Visualizing Time-Series on Spirals

Conference Paper · January 2001

DOI: 10.1109/INFVIS.2001.963273 · Source: DBLP

Wolfgang Müller
University of Education of Weingarten

Visualizing Time-Series on Spirals

Marc Weber
c-cop GmbH

Marc Alexa
Technische Universität Darmstadt

Wolfgang Müller
e4ib.com

Abstract

In this paper, we present a new approach for the visualization of time-series data based on spirals. Different to classical bar charts and line graphs, the spiral is suited to visualize large data sets and supports much better the identification of periodic structures in the data. Moreover, it supports both the visualization of nominal and quantitative data based on a similar visualization metaphor. The extension of the spiral visualization to 3D gives access to concepts for zooming and focusing and linking in the data set. The spiral comes with additional tools to further enhance the identification of cycles.

Keywords. *Information Visualization, Graph Drawing, Visualization of Time-Series Data, Data Mining*

1. Introduction

The analysis of time series data is one of the most widely appearing problems in science, engineering, and business.

Time-series data is analyzed in order to discover the underlying processes, to identify trends, and to predict future developments. Often, the analyzed data displays a periodic behavior, providing a model to better estimate such trends. Examples for time series data with periodic structures are natural phenomena such as temperatures and radiation of light in a month or year. Some theories assume that economic cycles also show periodic characteristics.

Visualization has been successfully used to analyze time-series data for a long time. Especially line graphs have proven to be very effective in this context.

New more sensitive sensors in science and engineering and the widespread use of computers in corporations have increased the amount of time series data collected by many magnitudes. Existing approaches to the visualization of such large data sets are insufficiently suited in supporting people in discovering underlying structures.

In this paper we present the Spiral Graph a new approach for the visualization of time-series data. The Spiral Graph can visualize large data sets and is ideally suited to support the human ability to detect structures. Such structures are clues to hidden, underlying cyclic processes behind the data

2. State of the Art

Time series data is characterized by data elements being a function of time. In general, this data takes the following form:

$$D = \{(t_1, y_1), (t_2, y_2), \ldots(t_n, y_n)\}$$

with

$$y_i = f\langle t_i\rangle$$

The data elements y_i can represent different data types. Usually we differentiate between nominal, ordinal, and quantitative data or tuples of these in the case of multivariate data. The purpose of a visualization is to detect and validate characteristic properties of the unknown function *f*.

The visualization of time-series data has a long history. Time series plots appear for the first time in the illustration of planetary orbits in a text from a monastery school [14]. In science, time-series charts have been rediscovered not earlier than in the 18th century by Lambert to display periodic variation in soil temperature in relation to depth under the surface [9]. Playfair was the first to an of the analyze the effectiveness of line graphs and bar charts [11]; he applied these graphs for the analysis of economic data. A detailed discussion of the history of time-series plots can be found in [14]. Today the visualization of time-series data differs only little from these early approaches.

The most important visualization techniques for time series data are sequence charts, point charts, bar charts, line graphs, and circle graphs:

Sequence charts represent time-dependent data on a one-axis chart in chronological order. Data elements are visualized by marks at the corresponding distances to the origin of the axis. Using marks for the visualization of the data elements, sequence charts are restricted to the visualization of nominal time-series data.

Point graphs extend sequence charts into the second dimension and use the remaining dimension to visualize quantitative data aspects by the distance from the main axis.

Bar charts use bars instead of points to represent the data enhancing the comparability of the data elements.

Line graphs extend point graphs by linking the data marks with lines to emphasize the temporal relation.

Different sequences can be combined in a single graph to allow for a comparison of these sequences, leading to multiple bar charts and multiple line graphs. Depending on the data, this combination is restricted to 2-8 sequences in one graph. In addition, cycle and cycle length of the data have to be known in advance to allow for a comparison.

Circle graphs map line graphs into the spherical domain. They are usually used to visualize quantitative data with (assumed) periodic background and with a known cycle length. Similar to multiple sequences ca be combined in one cycle graph. Hereby, multiple cycles of a data set can be compared.

