# Excentric Labeling:
# Dynamic Neighborhood Labeling for Data Visualization

**Jean-Daniel Fekete**
Ecole des Mines de Nantes
4, rue Alfred Kastler, La Chantrerie
44307 Nantes, France
Jean-Daniel.Fekete@emn.fr

**Catherine Plaisant**
Human-Computer Interaction Laboratory
University of Maryland
College Park, MD 20782, USA
Plaisant@cs.umd.edu

## ABSTRACT
The widespread use of information visualization is hampered by the lack of effective labeling techniques. A taxonomy of labeling methods is proposed. We then describe "excentric labeling", a new dynamic technique to label a neighborhood of objects located around the cursor. This technique does not intrude into the existing interaction, it is not computationally intensive, and was easily applied to several visualization applications. A pilot study indicates a strong speed benefit for tasks that involve the rapid exploration of large numbers of objects.

## KEYWORDS
Visualization, Label, Dynamic labeling

## INTRODUCTION
A major limiting factor to the widespread use of information visualization is the difficulty of labeling information abundant displays. Information visualization uses the powerful human visual abilities to extract meaning from graphical information [Card et al, 1998, Cleveland, 1993]. Color, size, shape position or orientation are mapped to data attributes. This visualization helps users find trends, and spot exceptions or relationships between elements on the display. Experimental studies have been able to show significant task completion time reduction and recall rate improvements when using graphical displays instead of tabular text displays (e.g., [Lindwarm-Alonso et al., 1998.]) However textual information in the form of labels remains critical in identifying elements of the display. Unfortunately, information visualization systems often lack adequate labeling strategies. Often labels are entirely missing and users have to peck at graphical objects one at a time. Sometimes labels overlap each other to the point of obscuring the data and being barely usable; or they are spread out in such a way that the relation between objects and labels becomes ambiguous. The problem becomes acute when the data density increases and the labels are very long.

To address this problem we propose "excentric labeling" as a new dynamic technique to label a neighborhood of objects (Figure 1 to 3). Because it does not interfere with normal interaction and has a low computational overhead, it can easily be applied to a variety of visualization applications.

The labeling problem is not new. It has been extensively studied for cartographic purposes [Christensen et al., 1998] where printing or report generation is the main purpose of the application. But very few solutions have been proposed to automate the labeling process of interactive applications. In this paper we propose a taxonomy of labeling methods, then describe our excentric labeling technique in detail, discuss its benefits and limitations, and illustrate how it can benefit a variety of applications.

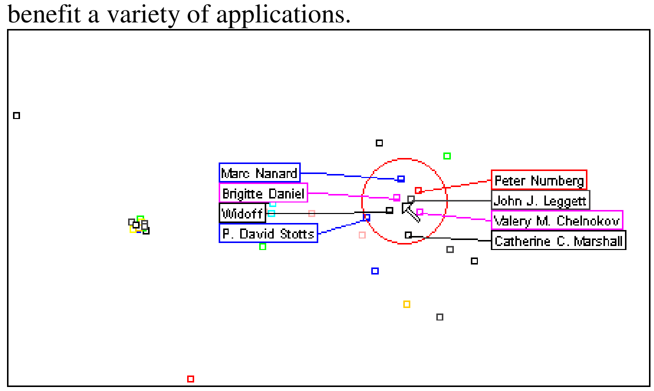


Figure 1: Excentric labeling provides labels for a neighborhood of objects. The focus of the labeling is centered on the cursor position. Labels are updated smoothly as the cursor moves over the display, allowing hundreds of labels to be reviewed in a few seconds. The color of the label border matches the object color.

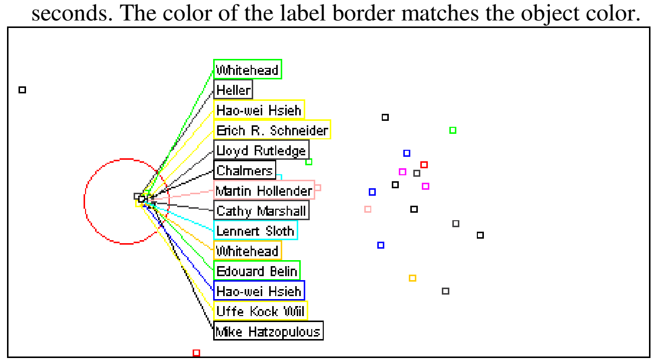


Figure 2: Labels are spread to avoid overlapping, possibly revealing objects clumped together on the display.

