge v_i v_j \in E(G)\} \rangle \quad (2)$$

Another important concept is the *adjacency matrix*  $A$  of a graph  $G$ .  $A$  is defined such that

$$A = [a_{ij}] \text{ where } a_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E(G) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The adjacency matrix is useful in understanding the relationship between viewshed analysis and the visibility graph. In a simple graph  $a_{ij} = a_{ji}$  in all cases. However, in a *directed graph* this is not necessarily the case, and the symmetry of linkages between vertices may be broken. Many manipulations of simple undirected graphs are also possible on directed graphs, although in our presentation all visibility graphs are undirected. Nevertheless, the generality of our approach is not affected by this possibility. Having completed these preliminaries, we may formally define the visibility graph concept and discuss its construction in a landscape.

### 3. Constructing a landscape visibility graph

A visibility graph is a graph where the vertex set  $V$  is some set of locations in the environment, and the edge set  $E$  consists of all those pairs of locations which are mutually visible. Such a graph may be constructed for any environment for any set of locations, suitably defined, in that environment. It may be easier to visualise a visibility graph by thinking of it as the set of locations and all the unimpeded lines-of-sight between those locations. Note however, that graphs are not intrinsically geographical entities, and that only the relational information recording whether two locations are mutually visible is recorded in a graph. It is not in general possible to reverse the process of visibility graph construction—that is, many sets of locations in many environments could produce the same visibility graph, so that the environment so represented can not be recovered from the visibility graph. Nevertheless, the visibility graph does record all the visibility information in an environment and can be usefully analysed, as we demonstrate in the next section.

Construction of a visibility graph for a landscape is a two-stage process. First, a set of locations  $V$  must be determined. Second, each pair of locations in  $V$  must be considered and their mutual visibility determined in order to construct the edge set  $E$ . The first step in this process is a matter of deciding on the investigators' priorities or requirements, as there is in general no way of knowing *a priori* which set of

locations in an environment are of particular interest. In an archaeological study, for example, it may be that the locations of a set of artefacts constitute an interesting set to consider. In studying the perception of landscape it might be that the peaks, saddle points, troughs, points along ridge lines, and so on represent a suitable vertex set (we briefly examine this possibility in §5.1 below). In the examples we consider, we have adopted the pragmatic approach of generating a vertex set by selecting locations on a regularly spaced grid. Using a set of spot height locations seems likely to be a sensible ‘first cut’ in many cases. In a DEM this would equate to our regular grid, but in a TIN, it would result in an irregular network of point locations. By adopting the grid sampling approach in our examples we do not intend to suggest that this is a preferred or best approach. The choice of the vertex set is a fertile area for further research, which could be usefully informed by studies of the impact of different DEM interpolation techniques, and spot height sampling strategies.

Having determined the set  $V$ , it remains to construct the edge set  $E$ , or equivalently, the adjacency matrix  $A$ . Figure 1 shows the most straightforward approach. Such an approach is easily implemented in many current desktop GIS, where a ‘CAN-SEE’ test is a standard visibility analysis tool which returns true if points  $v$  and  $w$  are mutually visible. This algorithm has time complexity  $O(n^2)$  where, as before,  $n$  is the number of locations in the vertex set  $V$ . If the landscape  $S$  is represented by a TIN of size  $s$ , then the overall complexity of the algorithm is  $O(sn^2)$ . An alternative approach (see figure 2) calculates the adjacency matrix of the visibility graph. This is completely equivalent to the previous approach, and also has time complexity  $O(sn^2)$ , but draws attention to the important point that the rows in a visibility graph adjacency matrix are equivalent to the binary viewshed of the location which that row describes. Thus, the neighbourhood of a location in a visibility graph corresponds to the viewshed from that location. The accuracy of the correspondence depends on the density of the chosen locations in the vertex set. In a sparse vertex set, a vertex neighbourhood is effectively a punctiform representation of the corresponding location’s viewshed. In a densely distributed vertex set (such as the regularly spaced grids we have used) the correspondence is sufficiently close for many purposes to constitute equivalence.

```

Algorithm CREATE_EDGE_SET(  $V, S$  );
//  $V$  is a set of vertices in the spatial environment  $S$ 
//  $E$  is the output set of edges
begin
   $E \leftarrow \emptyset$ 
   $V' \leftarrow V$  // target set of vertices
  for every  $v \in V$  do
     $V' \leftarrow V' \setminus v$ 
    for every  $w \in V'$  do
      if CAN_SEE(  $v, w, S$  ) then
         $E \leftarrow E + \{ vw \}$ 
      end if
    end for
  end for
end CREATE_EDGE_SET

```

Figure 1. Algorithm CREATE\_EDGE\_SET for determining the edge set for a given vertex set and landscape.

```

Algorithm CREATE_ADJACENCY_MATRIX(  $V, S$  ):
//  $V$  is a set of vertices in the spatial environment  $S$ 
// assume that vertices in  $V$  are indexed from 0 to  $n - 1$ 
//  $A$  is the output adjacency matrix of the visibility graph
begin
   $A \leftarrow \{ 0 \dots n - 1 \} \{ 0 \dots n - 1 \}$  i.e. the  $n \times n$  zero matrix
   $V' \leftarrow V$  // target set of vertices
  for every  $v \in V$  do
     $V' \leftarrow V' \setminus v$ 
     $P \leftarrow \text{VIEWSHED}(v, S)$  //  $P$  is a viewshed polygon or
    // collection of disjoint polygons
    for every  $w \in V'$  do
      if CONTAINS(  $P, w$  ) then
         $A[v][w] \leftarrow 1$ 
         $A[w][v] \leftarrow 1$  // assume mutual visibility
      end if
    end for
  end for
end CREATE_ADJACENCY_MATRIX

```

Figure 2. Algorithm CREATE\_ADJACENCY\_MATRIX. This demonstrates the relationship between viewsheds and the visibility graph.