Lately, 3D versions of line and bar charts have been used to visualize time-series data in relation to a second free variable. Animation is used to visualize temporal aspects. A good overview on graphical representation for time-series data can be found in [5].

Bertin [1] performed a broader analysis of visual attributes which can be exploited in the visualization of data and their effectiveness to communicate certain types of information. Cleveland [3] further improved these studies and gives measures for the efficiency of a number of graph types in various applications.

Standard graphs and charts can be enhanced by different interactive techniques, such as scrolling, zooming, brushing, as well as focusing&linking:

- Scrolling extends the display area and allows for the representation of larger data sets. However, a comparison of data elements is only possible in the currently visible subset.
- Zooming is another approach to the visualization of large data sets. Initially a low resolution view is presented and the user can decide to zoom into interesting regions. Again, comparisons are only possible across the visible subset and important detail might not be visible in the overview.
- Focusing&linking [2] extends the idea of zooming by providing not only zoomed versions of the detail data, applying also different, more effective visualization techniques for the selected frame.
- Brushing provides such additional information as pop-ups which are automatically displayed as a roll-over effect.
- The information mural [7] is a visualization technique providing an initial view of the whole data set as the basis for further analysis with the above-mentioned interactive techniques.

An overview on interactive visualization techniques for time series data can be found in [13].

While all these techniques proved to be very effective in many cases, some general problems stay for the visualization of time-series data. The visualization of large data sets is still difficult and all these techniques do not efficiently support the identification of serial and periodic aspects in the data effectively. The detection of a periodic behavior in the data – though one of the main intentions for the visualization from the very beginning – is still difficult and often only possible, if the cycles are relatively obvious. Comparisons between different cycles - needed to identify - are also difficult. For instance, the detection of small, varying offsets in the period is rarely possible with today's techniques. Also, comparisons between different periodic processes are difficult.

Until now, spirals have rarely been used for visualization purposes.

A first example of a Spiral Graph was presented by Gabaglio in 1888 [4].

Bertin [1] presented a single example for the visualization of time-series data using a Spiral Graph. However, he does not discuss this visualization technique in much detail.

Mackinlay et. al. [10] used a spiral for calendar visualization; iconic representations of past daily calendar entries are positioned on the spiral to display the development of the calendar. However, though using the spiral for temporal data, this solution does not present a general approach to the visualization of time-series data.

Keim et. al. [7] introduced a pixel-based visualization technique for data mining in databases, where entries are represented as colored pixels or color icons. Keim et al. positions these entries on a roughly approximated spiral depending on their relevance to generate 2d iconic displays. Yet, they do not apply the spiral to the visualization of time-series data.

Hewagamage et. al. [6] used 2d and 3d spirals to visualize temporal and spatiotemporal events.

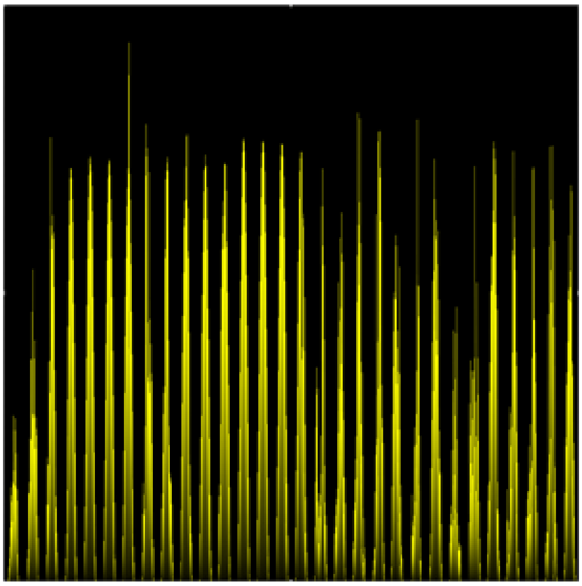
The Spiral Graph which we present in the following sections puts the focus on the put the focus on the goals of comparing data elements, both in a neighborhood and between cycles, and the identification of patterns and periodic behaviors in time-series data. These goals correspond to the aspects of comparative and summary reading introduced by Bertin [1]. We give a broader discussion of Spiral Graphs for time-series data and provide interactive 2d and 3d metaphors specifically for the purpose of analyzing with periodic behavior.

