Visualization of Large Volumes of Information Using Different Representations

Maria Beatriz Carmo
FCUL-DI, Bloco C5, Campo Grande
1700 Lisboa, Portugal
e-mail: bc@di.fc.ul.pt

João Duarte Cunha
LNEC-CI, Av. do Brasil 101
1799 Lisboa Codex, Portugal
e-mail: jdc@lnec.pt

Abstract

The visualization of large volumes of information requires mechanisms to reduce the amount of information to be displayed. In this paper we propose a technique to combine the usage of filters, like the degree of interest function [Furnas86], with different, scale dependent, representations. This eliminates the need to create new representations or use graphical attributes to show the degree of interest of each object. This approach is specially useful when the number of variables is large, or the information density is too high.

The method can be complemented with zoom mechanisms and distortion of the position of the objects in space, taking advantage of the reduction of the area occupied by the objects with less interest to increase the area available for those with a higher degree of interest, normally located near the focus.

1. Introduction

To visualize large volumes of information we need mechanisms to reduce the amount of information displayed: use of filters, identification of different levels of interest, reduction of the number of variables represented, use of multiple representations with amount of detail varying with the scale of representation or the degree of interest of the object.

In this paper we propose a technique to combine the use of degree of interest functions [Furnas86] with different, scale dependent, representations.

In order to generate a fisheye view of a structure the degree of interest function was introduced by Furnas [Furnas86]. This function is useful to eliminate non relevant information and also to determine the type of representation [Sarkar92], [Fairchild93].

The degree of interest function assigns to each point in the structure a number which is a measure of the user's interest in seeing that point. The value of the function in a point **x** depends on the *a priori* importance of the point, **API(x)**, and on the distance between **x** and the current focus, **y**: **DOI(x|y)=API(x)-D(x,y)**. The distance may correspond to euclidean distance or to a semantic distance. Only the points whose DOI is over a given threshold, **k**, are displayed. This means that the amount of information effectively displayed depends on the value of this threshold.

In this paper we consider the degree of interest function even if a focus is not specified, assuming in this case that the distance function is equal to zero everywhere, i.e., the DOI function exactly matches the *a priori* importance.

An object can have multiple representations with different levels of detail. When dealing with real world objects, the most detailed representations correspond to exact shape, dimension and position of the objects, while the less detailed ones may be conventional symbols. When visualizing abstract information, all the representations are symbolic and it is easier to change the position of the objects in order to make good use of the display area.

In any case we need a criterion to select the representation to be used. This criterion must take into account the scale of the representation, but also the density of information and the user's interest in visualizing a particular object. In order to obtain intelligible representations in the smaller scales, it is necessary to adapt the representation by reducing the level of detail, or eliminating some objects or grouping several objects into a unique representation.

If there is a large amount of information to visualize, the representation must be simplified and part of that information eliminated. Semantic criteria can be used to select the objects to eliminate. These criteria can be based on degree of interest functions. These functions can be used, not only to eliminate information, but also to establish different levels of interest. Objects with less interest can be associated with less detailed representations, normally reserved to smaller scales. In this way there is no need to create new representations or to choose attribute values in order to identify different levels of interest.

This approach is specially advantageous if there is a lot of variables to represent or the density of information is too high, because it reuses previously defined representations and also simplifies the representation of less relevant information. The ratio

quantity of relevant information / quantity of represented information

increases, since the amount of relevant information is the same, but the total amount of information displayed is reduced.

2. Related work

In [Sarkar92] fisheye views are used to visualize graphs. The position, size and amount of detail of each vertex of the graph are calculated using the *a priori* importance of each vertex and its distance to the focus. The method was applied to a particular type of abstract objects, graphs.

In [Stonebr.93] there is a hierarchy of representations with different levels of detail. Each representation has two associated constants: the maximum and minimum screen window sizes, specified in the application coordinate system, for which the representation is appropriate. This approach is used in a system for database support of scientific visualization applications.

In [Frank94] a hierarchy of representations, multi-scale tree, is used to organise the visualization of an object according to the scale of representation. As the scale grows, the level of detail increases. This approach is proposed in the scope of cartographic generalization.

Both Stonebraker et al.[Stonebr.93] and Frank and Timpf [Frank94] describe methods to deal with representations with different levels of detail for objects with spatial reference, they do not consider the inclusion of semantic criteria. In both cases the objects can be represented with a wide range of representations, from very complex ones to simplified symbols.

A different approach is presented in [Fairchild93]. Each object type is associated with a list of representations and the visualization is determined by the value of the degree of interest function of a particular object, without any restrictions derived from the scale of representation. This method is applied to abstract information, which is less restrictive in terms of positioning and dimensioning of the objects, as pointed out above.

The Virtual Reality Modelling Language [VRML96] provides a means of creating, viewing and interacting with 3D scenes. It is possible to build a list of representations for an object with various levels of detail or complexity associated with a list of distances from the object to the user. The appropriate version of the object is chosen automatically based on the distance from the user.

A VRML file contains information about the graphical representation of the object and the interaction between objects and the user. But it does not include semantic information about the object. So, all possible variations of the representation must be established previously.

3. Use of multiple representations

For each class of objects a list of representations is provided, each one with an associated constant. This constant is the minimum scale factor of the window-viewport transformation that keeps the representation intelligible. For smaller scales the representation will be too small to be meaningful. For larger scales the representation has too few details but is perfectly intelligible and can be used for less relevant objects to simplify the visualization.

In this way we build a hierarchy of representations based on the smallest scale factor allowed for each one. As the scale factor increases, representations showing more details will be available. For scales smaller than the scale associated with the less detailed representation, the objects would not be represented.

Adjacent representations in the list can differ in level of detail or in type of representation. For instance, when dealing with georeferenced information, the same object can be represented as a symbol in a smaller scale, and its exact shape, dimension and position used in larger scales. This point would be further developped in section 5.

We want to choose the representation taking into account not only the scale factor, but also semantic criteria. Since it would not be correct to use representations which are not suitable for a given scale, representations defined for scales greater than the current scale are always excluded. However, less detailed representations can be selected to visualize information with less interest.

4. Multiple representations and degree of interest functions

As mentioned above, we want to have the possibility of selecting representations, not only according to the scale, but also including semantic criteria expressed by the degree of interest function.

A way to achieve this goal is to divide the range of values of the degree of interest function (minDOI - maxDOI) by the number of available representations, where the minDOI is the threshold value. Objects belonging to the upper interval will be displayed with the representation associated with the current scale[(*)] or current representation for short. Objects belonging to lower intervals will be displayed using less detailed representations.

The worst case occurs when the current scale is associated with the simplest representation: all the objects of the same type will be visualized with the same representation. On the other hand, when the current scale corresponds to the most detailed representation, all the available representations can be used.

Let us assume that we have **n** representations, **1** through **n**, with the $\mathbf{n}^{\mathbf{th}}$ representation corresponding to the largest scale and the $\mathbf{1}^{\mathbf{st}}$ to the smallest one. The range of values of the degree of interest function will be split into **n** intervals, $\mathbf{s_1}$ through $\mathbf{s_n}$. If the representation associated with the current scale is **k**, then objects belonging to the upper interval, $\mathbf{s_n}$, will be displayed with representation **k**, objects in the interval $\mathbf{s_{n-1}}$ will be displayed with representation **k-1,** and so on. Objects in the intervals $\mathbf{s_{n-k+1}}$ to $\mathbf{s_1}$ will be displayed with representation **1**, the one with less detail. In other words, objects whose degree of interest, **DOI**, lies in the interval **[minDOI,maxDOI[** are displayed using representation

current representation - delta
where
delta=n-[(DOI-minDOI) div ((maxDOI-minDOI)/n)+1]

For **DOI=maxDOI**, delta=0 by convention and the current representation is used.

This approach does not make an optimal use of the usable subset of representations, **1** through **current**, since more than one interval can be mapped to representation **1**. To provide a more uniform distribution of the objects among the **k** usable representations, the range of values of the degree of interest function can be divided by **k** instead of the total number of representations, **n**.