Note that these approaches assume that locations are necessarily mutually visible, and that the resulting visibility graph is a simple undirected graph. If we wish to allow visibility relations which are not symmetric then every pair-wise visibility test must be carried out in both directions, and the corresponding directed edges  $v_i v_j$  and  $v_j v_i$  stored separately. There seems to be no trivial way of making the visibility graph construction process any more efficient, since although landscapes are typically locally correlated, there is no obvious way of using this information to improve the efficiency of line-of-sight or viewshed operations. In the examples considered in this paper visibility graphs have been constructed using a line-of-sight operation on a TIN representation of a landscape, for the pragmatic reason that this was the most efficient operation available through the GIS scripting language. In general, any convenient visibility analysis technique can be used. The important point is that the visibility graph requires significant pre-processing to be carried out, equivalent to calculating viewsheds from all the locations in the vertex set on the region under investigation. Often, much of the information produced by such an exhaustive visibility analysis process is discarded, and summarised in terms of the viewable area from each location. Storing the results in a visibility graph allows the possibility of further exploratory and analytical processing.

Once constructed, a visibility graph can be conveniently stored as a list of locations. Each vertex (location) object might have the following characteristics:

```

Object Vertex
begin
  integer id
  list-of-vertex-ids neighbours
end Vertex

```

The space complexity of this visibility graph data structure is  $O(kn)$  where  $k$  is the average size of each vertex neighbourhood, as before. A more compact data structure would not store each edge twice, but will be less conveniently searched. Depending on the characteristics and extent of the region,  $k$  may be linearly dependent on  $n$  (in a relatively small region covered at high resolution by  $V$ ), or independent of  $n$  (in a more extensive region where the effects of earth curvature tend to limit  $k$ ). It is clear in any case, that this is likely to compare unfavourably with the storage complexity  $O(s)$  of an underlying TIN landscape representation. The extent of the storage overhead depends on the relationship between  $n$  and  $s$ .

The above data structure is one simple possible approach. It may be more useful to store the line of sight relations in two lists—one of other vertices visible from the current one, and another of vertices from which the current one can be seen, and further duplicating the storage of edges. This will also facilitate search operations on the graph. Perhaps more importantly, it is preferable that the vertices be stored in a random access data structure, such as a hash-table or binary tree so that interactive exploration of the graph (as discussed in the next section) is relatively efficient. In a GIS it is likely that vertex geo-coordinates will also be stored, and that these be used as the keys for indexing the table or tree—for example, a quad-tree could be used—so that spatial search of the visibility graph is efficient.

#### **4. Using a visibility graph**

As has been mentioned, construction of a visibility graph involves considerable pre-processing of a landscape under study. The storage requirements are also significant. Whether or not such processing effort is worthwhile will depend on the purpose of the study under way, and will be influenced by the additional benefit which can be derived from exploration and analysis of the graph, as discussed in this section.

##### *4.1. Landscape exploration*

Once constructed, a visibility graph provides rapid access to the punctiform viewshed of any location. This is an immediate consequence of the correspondence between visibility graph vertex neighbourhoods and the viewshed visible from the location represented by that vertex. When a GIS user selects a particular location in a landscape for which a visibility graph has previously been generated and stored, the neighbourhood of the corresponding vertex in the graph can be rapidly retrieved and displayed. This process will generally be more efficient than typical viewshed determination processes. If a landscape is represented by a TIN of size  $s$  then the punctiform viewshed for one location in a set of  $n$  distinct locations, can be determined in  $O(sn)$  time, whereas an efficiently indexed visibility graph structure should be searchable—and thus pre-computed viewsheds are retrievable—in  $O(\log n)$  time. The significance of the time saving is again dependent on the relationship between  $n$  and  $s$ . An  $O(\log n)$  efficient process will usually make ‘point-and-click’ exploration of the visibility characteristics of a landscape feasible, and rapid enough to be interactive. Interactive exploration of landscape in this way is likely to promote faster understanding of its visibility characteristics.

Related to this improved interactive exploration capability, is the possibility of rapid response to more complex visibility queries. For example, the region in a landscape which is visible from two or more locations is the intersection of all the neighbourhoods of the associated vertices in the visibility graph. If  $Q$  is a subset of

the vertices in a landscape then the region  $X$  which is visible from all of them is given by

$$X(Q) = \bigcap_{v \in Q} N(v) \quad (4)$$

Similarly, the determination of the cumulative viewshed  $Y$  of all locations visible from any location in some set  $Q$  in a landscape is given by

$$Y(Q) = \bigcup_{v \in Q} N(v) \quad (5)$$

which is simply the repeated union of the neighbourhoods of the associated vertices in the visibility graph. Similar queries in a typical implementation of viewshed analysis may require the creation and manipulation of many intermediate data sets. Of course, the improved efficiency of such exploratory analysis is only possible because all the required viewsheds have already been generated, and are efficiently stored in the visibility graph. Whether or not the pre-processing involved in creation of the visibility graph is justified or not will depend on the task at hand. It seems likely that more ‘interpretative’ applications are likely to benefit most from the interactivity provided. Thus the archaeologist seeking to ‘get a sense’ of a historical landscape, or an environmental manager planning hiking trails in a national park seem most likely to benefit. Where the landscape in question is not readily accessible the visibility graph approach is likely to be most useful, and could supplement other approaches such as the construction of interactive virtual environments representing the actual landscape. On the other hand, where other considerations intervene, such as site suitability due to transport access or soil type—as for example in the choice of location for some facility—the number of potential sites is likely to have been sufficiently reduced beforehand to render the exploratory visibility graph process redundant. Thus, in their study of line-of-sight communications problems, where  $V$  is a relatively sparse set, De Floriani *et al.* (1994) conclude that this computational and storage overhead is not repaid, and suggest that an approach where intervisibility are calculated dynamically as required is preferable.