3. Using spirals for information visualization

To visualize and analyze time-series data and to allow for comparative and summary reading, the visual representation of the data set has to support the following aspects:

- provide a single visualization technique appropriate for nominal, ordinal, and quantitative data;
- support the visualization of large data sets;
- support comparative reading the data set, that is the comparison of values in a neighbourhood; also support the effective comparison of several cycles in a data set;
- support the analysis of the data set on the level of summary reading, that is the detection of periodic behaviours and trends in a data set as well as the determination of corresponding cycle lengths and intensity;
- allow for the comparison of multiple data sets, including the identification of offsets in their periodic behaviour.

Classical point graph, bar chart, and line graph have already proven to be very effective for the analysis of serial data. On the other hand, circle graphs have proven useful to expose and compare periodic behaviors in small sets of time-series data. Spiral graph visualizations represent a symbiosis of these techniques taking the advantages of both. The circu-



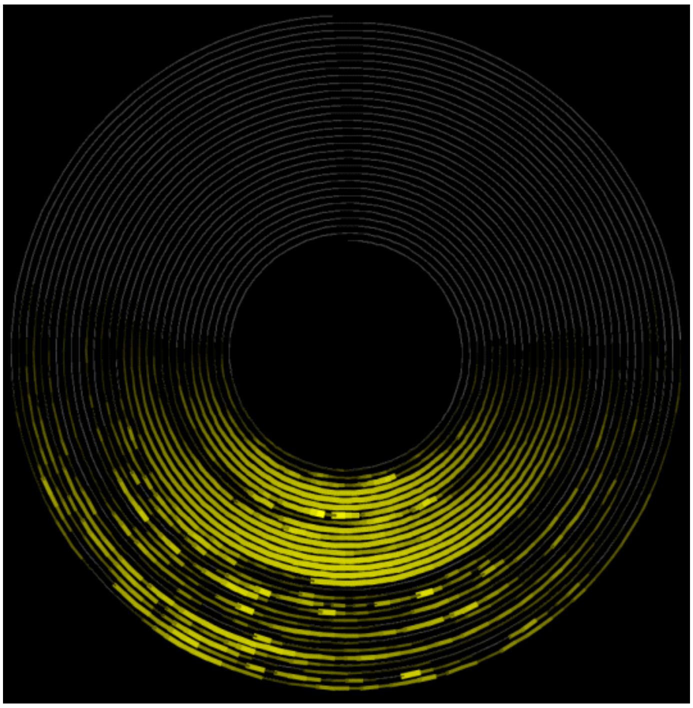


Figure 1:Two visualizations of sunshine intensity using about the same screen real estate and the same color coding scheme. In the spiral visualization it is much easier to compare days, to spot cloudy time periods, or to see events like sunrise and sunset.

lar structure of spirals allows for an easy detection of cycles and for the comparison of periodic data sets. Furthermore, the continuity of the data is expressed by using a spiral instead of a circle.

3.1. Mathematical description and types of spirals

A spiral is easy to describe and understand in polar coordinates, i.e. in the form $r = f(\varphi)$. The distinctive feature of a spiral is that *f* is a monotone function. In this work we assume a spiral is described by

$$r = f(\varphi), \qquad \frac{df}{d\varphi} > 0, \qquad \varphi \in R^{+}.$$

Several simple functions *f* lead to well-known types of spirals:

- Archimedes‘ spiral has the form $r = a\varphi$. It has the special property that a ray emanating from the origin crosses two consecutive arcs of the spiral in a constant distance $2\pi a$.
- The Hyperbolic spiral has the form $r = a/\varphi$. It is the inverse of Archimedes‘ spiral with respect to the origin.
- More generally, spirals of the form $r = a\varphi^{k}$ are called Archimedean spirals.
- The logarithmic spiral has the form $r = ae^{k\varphi}$. It has the special property that all arcs cut a ray emanating from the origin under the same angle.

