Proceedings of CHI'99 (Pittsburgh, PA, May 1999). 184-185

Exploring a Database through Interactive Visualised Similarity Scaling

Pieter Jan Stappers and Gert Pasman
Department of Industrial Design
Delft University of Technology
Jaffalaan 9, NL-2628 BX Delft, The Netherlands
p.j.stappers@io.tudelft.nl, g.pasman@io.tudelft.nl

ABSTRACT
Databases often force users to verbalize queries in terms of attribute values. In matters of taste or subjective judgement, as in picking a color, or a meal from a menu, this is difficult and distracting. In our solution, MDS-Interactive, the dialogue between user and database is mediated through a small, evolving set of sample objects, which are presented visually by a real-time multidimensional scaling (MDS) algorithm. Queries are posed by indicating positions between the samples in order to find an object having like similarities to the displayed objects. Three prototype systems have been built and evaluated.

Keywords
Query dialogue, interaction design, visual interfaces, Multidimensional Scaling, MDS.

INTRODUCTION: SEARCHING ON TASTE
People do not always have a rigorously defined question in mind when they exploratively search in a collection. Examples are a designer browsing through magazines for inspiration on a design project, or consumers picking the colour of wallpaper or furniture to match their taste.

The way databases present themselves to these users is often less than helpful. First, users are forced to verbalize their questions in terms of jargon they are not familiar with, in order to specify values for an attribute-based search. Secondly, the results of this search are often presented in a sequential way, forcing the user to focus on each finding individually, thereby losing the so often desired overview of the results as a set. Thirdly, the findings are presented with such a definiteness that the user finds him- or herself to blame if the findings are not to taste. This discourages further exploration.

When given a book of samples, e.g., sample colours, the user is more able to indicate that "it's like this yellow, but more reddish", or "it's somewhere between this one and that one". But sample books often overwhelm the user by presenting several hundreds of small samples together, making it hard to mantain the overview. With loose samples, the user prefers to keep just a few to guide the further search.

SOLUTION
A solution which better fits this way of working presents an accessible (read: small) number of samples in a clear way and lets users specify queries with respect to these samples. To develop such an interface the following two questions were adressed: 1) how to present the samples in a way that the user can keep an overview, and 2) how to let the user specify the query relative to a set of depicted samples? In solutions we try to use the information about the objects in the database without requiring the user to learn the jargon involved.

Visualisation
For the first question we turned to a statistical method for visualising similarity relationships between objects named Multi-Dimensional Scaling (MDS), which has been used for over a few decades in psychological and marketing research [1]. It creates a 2D, 3D, or even higher-dimensional, arrangement of objects based only on information about the amount of similarity between object pairs. