

IVPrototype — an Information Visualization Prototype

Maria Beatriz Carmo¹, João Duarte Cunha², Ana Paula Cláudio¹
¹FCUL-DI, Bloco C5, Piso 1, Campo Grande 1749-016 Lisboa, Portugal
²LNEC-CI, Av. do Brasil 101, 1700-066 Lisboa, Portugal
bc@di.fc.ul.pt, jdc@lnec.pt, apc@di.fc.ul.pt

Abstract

IVPrototype is an Information Visualization Prototype for the visualization of both geometrical and non-geometrical data. Our goal is to provide a generic system that automatically generates data visualization and dynamically adapts the interface according to the data in order to control the amount of information displayed. The information is organized in one or several information groups, which aggregate semantically related information. Icons and coloured surfaces are the visual mappings provided. Other features of this prototype are: a data independent architecture; several filtering mechanisms; use of degree of interest functions; automatic selection of icons based on scale factor and semantic criteria.

1. Introduction

Creating visualizations is a complex and time consuming process that requires knowledge about graphical software and interaction techniques. Moreover all this work is seldom re-used in other applications. So generic systems that can be used with different datasets are very useful mainly for those who are not graphical experts. By identifying common requirements in a wide range of cases it is possible to develop generic systems.

Three main types of generic systems can be identified: those, like vtk [15], that provide a kernel of routines to be included in the users' application; those, like AVS [9] or IBM's Data Explorer, that supply a set of modules which can be linked using visual programming to generate the visualization pipeline; and those, like Wavefront's Data Visualizer or Amtec's Tecplot, that allow the user to interactively select the visualization for the data.

These systems are typically of the Scientific Visualization field. In the Information Visualization domain there are some examples of generic visualization systems. For instance, systems like VANISH [11] or

Telescope [2] supply a set of abstract classes that define the behaviour of the system components. Some built-in classes are provided and the user can extend those classes to adapt to his/her own data.

On the other hand, there are systems that do not require any programming effort by the user. This is the case with the IVEE system [1] that provides the automatic creation of dynamic queries applications attaching to each object in a database relation a graphical object.

Our goal is the development of a generic visualization system to automatically create an interface for the visualization of abstract data but also allowing the inclusion of geometrical data. Furthermore, special emphasis is given to filtering mechanisms in order to reduce the amount of information displayed. This includes the use of the degree of interest function [10] to support semantic criteria to either suppress information or select different graphical representations.

The concept of information group is a basic feature of the prototype to organize semantically related information. An information group associates a set of attributes with a spatial reference and one or more graphical representations.

In section 2 the data types handled by the prototype are identified; section 3 describes the information group concept; in section 4 available graphical mappings are detailed; section 5 is devoted to interaction mechanisms; and finally, in section 6 several examples are discussed.

2. Data types

A main aspect in a visualization system is the type of data that can be treated by the system. We concentrate on tabular data, i.e., data that can be organized in data tables. This type of data plays a central role in the reference model for Information Visualization proposed by Card et al. [3]. The authors argue that in a pre-processing step the data can be transformed into tabular data and what is really important in the visualization process is the mapping from data tables to visual structures.

Sometimes it is useful to combine geometrical data and non-geometrical data. So the prototype allows also the visualization of geometrical data. This type of data has an intrinsic spatial reference and is described by a list of graphical primitives.

To make the prototype independent of the data to be represented, a file format was defined. Different data formats can be dealt with by providing a conversion module for each one. Non-geometrical data, i. e., numerical or string data, and their spatial reference are described in a data file. The geometrical data and the representations for non-geometrical data, i. e., the graphical data, are defined in a different file called mapping file.

3. Information groups

To visualize information we need for each element a graphical representation and a location in space. The concept of information group (or information class as defined in [5]) was introduced as a convenient mechanism to combine a set of attributes associated to a given spatial reference with a list of graphical representations [5]. Information groups may contain geometrical or non-geometrical data or combine both types of data. Information groups with geometrical data only are called geometrical information groups. The groups with non-geometrical data, which can possibly include geometrical data, are called global information groups. In this case the display of geometrical data cannot be separately controlled.