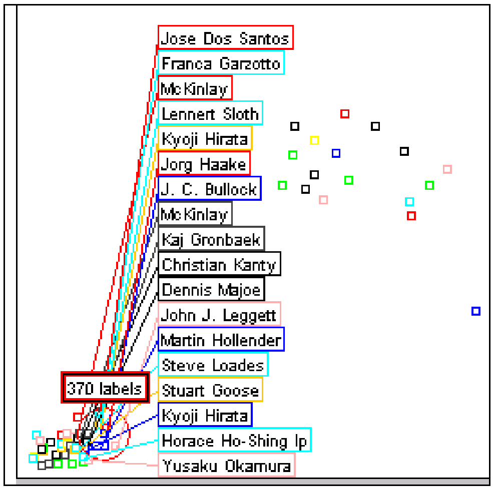

Figure 3: Special algorithms handle border effects (e.g., corners) When objects are too numerous, the total number of objects in the focus area is shown, along with a few sample labels.

## TAXONOMY OF LABELING TECHNIQUES

The labeling challenge can be stated as follows: given a set of graphical objects, find a layout to position all names so that each name (label) is:

1. Readable.
2. Non-ambiguously related to its graphical object.
3. Does not hide any other pertinent information.

Completeness (the labeling of all objects) is desired but not always possible.

Labeling techniques can be classified into two categories: static and dynamic. The goal of static labeling is to visually associate labels with a maximum (hopefully all) graphic objects in the best possible manner. But good static technique are usually associated with delays not suitable for interactive exploration. Dynamic labeling began with interactive computer graphics and visualization. Two attributes account for the “dynamic” adjective: the set of objects to be labeled can change dynamically, and the number and layout of displayed labels can also change in real time, according to user actions.

### Static Techniques

Static techniques have been used for a long time in cartography. Christensen et al. (to appear) wrote a recent summary of label placement algorithms. Cartography also needs to deal with path labeling and zone labeling, which is less widespread in visualization. We do not address those two issues in this article. But the same algorithms can be used for both cartography and general visualization. Since static techniques have to find "the" best labeling possible, the set of objects has to be carefully chosen to avoid a too high density in objects or labels. In cartography, this is achieved by aggregating some information and forgetting (sampling) others (this process is called "generalization"). This technique could be nicknamed the "label-at-all-cost" technique since one of the constraints is to label all objects of the display.

For data visualization, a similar process of aggregation can be applied to achieve a reasonable result with static techniques (e.g., aggregation is used in the semantic zooming of Pad++ [Bederson, 1994] or LifeLines [Plaisant et al., 1998]), but the logic of aggregation and sampling is mainly application dependent. Label sampling has been used occasionally (e.g., Chalmers et al., 1996).

The most common techniques remain the "No Label" technique, and the "Rapid Label-all" technique which leads to multiple overlaps and data occlusion [e.g., in the hyperbolic browser [Lamping et al, 1995]). Also common is the "Label-What-You-Can" technique in which only labels that fit are displayed; other labels that would overlap or occlude data objects are not shown (e.g., in LifeLines).

Some visualizations avoid the problem completely by making the labels the primary objects. For example, WebTOC [Nation et Al, 1997] uses a textual table of contents and places color and size coded bars next to each label.

### Dynamic techniques

Dynamic labeling techniques are more varied (see Table 1). The classic infotip or "cursor sensitive balloon label" consists at showing the label of an objet right next to the object when the cursor passes over it. The label can also be shown on a fixed side window, which is appropriate when labels are very long and structured.

In the "All or Nothing" technique, labels appear when the number of objects on the screen falls below a fixed limit (e.g., 25 for the dynamic query and starfield display of the film finder [Ahlberg et al., 94]). This is acceptable when the data can be easily and meaningfully filtered to such a small subset, which is not always the case. Another strategy is to require zooming until enough space is available to reveal the labels, which requires extensive navigation to see all labels. This technique can be combined elegantly with the static aggregation technique to progressively reveal more and more details - and refined labels - as the zoom ratio increases.