5. Symbolic vs. exact representations

When representations with exact shape, dimension and position become too small to be intelligible, they must be replaced by symbolic representations. Therefore the list of representations of each object type has two kinds of representations: exact and symbolic. When the choice of representation depends only on the scale, there is no conflict between the two kinds of representation since they cannot be used simultaneously.

But when semantic criteria are also used, the two kinds of representation can coexist for the same class of objects. Since symbolic representations are conceived to replace the exact ones at smaller scales, they are usually larger, in the user coordinate system, than their exact counterparts. Therefore, if symbolic representations are defined in the user coordinate system and subjected to the standard window-viewport transformation, it can very well happen that objects with a smaller DOI are displayed with symbols larger than the exact representation of other objects with greater DOI.

To keep control of the dimension of symbols in the image, symbolic representations must be defined in an intermediate coordinate system and subjected to an independent window-viewport transformation. However, the indication of the scale associated with each representation must always be made relative to the standard window-viewport transformation.

The dimensions of the symbolic representations must be specified taking into account the dimensions of the exact representations they are going to substitute.

6. Example

To illustrate the proposed approach, let us consider the display of public buildings in a town plan. In this example four symbolic representations are used to represent the buildings. For larger scales exact representations should be included, as described above.

The simplest representation uses a single letter to indicate the type and location of the building. The remaining three symbols are obtained by drawing up to three concentric circumferences around the basic symbol.

[(*)] The current scale is the scale corresponding to the image currently seen by the user. When using zoom techniques that keep context information like the one mentioned in section 6., the current scale may be different for each object.

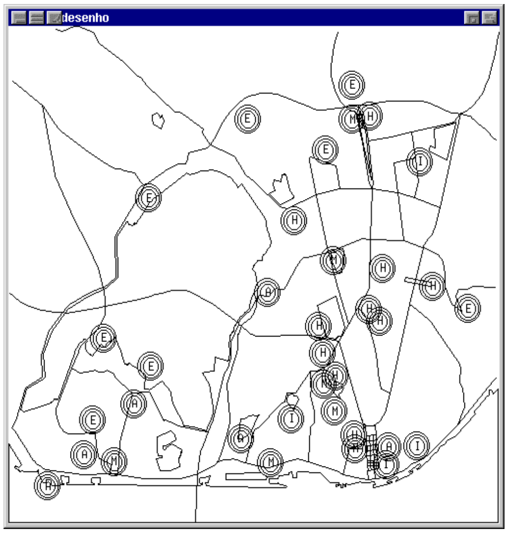

Fig.1. Representation determined by scale

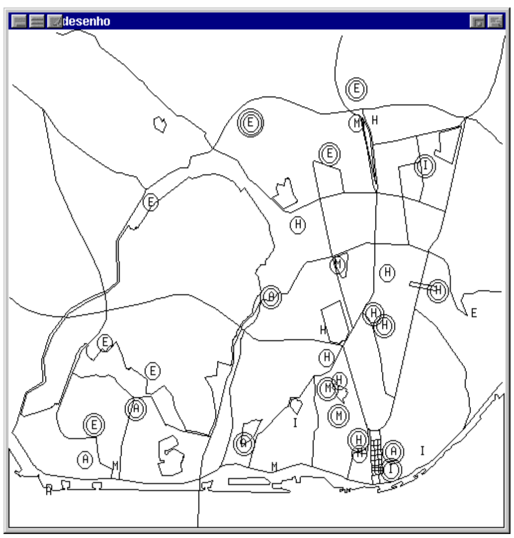

Fig.2. Representation determined by scale and *a priori* importance

Fig.1 shows all the public buildings with the representation determined by scale, i.e. without involving any semantic criteria. In Fig.2 the representation of each building is modified to reflect its *a priori* importance. Adding a focus (Fig.3) the value of the degree of interest function changes and several buildings have their representations modified. Finally, Fig.4 shows the effect of zooming the area near the focus. The zoom model used is the one proposed in [Robertson93] and [Cunha95]. In this model, the zoomed area is magnified without distortion but the global context is preserved. The scale factor is increased in the zoomed area and decreases uniformly when going from the border of the zoomed area to the periphery.