#### 4.2. Analysis of the visibility graph

A further benefit of the visibility graph is the opportunity it provides for further analysis of the visibility characteristics of a region. Following Turner *et al.* (2001), we can implement a series of measures to investigate visual properties of locations within the graph.

##### 4.2.1. Vertex neighbourhood size—viewable area

The most obvious immediate analytic measure presented by a visibility graph is that the size of each vertex neighbourhood is a measure of the visual accessibility of the associated locations, *from the other locations represented in the graph*. The italicised qualification is important, since it indicates that no direct correspondence between visibility graph neighbourhood size and the viewable area from a location can be assumed. This is because (i) the vertex set in the graph may not be uniformly distributed across the landscape, and (ii) many locations not included in the visibility graph may be visible from some or all of the vertices included in the graph. The first difficulty is minimised if the set of vertices are distributed at regular intervals across the landscape. The second of these difficulties (an edge effect) is an unavoidable difficulty of any visibility analysis technique, unless the region analysed is ‘embedded’

in a larger region by extension to the horizon in all directions. Locations in the extended analysis region then contribute to the determination of neighbourhood size (viewshed area) in the visibility graph but are not themselves analysed. That is to say, some regions included in analysis are included to ‘make up the numbers’ in the viewsheds of locations in the central region. The whole viewshed of such ‘peripheral’ locations in the analysis are not themselves considered.

Consider the landscape illustrated in figure 3. Figure 4 illustrates the vertices in a visibility graph generated from a regularly spaced grid across region A, where each

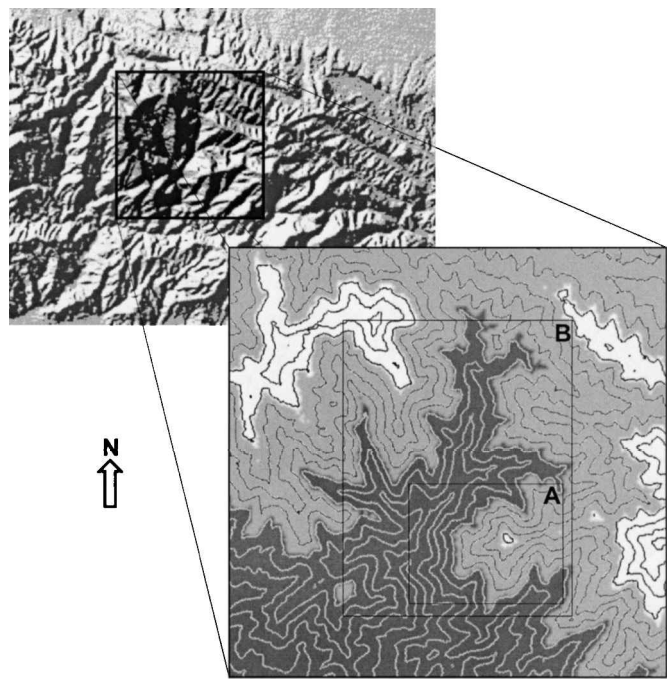


Figure 3. The study area. Analysis results for visibility graphs in region A are presented in figures 4, 5 and 7. Results for region B are presented in figures 8 and 9.

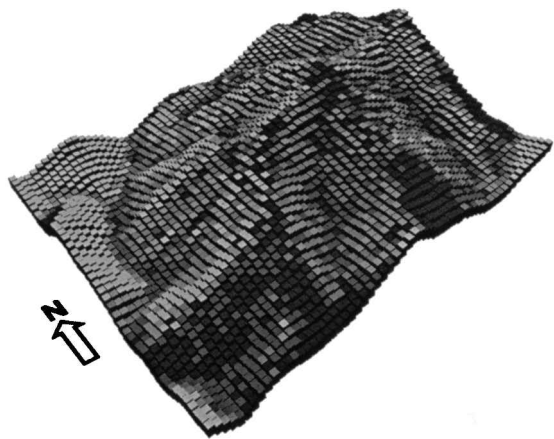


Figure 4. Visibility graph neighbourhood sizes for region A.



vertex has been coloured according to its neighbourhood size—paler greys correspond to larger visibility graph neighbourhoods—and rendered in 2.5-D so that the relationship between landscape and neighbourhood size can be better understood. No clear pattern is observable in this image, although there is a tendency for larger neighbourhood sizes to be available along ridges lines. The largest neighbourhood sizes are associated with the north-western corner of the region. This draws attention to the previously mentioned difficulty associated with the fact that not all the locations visible from each location in the region are included in the visibility graph. As a result the relatively low-lying region in the north-west from which most of the north-western face of the central massif is visible has the largest neighbourhood sizes within this graph.

Viewable area is often mapped by the method of cumulative viewshed analysis which does not rely on a visibility graph, so the above analysis is not novel. However, we now introduce two analytical measures of landscape visibility patterns which are not readily obtained *without* generating the visibility graph, and so may represent a more convincing argument for the approach. Both are based on measurement of the graph itself, and so draw on the wider literature on graph structural analysis (Harary 1969, Haggett and Chorley 1969, Barnes and Harary 1983, Buckley and Harary 1990, Wilson and Beineke 1979, Wassermann and Faust 1994), and in particular on the work of Watts and Strogatz (1998).

#### 4.2.2. Clustering coefficient

First, we consider the *clustering coefficient*, which is a measure based on the sub-graph induced by the neighbourhood of a vertex in the visibility graph— $N_i^*$  in the previously introduced terminology. Using this sub-graph we define the clustering coefficient  $c_i$  as

$$c_i = 2 \cdot \frac{|E(N_i^*)|}{k_i(k_i - 1)} \quad (6)$$

where  $k_i$  is the neighbourhood size as before, and  $0 \leq c_i \leq 1$ . The clustering coefficient is a measure of the extent to which all the lines of sight which *could* exist in the neighbourhood of a location in the visibility graph, *do* exist. If most of the locations visible from a location are mutually visible then  $c_i$  will approach 1. If many of the locations visible from a location are not mutually visible, then  $c_i$  will approach 0. The clustering coefficient for all  $n$  locations in a visibility graph can be calculated in  $O(nk^2)$  time. Note that in the form of equation (6) mutual visibility is implied by the factor of 2 in the expression, since this indicates that the largest number of lines of sight which could be present is  $k_i(k_i - 1)/2$ , which is only the case if all lines of sight are mutual. An equivalent expression where lines of sight are not mutual simply drops the factor of 2, and counts lines of sight in both directions i.e.  $v_j v_i$  is counted as distinct from  $v_i v_j$ . Figure 5 shows the clustering coefficient mapped for the same visibility graph as before. As might be expected, ridge lines and valleys in the region are distinguishable. This is particularly clear in the valley below the heights in the north-west corner of the region. Interestingly, more subtle concavity and convexity is also apparent, in the alternating north-south running ridges south of the central massif. This is what we would expect, since in a perfect ‘bowl’ all locations are mutually visible (high clustering coefficient), whereas at the ‘peak’ of a cone, none of the visible locations are mutually visible (low clustering coefficient). We can also note that the measure is less affected by the edge effects previously observed with

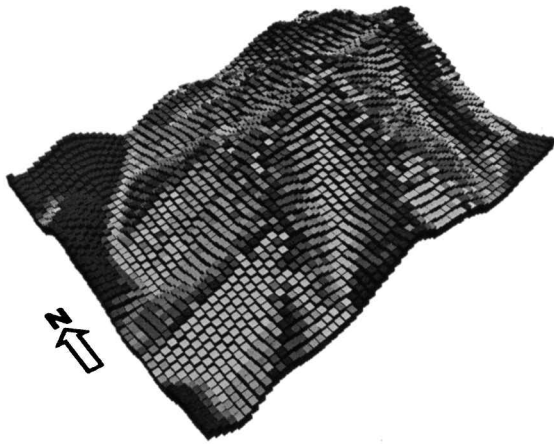


Figure 5. Clustering coefficients for region A.

neighbourhood size. This is because the measure is solely concerned with the interconnections of locations within the area under consideration, and because it is normalised with respect to area, it can be a useful tool to investigate geographic subsystems.

Another, more cognitive interpretation of this measure is that it reflects the extent to which a particular location seems to be 'in' its surroundings, rather than looking down on them from 'outside'. One factor in this interpretation is important to note. Two locations visible from location  $v_i$  may be mutually visible, without the line-of-sight between them being visible along its entire length from  $v_i$ . This may occur if there is some intervening obstacle such as a rock outcrop, or on a larger scale, a hill or mountain. Bearing this in mind, if we consider the impact on the clustering coefficient of introducing a series of differently shaped obstacles onto a flat plane of locations near the obstacle, we can begin to refine our understanding of the measure. Figure 6 shows the clustering coefficient determined at the same resolution for three artificial obstacles (coloured white) on a flat plane. The obstacles in these examples would all appear the same to an observer in corner P of the region shown. This similarity is reflected in the identical distribution of clustering coefficient values seen in the region PQBACW for obstacles (i) and (ii). However, obstacle (iii) produces a different pattern and range of clustering coefficient values in this region. This is because of the lines of sight beyond the back edge of obstacle (iii) in region STU.

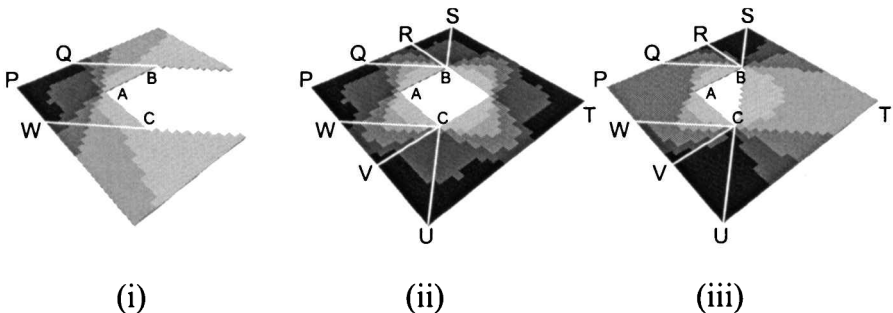


Figure 6. A series of obstacles (indicated by the white region in each case) on a flat plane, and their effect on the clustering coefficient measure. See text for commentary.