For the visualization of time-dependent data Archimedes‘ spiral seems to be the most appropriate. In most applications data from different periods are equally important. This should be reflected visually in that the distance to other periods is always the same.

3.2. Mapping data to the spiral

In general, markers, bars, and line elements can be used to visualize time-series data similar to standard point, bar, and line graphs on Spiral Graphs. For instance, quantitative, discrete data can be presented as bars on the spiral or by marks with a corresponding distance to the spiral. However, since the *x* and *y* coordinate are needed to achieve the general form of the spiral their use is limited for the display of data values. One might consider to map data values to small absolute changes in the radius, i.e.

$$r = a\varphi + bv(\varphi), \qquad b < \frac{\pi a}{\max(v(\varphi))}$$

Yet, we have found this way of visualizing to be ineffective. We conclude that the general shape of the spiral should be untouched and other attributes should be used, such as

- colour,
- texture, including line styles and patterns,

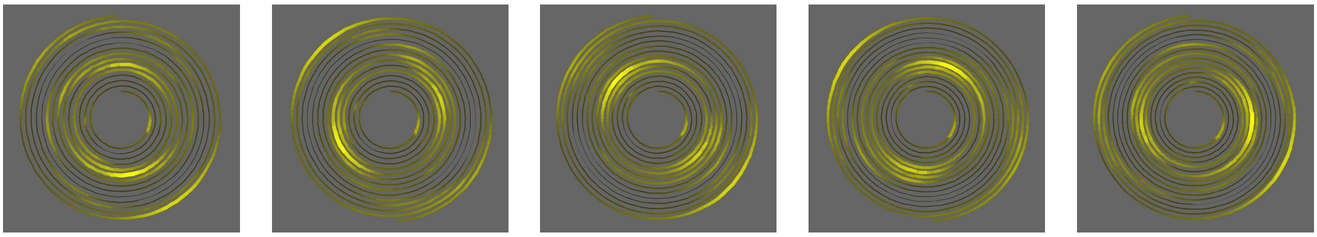


Figure 3: Visualizations of the same data with continously changing cycle length. The period in the data can be found visually, i.e. the visual system is used to detect periodic patterns in the data exploiting the spatial layout on the spiral.

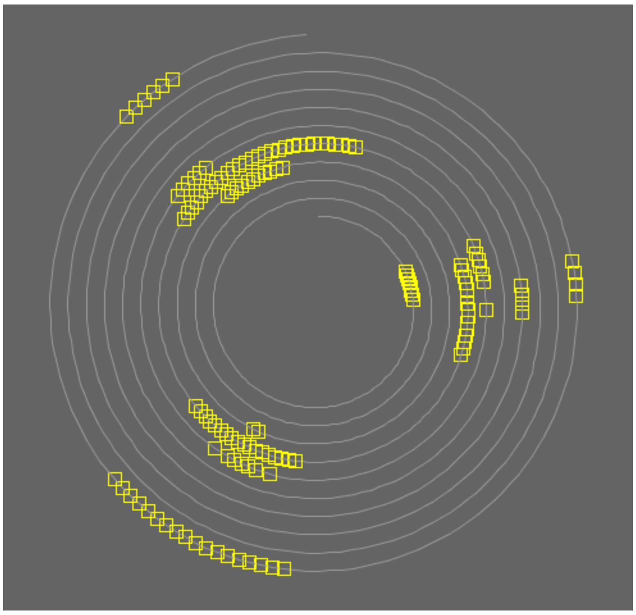


Figure 2: Nominal time-series data presented on a Spiral. The periodic behaviour of the underlying process is revealed.

- thickness of the line,
- or icons.

While line patters might be useful to distinguish different lines they are not appropriate to indicate data values. Therefore, the major visual scales considered here are color and thickness of the line.

First tests show that the most effective visualization of scalar data results from mapping data values to color and thickness (as in all the examples).

3.3. Spectra

If the periods in the data are known it is easy to scale the cycle length of the spiral so that data values with the same phase have the same phase in the spiral. For example, if a daily period is expected the data values could be mapped to angles so that 24 hours are represented by one cycle. The visualization will make periodic behavior apparent.

In many applications one might suspect or expect periodic behavior, however, it is not clear what the period could be or the data values have no timing information. In these cases it might help to compute the spectrum of the data values and use the frequencies with the highest amplitudes as possible candidates. These candidates could be inspected visually to confirm or deny the hypothesis.

Specifically, if the data values are known to be spaced regularly a standard Fourier transform is computed leading to the spectrum of the data. If the data values have timing information and are space irregularly it is still possible to compute a reasonable spectrum by least squares fitting of the sine and cosine functions to the data values (see [12] for more details). Once the spectrum is computer the user can choose frequencies with large amplitudes as cycle length.

3.4. Animating through cycles

A better approach is to utitlize the ability of the visual system of a human observer to discover structures: the spiral is animated by continuously changing the cycle length. Periodic behavior becomes immediately apparent during the animation as the visual appearance is changing from unstructured to structured (and to unstructured again). An illustration is given in Figure 3. The user can stop the animation when the period is spotted. In first experiments this approach has proven to work excellently.

3.5. Multi Spirals

Often one would like to compare the periodic behavior in a data set with cyclic patterns in other data or processes. Hereby, relations between different processes can be detected and presumptions on connections can be confirmed or rejected.

The comparison of different time-series can be achieved by rendering intertwined Spiral Graphs, for each data set a seperate graph at a time. Each data set should be visualized using a different visual encoding, such as color, texture, or line style. Figure 4 shows an example of a Multi Spiral with different color encodings. In our experience, 4 to 8 spirals can be combined in a single graph using color codings.

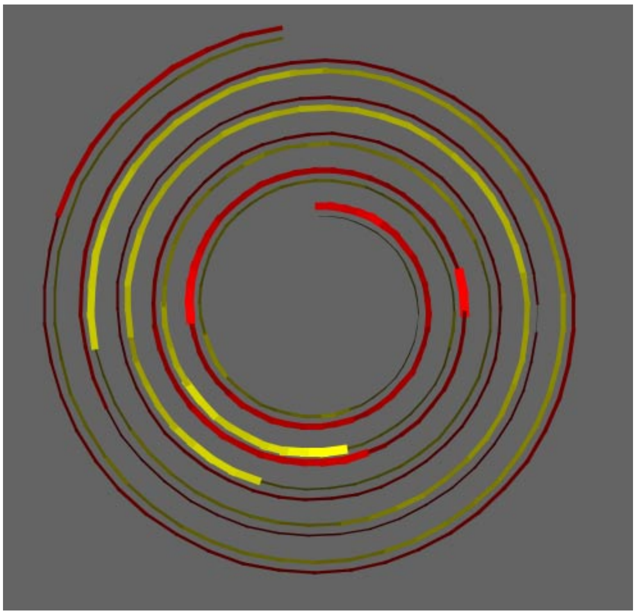


Figure 4:Stock prices of Microsoft (yellow) and Sun Microsystems (red) in five years on parallel spirals.

Individual interactive elements for the control of the spiral's cycle lengths have to be provided to allow for a period matching of the data sets.

4. Drawing the spiral

Most graphics programming languages or interfaces (e.g. OpenGL, Java2D) support line drawing in various styles and colors. Instead of trying to adapt the existing routines to draw a line in spiral form, the spiral is drawn as a compound of linear pieces. Each piece is an angular segment given by two angles φ_i and φ_{i+1}. These two angles yield the endpoints of the line segment with coordinates

$$x_i = a\varphi_i \cos\varphi_i \qquad y_i = a\varphi_i \sin\varphi_i$$

To control the visual error of this approximation, the angle of each segment is adjusted to the length in pixels of each segment. On the other hand, it is easier to use a constant number of segments per period. Therefore, the number of segments per cycle is chosen so that each segment is shorter than a predefined number of pixels. To compute the length of a segment a simple upper bound is the arc of a circle with the maximum radius of the segment. Thus, for cycle m with $\varphi \in [(m-1)\pi, m\pi]$ divided into d segments, an upper bound for the length of each segment is given by

$$l < \frac{am\pi^2}{d}$$

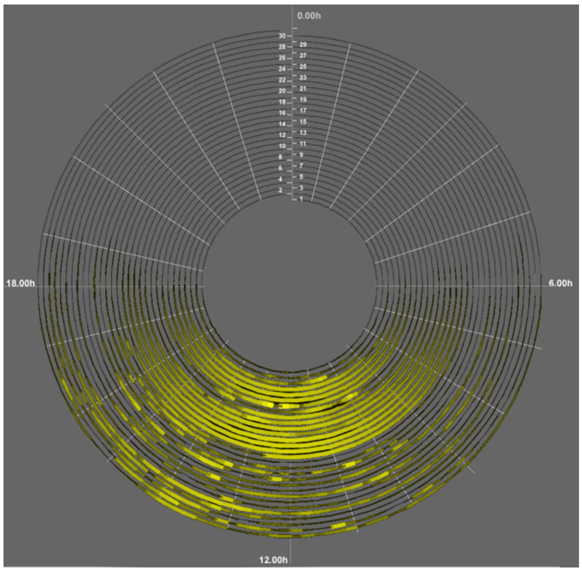


Figure 5:A possible way to add informative scales to the parametric dimensions of a spiral.

The number of pixels depends on the scale factor in the scan conversion process; however, it is linear in l meaning d is also linear in the unknowns.

It is possible to trade segment length for complexity of the segment, e.g. drawing circular arcs instead of line segments. However, it is more difficult to quantify the visual error and, thus, it is not apparent how long such segments should be chosen to satisfy a given error bound.

4.1. Scales on the spiral

If the meaning of the color coding scheme is not apparent, a legend should be provided. The meaning of locations on the spiral depends on the application. If, for instance, data is measured throughout a daily schedule, one cycle could be mapped to 24 hours and scales should inform which angles correspond to which times. Figure 5 gives an illustration.

5. Interaction

Interactive techniques are important for the visual exploration of data. Main point of most interaction metaphors is the navigation in large data spaces, i.e. exploring the data at large while easily accesing the necessary degree of detail.

One such approach is brushing, where an the data is exlored and detail is presented for the current focus. This could be easily integrated for the spiral layout since the computation of the data value from the spatial position of a pointing device is possible.

However, for large data spaces the spiral might not fit the screen. As an overview we have found a 3D helix to be use-

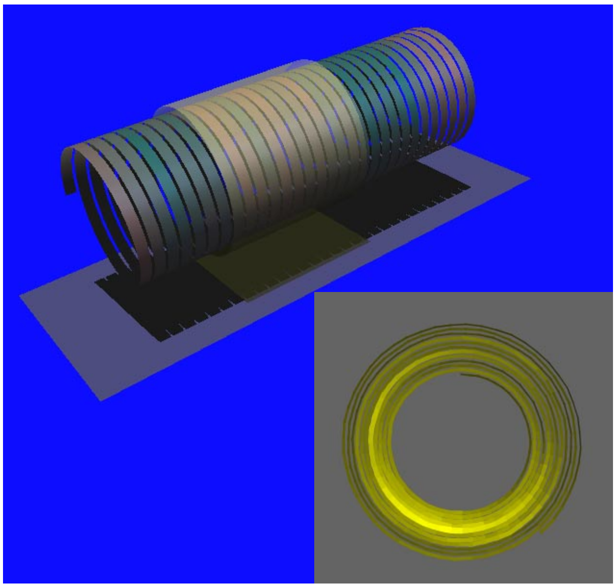


Figure 6:Using a helix in 3D to support intuitive browsing through a large data set.

ful. The complete data set is mapped on the cycles of the helix. It is possible to color-code the data values to give an impression of the data distribution.