The overview and detail view combination is an alternative zooming solution [Plaisant et al., 1994]. The detail view can also be deformed to spread objects until all labels fit (i.e., in the way of a labeling magic lens). Those last two techniques require either a tool selection or dedicated screen space.

Chalmers et al., proposed dynamic sampling where only one to three labels are displayed, depending on the user's activity. Cleveland describes temporal brushing: labels appear as the cursor passes over the objects (similarly to the infotip), but those labels remain on the screen while new labels are displayed, possibly overlapping older ones.

| Type | Technique | Comments/Problems |
|---|---|---|
| STATIC | No label | No labels! |
| | Label-only-when-you-can (i.e. after filtering objects) | Need effective filters. Labels are rarely visible. |
| | Rapid Label-All | High risk of overlaps or ambiguous linking to objects |
| | Optimized Label-All | Often slow - may not be possible |
| | Optimized Label-All with aggregation and sampling | Effective but application dependant- may not be possible |
| DYNAMIC | | |
| **One at a time** | Cursor sensitive balloon label | Requires series of precise selection to explore space (slow), cannot reach overlapped objects. |
| | Cursor Sensitive label in side-window | Same as above. Constant eye movement can be a problem, but avoids occlusion of other objects. |
| | Temporal brushing (Cleveland) | More labels visible at a time, but overlapping problem. |
| **Global display change** | Zoom until labels appear | May require extensive navigation to see many labels (can be effectively combined with semantic zooming, e.g., Pad++) |
| | Filter until labels appear | May require several filtering to see labels (can be effectively combined with Zooming, e.g., starfields) |
| **Focus + context** | Overview and detail view without deformation | Effective when objects are separated enough in the detail view to allow labels to fit (not guaranteed.) |
| | Overview and detail with deformation/ transformation (i.e.fisheye or magic lenses) | Deformation might allow enough room for labels to fit. (not guaranteed). May require tool or mode to be selected. |
| | Global deformation of space (e.g., Hyperbolic Browser) | Requires intensive navigation and dexterity to rapidly deform the space and reveal all labels (e.g., by fanning the space). |
| **Sampling** | Dynamic sampling (Chalmers et al.) | Few labels are visible. |
| **NEW** | Excentric labeling | Fast, no tool or special skill needed. Spread overlapping labels, and align them for ease of reading. |

Table 1: Taxonomy of labeling techniques

## EXCENTRIC LABELING

Excentric labeling is a dynamic technique of neighborhood labeling for data visualization (Figure 1 to 3). When the cursor stays more than one second over an area where objects are available, all labels in the neighborhood of the cursor are shown without overlap, and aligned to facilitate rapid reading. A circle centered on the position of the cursor defines the neighborhood or focus region. A line connects each label to the corresponding object. The style of the lines matches the object attributes (e.g., color). The text of the label always appears in black on a white background for better readability. Once the excentric labels are displayed, users can move the cursor around the window and the excentric labels are updated dynamically. Excentric labeling stops either when an interaction is started (e.g., a mouse click) or the user moves the cursor quickly to leave the focus region. This labeling technique does not require the use of special interface tool. Labels are readable (non overlapping and aligned), they are non-ambiguously related to their graphical objects and they don't hide any information inside the user's focus region.

### Algorithm and Variations

To compute the layout of labels, we experimented with several variants of the following algorithm:

1. Extract each label and position for interesting graphic objects in the focus region.
2. Compute an initial position.
3. Compute an ordering.
4. Assign the labels to either a right or left set.
5. Stack the left and right labels according to their order.
6. Minimize the vertical distance of each set from the computed initial position.
7. Add lines to connect the labels to their related graphic object.

So far, we have used three main variations of this algorithm: non-crossing lines labeling, vertically coherent labeling and horizontally coherent labeling (the last two can be combined). Each uses a different method to compute the initial position, the ordering, to assign the labels to the stacks and to join the labels to their related graphic objects.