These give rise to higher clustering coefficients in PQBACW with obstacle (iii), *even though these lines of sight are not visible from P directly*. To an observer moving around in the environments represented here, this difference would not be apparent until either line QB or WC was crossed, when the differences between the various obstacles would start to become apparent. Similar comments apply to similarities in distribution and differences in range in regions QRB and WCV between cases (ii) and (iii). It is also notable that locations very close to the shorter faces AB and AC of all three obstacles have very similar patterns of the clustering coefficient.

The intriguing point is that in an actual landscape—disregarding prior knowledge from maps—it is the *movement* of an observer which allows her to know the shape and scale of an obstacle. The clustering coefficient, by taking into account secondary visibility relations, may allow some of these perceptual effects to be described from *static* data only. This suggests that the measure may be useful in describing the extent to which the surroundings are perceived *as surroundings* rather than as collections of objects. This is similar to the previously suggested and complementary interpretation of ‘inside’ and ‘outside’ the landscape.

#### 4.2.3. Mean shortest path length

Another possible measure of a graph structure relates to *path lengths* between vertices. A *path* between two vertices  $v_A$  and  $v_B$  in a graph consists of a sequence of distinct edges  $\{v_A v_1, v_1 v_2, \dots, v_x v_{x+1}, \dots, v_{r-1} v_B\}$  between the two. The number of edges in the path is the *path length* in this case equal to  $r$ . The length of the shortest possible path between two vertices is the *distance* between them frequently represented by  $d_{AB}$ . Finding the shortest path and hence the distance between two vertices in a graph is a familiar problem with an algorithmic solution first proposed by Dijkstra (1959). Note that no geographical distance is included in the determination of a graph distance measure—each edge (line-of-sight) is considered of equal weight in this interpretation. If we determine the lengths of the shortest paths from every vertex in a graph to every other vertex, the results can be used as a measure of the *centrality* or *accessibility* of each vertex in the overall graph structure. The simplest way to do this is to sum all the shortest path lengths to all other vertices, to produce a *total path length* measure for each vertex. This is calculable in  $O(n^2 \log n)$  time. In figure 7 the total path length from each vertex to every other is plotted for the same landscape as previously. This measure is difficult to interpret, and shows no clear pattern relative to landscape features. It is notable that, in contrast to the clustering coefficient, a measure of this sort is particularly subject to edge effects due to the need to restrict the visibility graph to some local region in the landscape. The measure does, however, indicate the potential for further analyses based on the visibility graph which relate to all the visibility relations among a whole set of locations in a landscape.

### 5. Discussion and suggestions for further work

The previous section demonstrates the potential usefulness of the visibility graph approach in the analysis of landscape visibility patterns. We now examine some of the issues raised by the method in more detail, and suggest some possible avenues for further research.

#### 5.1. Vertex selection, and scale and resolution issues

The single most problematic aspect of the proposed method is determining an appropriate vertex set for the visibility graph. In the examples shown so far, we have

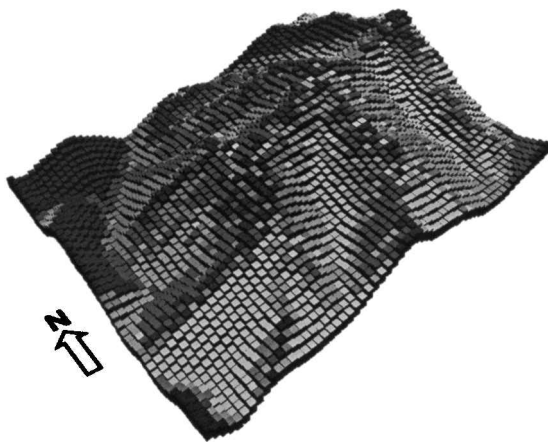


Figure 7. Mean shortest path lengths for region A.

adopted the ‘brute force’ approach of frequent sampling across a landscape. In the cases illustrated, the visibility analysis was carried out on a TIN, which was generated by close fitting to underlying DEM data. The visibility graph vertices in turn were located at the centre points of the pixels in the original DEM, so that the selected locations are as densely distributed as the data on which the landscape representation used was based. This approach has been used simply to demonstrate the possibilities of the approach, and it is not at all clear that it is ideal. Practically, such an approach leads rapidly to very large visibility graphs. A DEM with data points every (say) 50 m will yield 400 vertices per square kilometre. Very large visibility graphs result, and given the  $O(n^2)$  complexity of any algorithm to generate a graph the computational requirements rapidly escalate—even allowing for the fact that the computational complexity increases approximately linearly with  $n$  over (much) larger regions due to earth curvature and atmospheric effects. In the examples above for region A, there are 2820 vertices in the graph. With visibility graph generation implemented in a GIS macro-language, producing this graph is an overnight process. Much improved performance could be achieved by implementing the required routines in GIS source code, and providing the required graph data structures in standard GIS data layers. Nevertheless, there are always likely to be limits to the size of visibility graph which can be generated, analysed, and stored efficiently.