To select a subset of the data set, a user can drag an extendable, semi-transparent 'marker' over the helix. The selected region is then visualized in a 2D spiral.

Due to the similar metaphors of helix and Spiral Graph, users in our experience have no problems associating the different elements of the visualization.

One might expect that the helix might also be used for data visulization. Note, however, that the identification of patterns and periodic behaviour requires the user to see full cycles, which is difficult to achieve in the case of the helix. Even a perspective projection along the axis of the helix is not identical to Archimedes' spiral and introduces distortion. Consequently, the 3D helix is best used for navigation only. Figure 6 shows an illustration.

6. Conclusions

In this paper we presented techniques for the visualization of time-series data with periodic behavior using Spiral Graphs. Spiral Graphs support the analysis of time-series data with the goals of detecting and confirming periodic behaviors in the data. In addition, this technique allows for the comparison of data points and cycles.

Spiral Graphs can be easily extended to Multiple Spiral Graphs to support the comparison of different data sets and their periodic behaviors.

We provided interactive techniques to support the visualization based on the interactive manipulation of spiral parameters and the mapping of data onto color, texture, and line thickness. Analysis tools based on a fourier analysis help to detect interesting cycle lengths and to identify periodic patterns in the data. Animations based on the variation of the cycle lengths extend this idea further.

For the selection of data segments we presented an interactive helix, which makes use of a similar metaphor for data visualization as the Spiral Graph. This supports further the idea of focusing&linking.

7. References

[1] J. Bertin. *Semiology of Graphics.* The University of Wisconsin Press, 1983

[2] A. Buja, J. A. McDonald, J. Michalak, W. Stützle. *Interactive Data Visualization Using Focusing & Linking.* Proceedings Visualization '91, IEEE Computer Society Press, Los Alamitos, 1991, pp. 156 - 163

[3] William S. Cleveland. *The Elements of Graphing Data.* Wadsworth Advanced Book Program, Pacific Grove, 1985

[4] Antonio Gabglio. *Theoria generale della statistica*. Milan, 2nd Edition, 1888

[5] Robert L. Harris. *Information Graphics - A Comprehensive Illustrated Reference*. Management Graphics, Atlanta, Georgia, 1996

[6] K. Priyantha Hewagamage and Masahito Hirakawa. *Pattern Browser: Spiral-Based Interactive Visualization Using Timelines on a Geographical Map.* IEICE Trans. Inf. & Syst., Vol. E83-D, No. 8, pp. 1679-1686

[7] Dean Jerding and John T. Stasko. *The Information Mural: A Technique for Displaying and Navigating Large Information Spaces.* Proceedings IEEE Symposium on Information Visualization, Atlanta, GA, 1995, pp. 43-50

[8] D.A. Keim and H.P. Kriegel. *VisDB: Database Exploration Using Multidimensional Visualization*. IEEE CG&A, Sept. 1994, pp. 40-49

[9] J.H. Lambert. *Pyrometrie.* Berlin 1779

[10] J.D. Mackinlay, G.G. Robertson and R. DeLine. *Developing Calendar Visualizers for the Information Visualizer*. Proc. UIST '94, 1994

[11] William Playfair. *The Commercial and Political Atlas.* London, 1786. From:

[12] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes*. Cambridge University Press, Cambridge, 1991

[13] Heidrun Schumann and Wolfgang Müller. *Visualisierung - Grundlagen und allgemeine Methoden.* Springer Verlag, Heidelberg, 2000

[14] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut, 1983