#### *Non-Crossing Lines Labeling – Radial Labeling*

The non-crossing lines labeling layout (Figure 4) does not maintain the vertical or horizontal ordering of labels, but avoids line crossings. This technique facilitates the task of tracing the label back to the corresponding object. It can be used in cartography-like applications where ordering is unimportant. The initial position on the circle (step 2 of previous section) is computed with a radial projecting onto the circumference of the focus circle. It is always possible to join the object to the circumference without crossing another radial spoke (but two radii - or spokes- may overlap). Then, we order spokes in counter-clockwise order starting at the top (step 3). The left set is filled with labels from the top to the bottom and the right set is filled with the rest.

Labels are left justified and regularly spaced vertically. We maintain a constant margin between the left and right label blocks and the focus circle to draw the connecting lines.
For the left part, three lines are used to connect objects to their label: from the object to the position on the circumference, then to the left margin, and to the right side of the label box. This third segment is kept as small as possible for compactness, therefore barely visible in Figure 4, except for the bottom-left label. For the right labels, only two lines are used from the object to the initial position to the left of the label. The margins contain the lines between the circumference and the labels.

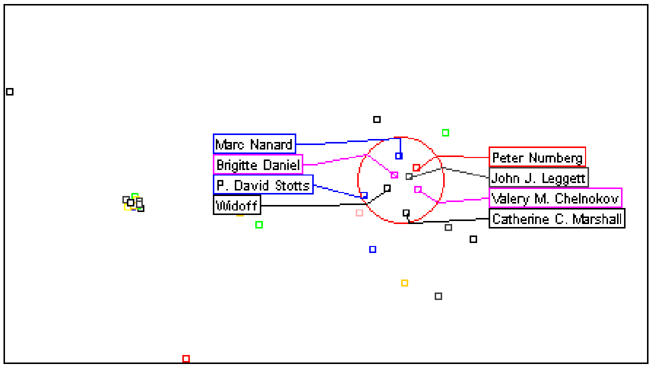


Figure 4: This figure shows the same data as in Figure 1 but using the non-crossing - or radial - algorithm.

### *Vertically Coherent Labeling*

When the vertical ordering of graphic objects has an important meaning we use a variant algorithm that does not avoid line crossing but maintains the relative vertical order of labels. This will be appropriate for most data visualization, for example, in the starfield application FilmFinder [Ahlberg, 1994], films can be sorted by attributes like popularity or length, therefore labels should probably be ordered by the same attribute. Instead of computing the initial position in step 2 by projecting the labels radially to the circumference, we start at the actual Y position of the object. The rest of the algorithm is exactly the same. Figure 1 and 2 shows examples using the vertically coherent algorithm, which is probably the best default algorithm. Crossing can occur but we found that moving slightly the cursor position animates the label connecting lines and helps find the correspondence between objets and their labels.

### *Horizontally Coherent Labeling*

When the horizontal ordering of graphic objects has a special meaning, we further modify the algorithm in step 5. Instead of left justifying the labels, we move them horizontally, so that they follow the same ordering as the graphic objects in Figure 5.

### *Dealing with window boundaries*

When the focus region is near the window boundaries, chances are that the label positions computed by the previous algorithms will fall outside of the window and the labels appear truncated (e.g., the first characters of the left stack labels would not be visible when the cursor is on the left side of the window).

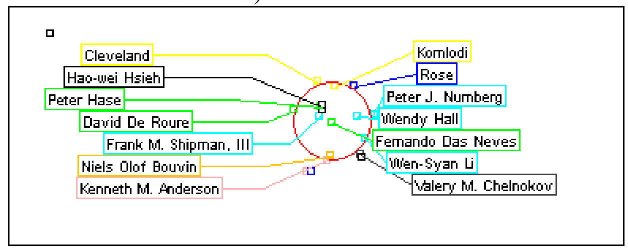


Figure 5: Here the labels order respect the Y ordering and the indentation of the labels reflects the X ordering of the objects.

To deal with window boundaries the following rules are applied. If some labels are cut on the left stack, then move them to the right stack (symmetric for the right side.) When labels become hidden on the upper part of the stack (i.e., near the upper boundary), move them down (symmetric for the bottom). Combining those rules takes care of the corners of the window (Figure 6).