This raises the question of whether it is possible to choose the graph vertex set more intelligently (so that  $n \ll s$ ) in order to reduce the computational requirements. Unfortunately, this is a difficult question to answer—since it is difficult to know *a priori* which locations in a region give the best ‘overview’ of the visibility pattern, without examining all the possibilities. We have begun to investigate this issue by examining the pattern of measures on the visibility graphs produced for different ‘sub-samples’ of a region. Preliminary results are shown in figure 8. This image has been produced by combining clustering coefficient results from four different visibility graphs generated for region B in figure 3. Each of the four visibility graphs was generated by placing vertices at every second DEM data point in both directions. The graphs are thus offset from each other by one DEM cell in one or both cardinal directions, and analysed completely independently. The results have been recombined only by plotting the results in one image. It is notable that the overall pattern of the

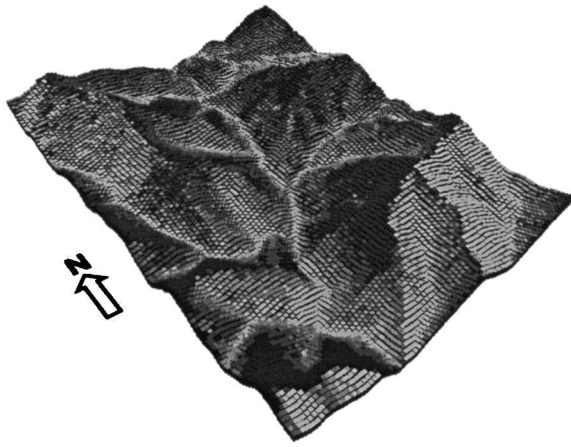


Figure 8. Clustering coefficient for four ‘interleaved’ visibility graphs generated in region B. This image shows that each sub-sampled graph reveals a similar overall pattern in this measure.

clustering coefficient measure is similar for all four graphs. This suggests that, for this measure at least, less intensive sampling of the region can give a good overall impression of the visibility patterns. As an aside, it is notable that the characteristics of the clustering coefficient measured over this larger region are clearer than in figure 6. The river valley is particularly clearly picked out as highly clustered in this image.

As an alternative approach to the question of the adequacy or suitability of different selections of visibility graph vertices, figure 9(i) shows results for region B with 1000 randomly selected locations, along with results from 2610 vertices at

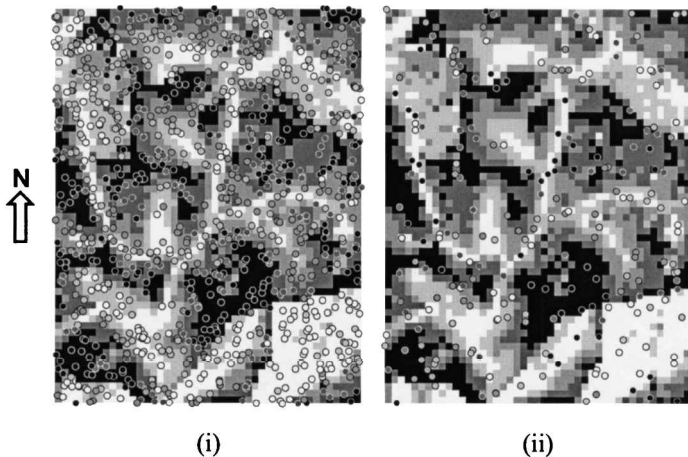


Figure 9. Comparisons of the clustering coefficient measure for different vertex sets in the same landscape: (i) a 1000 vertex random graph (foreground blobs), and a 2610 vertex regular graph (background grid); and (ii) a 228 vertex graph including local peaks and pits together with some random vertices (foreground blobs), and a 2610 vertex regular graph (background grid). All data sets have been classified into quintiles with pale colours representing higher values.

regularly spaced intervals (this is one of the graphs in figure 8). These results are presented as a single 2-D image with the clustering coefficient for the randomly selected vertices plotted as grey-scaled 'blobs' against a background grid coloured according to the clustering values for the regularly spaced locations. The grey scale colours are both in five classes based on quintiles with respect to each data set. In general, there is a good visual 'fit' between the classification of the randomly distributed and the regularly spaced locations. In this example the randomly selected graph can be produced and analysed in around one-sixth of the time required for the more complete representation of the regularly spaced graph. Since the regularly-spaced graph shown includes only one-quarter of all the DEM grid cells in region B, this result suggests that there is considerable potential for improved selection of visibility graph vertices.

Potentially, a better approach than determining some minimal resolution, or vertex set size which provides a consistent 'picture' of the landscape is to choose locations which have topological significance. The most obvious possibility in this connection is to use the significant points in a surface network (Pfalz 1976), such as peaks, pits and saddle points, and points along ridge and channel lines. It is likely that such points will themselves have extreme visibility properties, so it is unlikely that a set based *solely* on such features will be sufficient. Additional, more 'typical' locations are required. A set of vertices made up of 89 peaks and pits determined from simple local analysis of elevations, together with an additional 139 randomly chosen points has been analysed. In figure 9(ii), the clustering coefficient for this graph is plotted, as in figure 9(i), against a background showing the clustering coefficient from a regularly spaced 2610 vertex graph. Much more divergence in the pattern of values is evident in this case. In particular, vertices in the valleys in region B may have a high clustering coefficient (coloured white) in the grid-based visibility graph, but a low clustering coefficient (coloured black) in the sparser graph. It seems likely that the relatively small areas visible from such locations are inadequately sampled in the sparser graph. It is clear from this that there is considerable scope for further research into the issue of determining a representative set of locations for analysis of the visibility patterns in a landscape, and further, that locations of obvious topological importance may not be the most significant in terms of visibility relations.