The spatial reference is represented as a grid composed of locations and implicit or explicit connections. Following the grid classification proposed in [16], there are four types of grids available: Cartesian, regular, structured and unstructured grids. Only unstructured grid connections need to be explicitly indicated. Otherwise connections are implicitly defined based on the number of points in each axis direction. However, connections are only required when a surface is used to represent the data. Attributes of the information groups can be used to define grid positions. If the data does not include any spatial information or if none of the attributes is chosen to define the grid, Cartesian or regular grids are adopted. Geometrical information groups do not require a grid, because this data includes the information that is necessary to be displayed. When geometrical and non-geometrical data are combined in the same information group, they do not need to share the same spatial reference, in the sense that they do not have to share the same locations. Nevertheless, they must be defined in the same coordinate system.

Besides spatial reference, non-geometrical data needs a graphical representation that is described in the list of representations of the corresponding information group. This list includes one or more representations with

different levels of detail and one of these representations will be displayed in each point of the grid. Each representation has a range of allowed representation scales and is composed of a list of graphical primitives. The dimension and colour of graphical primitives may depend on the values of the group attributes.

Geometrical data is described by a list of geometries that includes the locations of their elements. In fact, this list is similar to a representation list. It is also composed of graphical primitive lists, but is treated in a different way by the degree of interest function [5].

A degree of interest function, as defined in [10], is used to deal with semantic criteria. The values of one of the attributes of the information group can be used as *a priori* importance values. This attribute is called the API attribute. For geometrical information groups the API attribute qualifies each representation of the list of geometries. For global information groups the API attribute establishes the importance of each point of the grid. In this case, the elements of the list of geometries will not be affected by the degree of interest function and their visibility depends exclusively on the current scale of representation. To sum up, a global information group can be composed of several attributes, one of them being the API attribute, a grid, a list of representations and, optionally, a list of geometries. A geometrical information group has a list of geometries and can include attributes to define the API attribute. The information groups are defined in the data file.

4. Mappings

Another important feature of a visualization system is the kind of visualizations provided. Two different visual mappings are available in the prototype for global information groups: icons and coloured surfaces.

In the first case, each element in an information group, which is associated with a point of a grid, is represented by one of the representations included in the list of representations of its group. The representation is chosen according to the current representation scale and to semantic criteria expressed by the degree of interest function [4]. In this way less detailed representations are selected to visualize information with a lower degree of interest. If a mapping file, where the representation lists are defined, is not provided or if some of the information groups do not have a representation list, a default graphical representation is available for global information groups. Elements in the lists of geometries are displayed if the current scale belongs to its range of allowed scales. Geometrical information groups are only displayed when the visualization by icons is selected. In this case, the elements of the list of geometries are displayed taking into account the current scale and the degree of interest function.

In the second case, a three-dimensional coloured surface is displayed. This surface includes the points of a two dimensional grid or a cutting plane of a three dimensional grid. The colour is calculated according to one of the attributes of the information group. Two colouring methods are available: with or without interpolation between adjacent values. Another attribute or occasionally the same can be used to define the relief. This type of visual mapping is only available for groups with a grid defined, i. e., global information groups. To construct the coloured surface the grid connections are needed. As already stated, for Cartesian, regular and structured grids the connections are implicitly defined, but for unstructured grids the connections must be explicitly provided. So for three-dimensional grids they have to be supplied in the data file, but for two-dimensional grids, if they are not provided in the data file, a Delaunay triangulation is generated. The algorithm implemented, a randomized incremental Delaunay insertion algorithm, is described in [12].

By default the representation by icons is displayed. The user can interactively switch between the two types of mappings.

5. Data Interaction

In order to explore the information the user must interact with the data. Besides zooming and panning the prototype provides several filtering mechanisms.

5.1 Filtering mechanisms

Several filtering mechanisms are included, as already described in [6]: suppression of one or more information group; limitations in the attribute domains; use of degree of interest function to reduce the amount of information displayed and also to select the representation; display of the information by levels of interest.

The user can interactively suppress the representation of one or more information groups marking them as inactive. By default, all the information groups are active and will be displayed.

Domain restrictions can be applied to one or more attributes of each global information group. They can be combined with and and or operators. For each group only the elements that satisfy the conjunction of the restrictions imposed for each attribute are selected. The final result is the union of the sets of the elements selected in all active groups. This corresponds to the disjunction of the conditions defined in each group. In order to combine both conjunctions and disjunctions in the same group, the user can interactively create derived global information groups. These groups are copies of the base group subjected to different restrictions. They are treated like the other groups, so inside each group the and operator is applied

and the or operator combines the results of the different groups [5].