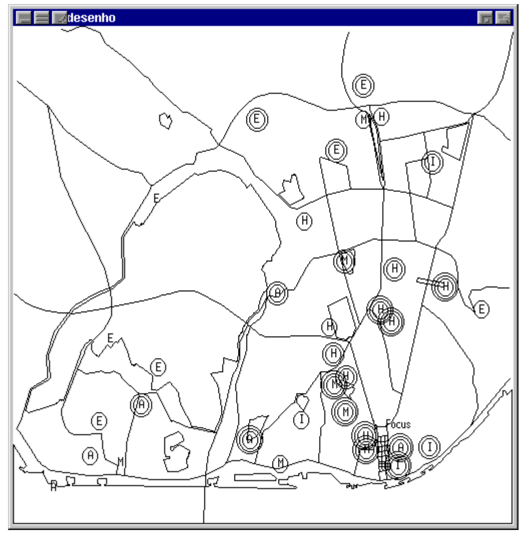

Fig.3. After selecting a focus the representation is determined by the scale and degree of interest function

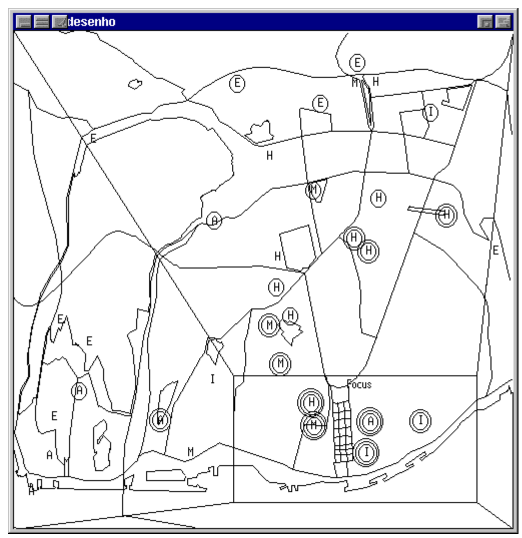

Fig.4. Zooming in near the focus

7. Conclusions

The combination of the degree of interest function with the selection of representations according to scale provides a powerful mechanism to improve the visualizations by:

- eliminating non relevant information;
- simplifying the representation of less relevant objects;
- enhancing the relevant information;
- keeping the visualization intelligible.

Including zoom tools that provide local detail and global context, selected areas can be amplified to occupy the space provided by the elimination or simplification of less relevant information.

References

[Cunha95] J.D. Cunha, M.B. Carmo, “Modelo da Pirâmide Truncada”, *VII Encontro Português de Computação Gráfica*, Monte da Caparica, 1995, pp. 35-46

[Fairchild93] K.M. Fairchild, “Information Management Using Virtual Reality-Based Visualizations”, *Virtual reality applications and explorations* edited by Alan Wexelblat, Academic Press, 1993, pp.45-74

[Frank94] A.U. Frank, S. Timpf, “Multiple Representations for Cartographic Objects in a Multi-Scale Tree - an Intelligent Graphical Zoom”, *Computer and Graphics*, vol.18, nº6, 1994, pp. 823-829

[Furnas86] G. Furnas, “Generalized Fisheye Views” *Proceedings CHI'86*, 1986, pp. 16-23

[Robertson93] G. Robertson, J.D. Mackinlay, J.D., “The Document Lens”, *Proceedings UIST’93*, 1993, pp. 101-108

[Sarkar92] M. Sarkar, M.H. Brown, “Graphical Fisheye Views of Graphs”, *CHI ‘92 Conference Proceedings*, May 1992, pp. 83-91

[Stonebrak.93] M. Stonebraker, J. Chen, N. Nathan, C. Paxson, A. Su, J. Wu, "Tioga: A Database-Oriented Visualization Tool", *Proceedings Visualization '93*, G.M. Nielson, D. Bergeron (eds), 1993, pp.86-93

[VRML96] “The Virtual Reality Modelling Language Specification”, v2.0, 1996, http://vrml.sgi.com /moving-worlds/spec/part1/concepts.html