## 5.2. Other possible analytical measures of the visibility graph

We have only considered three possible measures derivable from a visibility graph; there are many more. Any graph measure relates to some aspect of the structure of the graph, and it is useful to think of these as falling into various categories. The extensive graph metric literature in social network analysis (Wassermann and Faust 1994) considers graph analysis methods as falling into three distinct categories: *centrality*, *cohesive sub-groups*, and *structural equivalence*. The first of these is obvious and has already been demonstrated by the total shortest path length measure above (and to a lesser extent by the neighbourhood size measure), although there are other possible approaches (Nieminen 1974, Freeman 1979, Stephenson and Zelen 1989).

Cohesive sub-groups in a graph are sets of vertices which are particularly strongly interrelated. This is clearly related to the clustering coefficient measure described above. However the concept may also be extended in various ways. A *clique* in a graph is a set of vertices each of which is connected to every other, and clique analysis consists of identifying the various large cliques which are present in a graph.

Cliques in a visibility graph may refer to perceptually interesting properties of particular areas in a landscape, since they would identify sets of points all of which are mutually visible. Looser definitions of cohesive sub-groups also exist (Borgatti *et al.* 1990), and may also be of interest. The most likely interpretation of cohesive sub-groups in a visibility graph is that they would represent regions in an environment within which there is some sense of 'enclosure'. The size of cohesive sub-groups which would be of interest is closely related to the process by which vertices in the graph are chosen, since sets of just three mutually visible locations are likely to be very common in all but the sparsest of vertex sets. A related definition which may be of interest here is the *neighbourhood kernel*  $K_i$  at  $v_i$  in a graph, which is the set of vertices all of which are visible from the vertices in  $N_i$ . Like the clustering coefficient  $c_i$ ,  $K_i$  relates to the experience of movement in an environment since it represents a set of locations which will remain in view as a person leaves location  $v_i$ .

Structural equivalence in a graph is best understood in terms of family trees in which sets of siblings are structurally equivalent—that is, they have similar relations to other elements in the graph. The interest in equivalence groups in social networks is obvious (Borgatti and Everett 1989, Sparrow 1993). Whether or not sets of equivalent vertices in a landscape visibility graph would have any particular meaning is unclear, although it is possible that such sets might consist of similarly located points, and refer to perceptual qualities of landscape. Again there is a great deal of scope for further research on these issues.

All of the measures discussed in the previous paragraph may be derived *locally* for each vertex in a graph, or *globally* for a whole graph, in which case they describe the extent to which a particular graph exhibits properties such as 'centralisation' or 'cliquishness'. Such characterisation is of potential interest in understanding the overall perceptual qualities of different landscapes. Additionally, the extent to which different measures correlate with one another in different landscapes is potentially of interest. Thus, we would normally expect high clustering coefficients to be associated with smaller neighbourhood sizes, but this relationship is likely to be different in different types of landscape. Equally we would expect some measures to be more or less correlated with the elevation at different locations, and variations in this relationship between different landscapes may be of interest.

### 5.3. Directional visibility graphs and measures

One evident difficulty of the measures presented so far is that they relate to all the lines-of-sight in all directions from the locations in an environment. While this is interesting information to study, it is not related in any simple way to the information received by an individual moving around in the corresponding real-world environment. A person moving in a real environment directs her attention in particular directions as she moves around the environment, so that only a 'path-filtered' sub-graph of the visibility graph is experienced. Investigation of how the paths followed by individuals are related to the properties of the full visibility graph may be of interest. Equally, generation of 'path-centred' visibility graphs and appropriate measures thereon, seems likely to be a rewarding direction for research.

## 6. Conclusions

It is clear from the previous section that there is a great deal of scope for further research on the visibility graph concept in landscape visibility analysis. The examples in this paper suggest that further research may be fruitful, especially for the expanding

field which attempts to describe and understand environmental cognition using GIS. The approach may also have relevance to archaeological studies of historical landscapes. A specific example of the use of a clustering coefficient-type measure might be to ascertain how self-contained is the environmental impact in planning applications. However, the most important point arising from this work is that the visibility graph approach, by storing all of the information uncovered in exhaustive viewshed analyses, offers considerable potential for furthering our understanding of the visibility characteristics of spatial environments.

### **Acknowledgments**

This research was carried out while David O'Sullivan held an Engineering and Physical Sciences Research Council studentship at the Bartlett Faculty for the Built Environment, University College London, administered through the Centre for Advanced Spatial Analysis.