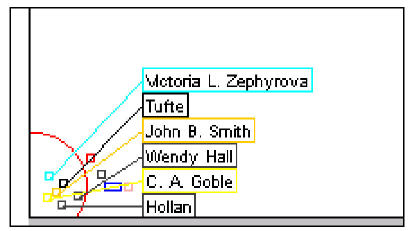


Figure 6: When the focus is close to the window boundaries, labels are moved so that they always fall inside the window.

## DISCUSSION

Excentric labeling fills a gap in information visualization techniques by allowing the exploration of hundreds of labels in dense visualization screens in a matter of seconds. Many labels can be shown at once (optimally about 20 at a time.) They are quite readable and can be ordered in a meaningful way. Links between objects and labels remain apparent. The technique is simple and computationally inexpensive enough to allow for smooth exploration while labels are continuously updated. Of course these algorithms don't solve all the problems that may occur when labeling. Three important challenges remain, and we propose partial solutions for them:

### *Dealing with too many labels*

We estimate that about 20 excentric labels can reasonably be displayed at a time. When more objects fall in the focus region, the screen becomes filled by labels and there is often no way to avoid that some labels fall outside the window. We implemented two "fallback" strategies: (1) showing the number of items in the focus region, and (2) showing a sample of those labels in addition to the number of objects (see Figure 3). The sample could be chosen randomly or by using the closest objects to the focus point. Although not entirely satisfactory, this method is a major improvement over the usual method of showing no labels at all, or a pile of overlapping labels.

The dynamic update of this object counts allows a rapid exploration of the data density on the screen. Of course (this is data visualization after all) the number of objets could also been shown graphically by changing the font or box size to reflect its level of magnitude.

### *Dealing with long labels*

Labels can be so long that they just don't fit on either side of the focus point. There is no generic way to deal with this problem but truncation is likely to be the most useful method. Depending on the application, labels may be truncated on the right, or on the left (e.g., when the labels are web addresses), or they may be truncated following special algorithms. Some applications may provide a long and a short label to use as a substitute when needed (e.g., Acronyms). Using smaller fonts for long labels might help in some cases. If long lines occur infrequently, breaking long labels in multiple lines is also possible.

### *Limiting discontinuities*

One of the drawbacks of the dynamic aspect of excentric labeling is that the placement of an object’s label will vary while the cursor is moving around the object. This is needed to allow new labels to be added when the focus area covers more objects, but can lead to discontinuities in the placement of labels. For example when the cursor moves from the left side of an object to its right side, the label will move from the right to the left stack. This effect is actually useful to confirm the exact position of a label but might be found confusing by first time users. We found that discontinuities were more common with the non-crossing algorithm than the Y coherent algorithm which we favor despite the risk of lines crossing.

## POSSIBLE IMPROVEMENTS

Depending on the application, several improvements might be considered :

- Changing the size and shape of the focus area can be allowed, either at the user's initiative, or dynamically as a function of the label density;
- When too many items are in the focus area, excentric labels can show not only the number of objects but also a glyph or bar chart summarizing the contents of the area (e.g., showing the color distribution of the points in the focus).
- Labels can inherit more graphic attributes from the objects they reference, as is often done in cartography. We currently map the color of the label border to the object’s color. But text font size, style or color can also be used if clear coding conventions exist and if adequate readability is preserved.
- Excentric labels can easily be used as selection menus. For example pressing a control key can temporarily "freeze" the excentric labeling, free the cursor, allowing users to select any of the labels.

## USE WITHIN EXISTING VISUALIZATION APPLICATIONS

We have implemented excentric labels within three different applications: a java version of starfield display/dynamic query visualization [Ahlberg et al, 1994] (Figure 7), a Java implementation of LifeLines (Figure 9), and a map applet to be used for searching people in a building. The addition of excentric labeling to the first two applications was done in a few hours. The last program was built from scratch as an evaluation tool.