Using a degree of interest function allows the control of the amount of information displayed. The value of the function in a point x depends on the *a priori* importance of the point, $API(x)$, and on the distance, $D(x,y)$, between x and the current focus, y : $DOI(x|y)=API(x)-D(x,y)$. Only the points whose DOI is over a given threshold, k , are displayed. As already mentioned, for each information group an attribute is chosen to define the *a priori* importance. If none of the attributes is chosen as API attribute, the *a priori* importance is, by default, equal to one. If a focus is not defined the degree of interest function is equal to the *a priori* importance. Both the distance and the *a priori* values are normalized in $[0,1]$.

The degree of interest function is also used, along with the current scale factor, to select one of the representations available in the list of representations of each global information group. Less detailed representations are chosen for elements with less degree of interest as described in [4].

The order in which the elements are displayed can take into account its degree of interest. The range of the degree of interest values can be divided into a given number of intervals called levels of interest. The elements of each global information group are separated in different levels according to their degree of interest value. Interactively the user can decide to display the data group by group or level by level. In the last case, the elements in the same level of all active groups are displayed starting with levels with higher degree of interest. The user can interactively stop the display process controlling the amount of information displayed and omitting the less relevant one [6]. By default the data is displayed group by group.

5.2 Zooming

Besides the conventional zooming in method that applies a selected area to the whole drawing area or interactively enlarges the contents of the drawing area centred in the pointer position by a constant factor, another zooming in model is available for the representation by icons. This zooming model is the one proposed in [14] and adapted in [8]. In this model a selected area is enlarged without distortion but the global context is preserved. The scale factor is constant in the magnified area and decreases uniformly when going from the border of this area to the periphery of the drawing area. The user can interactively select the scale factor, which is restricted to a maximum value in order to reserve space to display the context. This maximum value depends on the dimension of the selected area and is shown to the user.

A zooming out operation is available and it is performed using a constant scale factor. An overview operation is also implemented allowing the display of all

the information at once.

Zoom operations combined with panning allow the exploration of the whole data space.

5.3 Interface

As can be seen in figures 1 to 4, the interface window is divided in a display area, where the data information is visualized, and four interaction areas: a menu bar with menus to select the data and mapping files and the visual mapping; a dialog area to interact with information groups; an area to control zoom and DOI parameters; and, finally, a button area to command several actions.

The button area includes buttons to redraw the display area (redraw button), redefine users' window in order to display all the grids entirely and redraw the data (fitview button), turn on the degree of interest function (DOI button), enable zoom (zoom button), update or define colour table (colortable_def button) and show colour table (colortable button). When DOI button is pressed the dialog to control DOI parameters is shown. Similarly, pressing zoom button shows zoom dialog.

In the area reserved to control DOI parameters the user can turn off the degree of interest function, select its threshold, turn on/off the focus, command the selection of a new focus and choose between displaying the data by group or by level of interest. In the zooming dialog area the user can select the type of zoom: zoom out, zoom in without context or zoom in with context. In the last case a scale factor must be specified.

It is important to point out the relevance of the information group concept to the automatic generation of the interaction dialog to control the amount of information. This control is more effective when the representation by icons mode is selected. In this case, by default, there is a zone in the information group dialog area for each global information group and the geometrical information groups are not shown. However, in the data file, where information groups are defined, the visibility of a group in this area can be changed, i. e., a geometrical information group can be included in the dialog area and a global information group can be discarded. By default, as already mentioned, all groups are active and only those shown in the interface can be marked as inactive.

The area reserved for a global information group includes: the name of the group; a toggle button to switch between active and inactive state; a button to create a derived group, or, in the case of a derived group, a button to delete the group; a button to create a dialog window to select the API attribute for the group; and a line for each attribute of the group containing its name and variation limits, for float values, or a list of values, for string values. The user can change the variation limits of a float type attribute or unselect/select occurrences of a string type attribute. By default all attributes of a group are shown,

however in the data file the user can select some or all the attributes to be hidden.

For geometrical information groups the dialog area is slightly different: the variation limits of the attributes cannot be changed and there is no button to create derived groups. By default, in this type of group none of the attributes is shown, but the user can change this state in the data file.