## References

- Amidon, E. L., and Elsner, G. H., 1968, Delineating landscape view areas: a computer approach, Forest Research Note PSW180, US Department of Agriculture, Berkeley CA.
- Baldwin, J., Fisher, P., Wood, J., and Langford, M., 1996, Modelling environmental cognition of the view with GIS. Presented at *Third International Conference on Integrated GIS and Environmental Modelling*, University of California at Santa Barbara, CD Rom.
- Barnes, J. A., and Harary, F., 1983, Graph theory in network analysis. *Social Networks*, 5, 235–244.
- Benedikt, M., and Burnham, C. A., 1981, Perceiving architectural space: from optic arrays to isovists. In *Persistence and Change: Proceedings of the First International Conference on Event Perception*, edited by W. H. Warren and R. E. Shaw (Hillsdale NJ: Lawrence Erlbaum Associates), pp.103–114.
- Borgatti, S. P., and Everett, M. G., 1989, The class of all regular equivalences: algebraic structure and computation. *Social Networks*, 12, 337–358. Borgatti, S. P., Everett, M. G., and Shirey, P. R., 1990, LS sets, lambda sets, and other cohesive subsets. *Social Networks*, 11, 65–88.
- Buckley, F., and Harary, F., 1990, *Distance in graphs* (Redwood City: AddisonWesley).
- Burrough, P. A., 1986, *Principles of Geographical Information Systems for Land Resources Assessment*. Monographs on Soil and Resources Survey (Oxford: Clarendon Press).
- Camp, R. J., Sinton, D. T., and Knight, R. L., 1997, Viewsheds: A complementary management approach to buffer zones. *Wildlife Society Bulletin*, 25, 612– 615.
- Davidson, D. A., Watson, A. I., and Selman, P. H., 1993, An evaluation of GIS as an aid to the planning of proposed developments in rural areas. In *Geographical Information Handling: Research and Applications*, edited by P. M. Mather (London: Wiley), pp. 251–259.
- de Berg, M., van Kreveld, M., Overmars, M., and Schwarzkopf, O., 1997, *Computational geometry: algorithms and applications* (Berlin, New York: Springer).
- De Floriani, L., and Magillo, P., 1994, Visibility algorithms on triangulated terrain models. *International Journal of Geographical Information Systems*, 8, 13–41.
- De Floriani, L., Marzano, P., and Puppo, E., 1994, Line of sight communication on terrain models. *International Journal of Geographical Information Systems*, 8, 329– 342.
- Dijkstra, E. W., 1959, A note on two problems in connection with graphs. *Numerische Mathematik*, 1, 269–271.
- Fisher, P. F., 1991, First experiments in viewshed uncertainty - the accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing*, 57, 1321–1327.
- Fisher, P. F., 1995, An exploration of probable viewsheds in landscape planning. *Environment and Planning B: Planning & Design*, 22, 527– 546.
- Fisher, P. F., 1996, Extending the applicability of viewsheds in landscape planning. *Photogrammetric Engineering and Remote Sensing*, 62, 1297–1302.

- Freeman, L. C., 1979, Centrality in social networks: Conceptual clarification. *Social Networks*, **1**, 215-239.
- Gibson, J. J., 1979, *The Ecological Approach to Visual Perception* (Boston, MA: Houghton Mi, en).
- Goodchild, M. F., and Lee, J., 1989, Coverage problems and visibility regions on topographic surfaces. *Annals of Operations Research*, **18**, 175-186.
- Haggett, P., and Chorley, R. J., 1969, *Network Analysis in Geography* (London: Edward Arnold).
- Harary, F., 1969, *Graph Theory*, Addison Wesley series in Mathematics (Reading, MA: AddisonWesley).
- Jiang, K., Seneviratne, I. D., and Earles, S. W. E., 1996, Three dimensional shortest path planning in the presence of polyhedral obstacles. *Proceedings of the Institute of Mechanical Engineers Part C- Journal of Mechanical Engineering Science*, **210**, 373-381.
- Jiang, K., Seneviratne, I. D., and Earles, S. W. E., 1999, A shortest path based path planning algorithm for nonholonomic mobile robots. *Journal of Intelligent and Robotic Systems*, **24**, 247-366.
- Lake, I. R., Lovett, A. A., Bateman, I. J., and Langford, I. H., 1998, Modelling environmental influences on property prices in an urban environment. *Computers, Environment and Urban Systems*, **22**, 121-136.
- Lee, J., and Stucky, D., 1998, On applying viewshed analysis for determining least cost paths on digital elevation models. *International Journal of Geographical Information Science*, **12**, 891-905.
- Llobera, M., 1996, Exploring the topography of mind: GIS, social space and archaeology. *Antiquity*, **70**, 612– 22.
- Lynch, K., 1976, *Managing the Sense of Region* (Cambridge MA: MIT Press).
- Nieminen, J., 1974, On the centrality in a graph. *Scandinavian Journal of Psychology*, **15**, 322-336.
- Pfalz, J. L., 1976, Surface networks. *Geographical Analysis*, **8**, 77– 93.
- Sparrow, M. K., 1993, A linear algorithm for computing automorphic equivalence classes: the numerical signatures approach. *Social Networks*, **15**, 151-170.
- Stephenson, K., and Zelen, M., 1989, Rethinking centrality: methods and examples. *Social Networks*, **11**, 1-37.
- Tandy, C. R. V., 1967, The isovist method of landscape survey. In *Symposium: Methods of Landscape Analysis*, edited by H. C. Murray (London: Landscape Research Group), pp. 9-10.
- Turner, A., Doxa, M., O' Sullivan, D., and Penn, A., 2001, From isovists to visibility graphs: a methodology for the analysis of architectural space. *Environment and Planning B: Planning & Design*, **28**.
- Wang, J. J., Robinson, G. J., and White, K., 1996, A fast solution to local viewshed computation using grid-based digital elevation models. *Photogrammetric Engineering and Remote Sensing*, **62**, 1157–1164.
- Wassermann, S., and Faust, K., 1994, *Social Network Analysis, Methods and Applications* (Cambridge University Press: Cambridge).
- Watts, D. J., and Strogatz, S. H., 1998, Collective dynamics of 'small-world' networks. *Nature*, **393**, 440– 442.

- Wheatley, D., 1995, Cumulative viewshed analysis: a GIS based method of investigating intervisibility and its archaeological application. In *Archaeology and GIS: A European Perspective*, edited by G. Lock and Z. Stancic (London: Taylor and Francis), pp. 171–185.
- Wilson, R. J., 1996, *Introduction to graph theory* (Harlow, England: Longman).
- Wilson, R. J., and Beineke, L. W. (editors), 1979, *Applications of Graph Theory* (London: Academic Press).