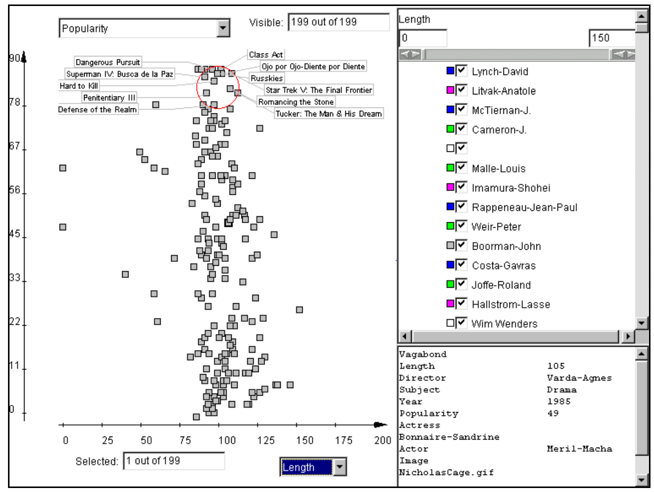


Figure 7: Excentric labeling was found effective in a java implementation of a starfield/dynamic query environment. It provides a rapid way to review the names of the data objects and to fathom the density of the overlapping areas.

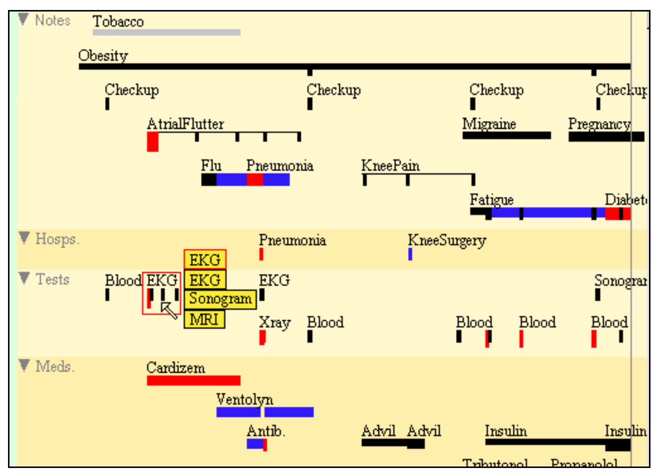


Figure 8: In LifeLines excentric labeling can be useful as it guarantees that all events in the focus are labeled, even if events overlap. Chronological order is best for ordering labels. In this example the focus area is rectangular (i.e., a time range) and no connecting lines are used. The label background is yellow to make them more visible.

## EARLY EVALUATION

We are in the process of comparing excentric labeling with a purely zoomable interface. The map of a building is displayed with workers names assigned randomly to offices. Subjects have to figure out if a given person is assigned to a room close to one of three red dots shown on the map (symbolizing three area of interest that a visualization would have revealed, e.g., areas close to both vending machines and printers). Each subject has to repeat the task ten times with new office assignments and red dot locations. The questions asked are of the form: "is <the person> in the neighborhood of one of the red dots?" Subjects reply by selecting "yes" or "no". The time to perform each task and the number of errors are recorded. Subjects using excentric labels (Figure 9) have to move the cursor over and around each highlighted point and read the labels. Subjects using the zooming interface have to move the cursor over each highlighted point, left click to zoom until they can read the labels (one or two zoom operation), right click to zoom back out or pan to the next point.

Our initial test of the experiment highlighted how speed and smoothness of zooming is crucial for zooming interfaces. In our test application a zoom or pan takes about 3/4 seconds to redraw. This is representative of many zooming interfaces, but in order to avoid any bias in favor of the excentric labeling we chose to ignore the redisplay time (the clock is stopped during redraws in the zooming interface version).

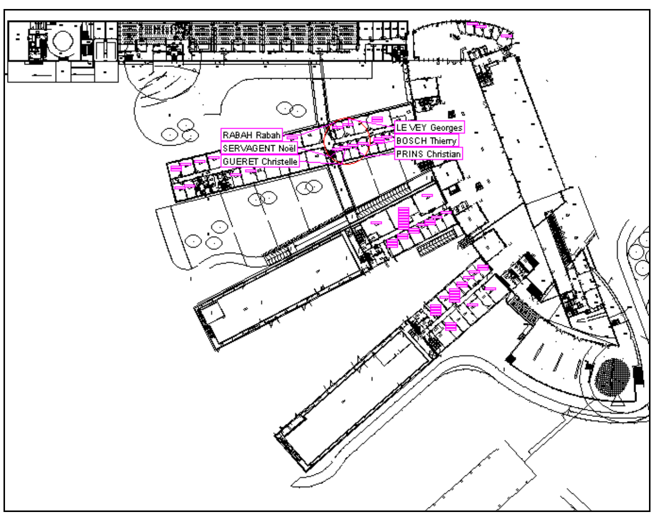


Figure 9: Map with a section of the building dynamically labeled. This application is being used to compare excentric labeling with a plain zooming interface when performing tasks that require the review of many labels.