When the coloured surface mapping mode is selected, the information group dialog is different as the parameters that can be selected are not the same. In this case only the global information groups with known connections and not marked as invisible in the data file are shown in the interface. The dialog area for each global information group contains: the name of the group; a toggle button to switch the group between active and inactive; two scroll lists with the float attributes of the group in order to associate the attributes with colour and relief; and, in the case of 3D grids, a button to generate the dialog to define a cutting plane.

6. Examples

The examples show three data sets. The first one contains data generated by a debugging tool for message passing distributed applications [7], the second one includes locations and attributes of public buildings in a city, and the last one temperatures measured on an isobaric surface between the parallels 80° and 20° and the meridians -45° and 45° at a given time.

In the first case we have abstract data without spatial locations. To represent the information we need to associate the data to graphical elements at some position in space. The data includes processes, communication events, messages and information related to them. Each process has an identifier and contains a sequence of communication events. Send, receive and consume events are the three kinds of communication events. A send event occurs when a process sends a message, a receive event when the process is ready to consume a message and a consume event when the process effectively consumes a message.

A receive event is classified as blocking if the execution of the process is suspended until the expected message arrives; otherwise, it is non-blocking. A receive from any corresponds to a receive call that does not specify any particular process identifier for the message sender. Therefore, there are four distinct types of receiving events: blocking receive, non-blocking receive, blocking receive from any and non-blocking receive from any.

A consume event is always associated with a receive that occurs immediately before it. However, a receive may not be followed by a consume because the process may be waiting for a specific message that is not delivered to it.

Program's execution is frequently represented with a time-space diagram [13]. This type of diagram contains two orthogonal axes, one for time and the other for processes. Time grows from left to right and an horizontal line represents every process. Event symbols are displayed above each process line. An oblique line connects every message send event to the corresponding consume event.

The diagram is not intended to reflect the exact time of event occurrence. However it is consistent with the *happened before* relation defined by Lamport [13]. That is, in each process line, events are represented from left to right according to their internal order of occurrence and a consume event is always represented to the right of the corresponding send event.

To obtain the time-space diagram a global information group for each type of event has been used. In each group the grid contains the positions of the corresponding event. The representation defined for the events in each group is very similar. A letter is used to identify the type of the event: S for send, C for consume, b for blocking receive, n for non-blocking receive, B for blocking receive from any, N for non-blocking receive from any. This letter is included in the group name in order to identify the symbol associated with the event (Fig. 1).

Filtering using attribute domain restrictions is available for groups concerning non-blocking receive events. In this case the attribute that indicates whether a message has been received or not is shown. It has value one if a message has been received and consumed, and zero otherwise. The user can interactively exclude all receive events that are not immediately followed by a consume event omitting value zero, i. e., setting value one as maximum and minimum value.

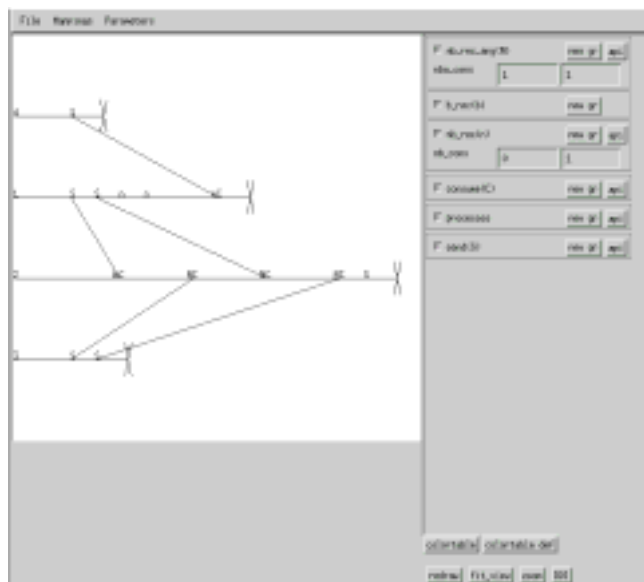


Fig.1 Abstract data without spatial location

Processes and messages connect points of the grids defined by each type of event. Geometrical groups handle processes and messages. The groups associated with processes also include the process identifier and a symbol (X) to mark the end of the process.

This example shows, on one hand, how the information group concept allows the organization of data in a flexible way and, on the other hand, how the two types of information groups can be combined together in order to obtain an adequate representation without the need of any special kind of mapping.

In the second example we have abstract data with a spatial reference, several types of buildings with associated attributes, together with information about the location of main streets. Each type of building is handled with a global information group and geometrical information groups are used for sets of streets. Several representations can be associated with each global information group and one of them is selected according to current scale and to the value of the degree of interest function, if this feature is active. To illustrate this, four different, although similar, representations were defined: a letter that identifies the type of building, which can be surrounded by one, two or three ellipses. In figure 2 all information is represented and in figure 3 the degree of interest function is active and a focus is defined.

Finally, another example of abstract data with spatial reference is shown but, in this case, the data corresponds to observations of a continuous variable. There is only one global information group with one attribute – the temperature. Figure 4 shows the data represented with a coloured surface with relief and colour associated to the temperature.

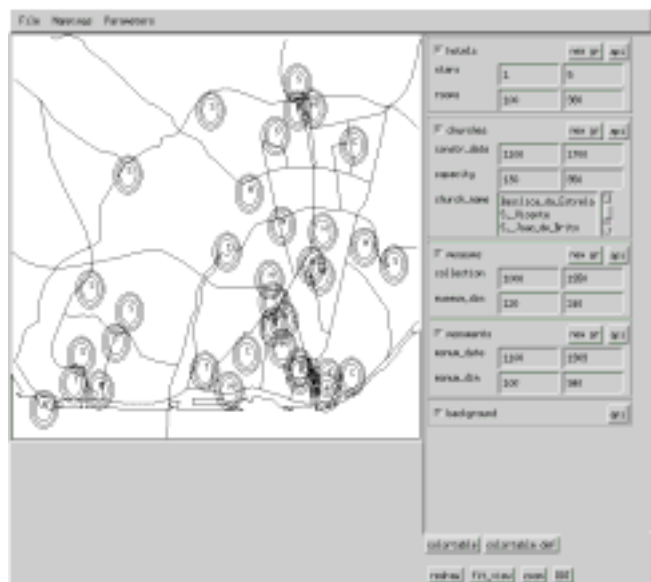


Fig.2 Abstract data with spatial location

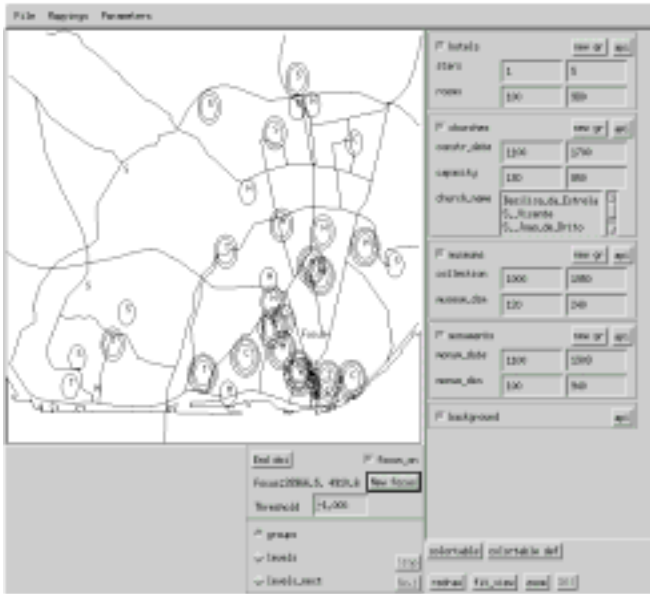


Fig.3 Icons selected according to degree of interest function

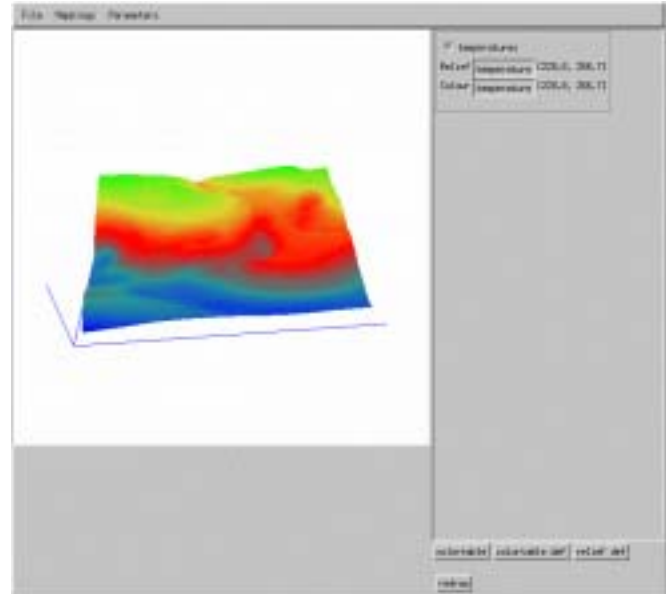


Fig.4 Coloured surface mapping

7. Conclusions and future work

This information visualization prototype automatically generates the visualization of abstract data that can be described in a table along with geometrical data. The interface is dynamically adapted according to the data. Several filtering mechanisms are included providing user control of the volume of information displayed. Moreover, based on degree of interest functions, the user ensures that the most relevant information is visible.

The concept of information group provides a powerful structuring mechanism to combine data, graphical representations and spatial references. Information groups aggregate semantically related information that is displayed together. Furthermore, it is possible to define different graphical representations to each global information group.

As future work we intend to implement an algorithm to generate the connections of 3D scattered points, develop new mappings and include new degree of interest functions, taking advantage of the object oriented paradigm used to implement the prototype.