Initial results of the pilot test (with 6 subjects repeating the task 3 times with each interface) shows that users performed the task 4 times faster when using the excentric labeling than with the zooming interface. Any delay in zooming and panning would further increase the effect in favor of excentric labeling. Our informal observations suggest that users sometime get lost using the zoom and pan, which does not happen when using the excentric labeling. On the other hand some subjects commented on the discontinuity problem.

## CONCLUSION

Despite the numerous techniques found in visualization systems to label the numerous graphical objects of the display, labeling remains a challenging problem for information visualization. We believe that excentric labeling provides a novel way for users to rapidly explore objects descriptions once patterns have been found in the display and effectively extract meaning from information visualization. Early evaluation results are promising, and we have demonstrated that the technique can easily be combined with a variety of information visualization applications.

## ACKNOWLEDGEMENT

This work was mainly conducted while Jean-Daniel Fekete visited Maryland during the summer 1998. We thank all members of the HCIL lab for their constructive feedback, especially Julia Li for her initial research of the labeling problem, and Ben Shneiderman for suggesting the main-axis projection. This work was supported in part by IBM through the Shared University Research (SUR) program and by NASA (NAG 52895).

## REFERENCES

1. Ahlberg, Christopher and Shneiderman, Ben, Visual information seeking: Tight coupling of dynamic query filters with starfield displays, *Proc. CHI'94 Conference: Human Factors in Computing Systems*, ACM, New York, NY (1994), 313-321 + color plates.
2. Bederson, Ben B. and Hollan, James D., PAD++: A zooming graphical user interface for exploring alternate interface physics, *Proc. User Interfaces Software and Technology '94* (1994), 17-27.
3. Chalmers M., Ingram R. & Pfranger C., Adding imageability features to information displays. UIST'96, Seattle Washington USA, ACM.
4. Christensen J., Marks J., Shieber S. Labeling Point Features on Map and Diagrams, to appear in *Transactions of Graphics*.
5. Cleveland, William, *Visualizing Data,* Hobart Press, Summit, NJ (1993).
6. Card, S, Mackinlay, J., and Shneiderman, Ben, *Readings in Information Visualization: Using Vision to Think*, Morgan Kauffman Publishers, to appear.
7. Cleveland, William, *Visualizing Data,* Hobart Press, Summit, NJ (1993).
8. Lamping, John, Rao, Ramana, and Pirolli, Peter, A focus + context technique based on hyperbolic geometry for visualizing large hierarchies, *Proc. of ACM CHI'95 Conference: Human Factors in Computing Systems*, ACM, New York, NY (1995), 401-408
9. Lindwarm D., Rose, A., Plaisant, C., and Norman, K., Viewing personal history records: A comparison of tabular format and graphical presentation using LifeLines, *Behaviour & Information Technology* (to appear, 1998).

10. Nation, D. A., Plaisant, C., Marchionini, G., Komlodi, A., Visualizing websites using a hierarchical table of contents browser: WebTOC, *Proc. 3rd Conference on Human Factors and the Web*, Denver, CO (June 1997).
11. Plaisant, C., Carr, D., and Shneiderman, B., Image-browser taxonomy and guidelines for designers, *IEEE Software 12*, 2 (March 1995), 21-32.
12. Plaisant, Catherine, Rose, Anne, Milash, Brett, Widoff, Seth, and Shneiderman, Ben, LifeLines: Visualizing personal histories, *Proc. of ACM CHI96 Conference: Human Factors in Computing Systems*, ACM, New York, NY (1996), 221-227, 518.
13. Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., and Shneiderman, B.(1998), LifeLines: Using visualization to enhance navigation and analysis of patient records, to appear *in Proc. of American Medical Informatics Association Conference,* Nov. 1998, AMIA, Bethesda, MD.

## DEMONSTRATION

Excentric labeling is implemented in Java. A demo program has been placed in a neutral site for demonstration.

**URL: http://www-ihm.lri.fr/excentric/**