8. References

[1] C. Ahlberg, E. Wistrand, "IVEE: An Information Visualization & Exploration Environment", Proceedings InfoVis'95, Atlanta, 1995, pp. 66-73.
 [2] M. Campo, R. Orosco, A. Teyseyre, "Automatic Abstraction Management in Information Visualization Systems", Proceedings IV'97, London, 1997, pp. 50-56.
 [3] S.K. Card, J.D. Mackinlay, B. Shneiderman, *Readings in Information Visualization Using Vision to Think*, Morgan

Kaufmann Publishers, Inc., 1999.
 [4] M.B. Carmo, J.D. Cunha, "Visualization of Large Volumes of Information Using Different Representations", Proceedings IV'97, London, 1997, pp. 101-105.
 [5] M.B. Carmo, J.D. Cunha, A.P. Cláudio, "Visualization of geometrical and non-geometrical data", Proceedings WSCG'99, Plzen, 1999, pp. 34-46.
 [6] M.B. Carmo, J.D. Cunha, A.P. Cláudio, "Filtering mechanisms for the visualization of non-geometrical and geometrical data", Short Papers and Demos, Eurographics '99, M.A. Alberti, G. Gallo, I. Jelinek, (eds), Milano, 1999, pp. 213-215.
 [7] A.P. Cláudio, J.D. Cunha, M.B. Carmo, "Monitoring and Debugging Message Passing Applications with MPVisualizer", Euromicro Workshop PDP 2000 Proceedings, Rhodes, 2000
 [8] J.C. Cunha, M.B. Carmo, "Modelo da Pirâmide Truncada", VII Encontro Português de Computação Gráfica, Monte da Caparica, 1995, pp. 35-46.
 [9] I.J. Curington, M.D. Coutant, "AVS - a flexible interactive distributed environment for Scientific Visualization applications", Workshop on ViSC, Delft, 1991
 [10] G. Furnas, "Generalized Fisheye Views", Proceedings CHI '86, 1986, pp. 16-23.
 [11] R. Kazman, J. Carrière, "An Adaptable Software Architecture for Rapidly Creating Information Visualizations", Proceedings of Graphics Interface '96, Toronto, 1996, pp. 17-27.
 [12] I. Kolingerová, B. Zalik, "Improvements to Randomized Incremental Delaunay Insertion", accepted for Computers and Graphics, 2002
 [13] L. Lamport, "Time, Clocks, and the Ordering of Events in a Distributed System", CACM, Vol. 21, n° 7, 1978, pp. 558-565.
 [14] G. Robertson, J.D. Mackinlay, "The Document Lens", Proceedings UIST'93, 1993, pp. 101-108.
 [15] W.J. Schroeder, L.S. Avila, W. Hoffman, "Visualizing with VTK: A Tutorial", IEEE CG&A, Sep./Oct. 2000, pp. 20-27.
 [16] D. Speray, S. Kennon, "Volume Probes: Interactive Data Exploration on Arbitrary Grids", Computer Graphics, vol. 24, n° 5, 1990, pp. 5-12.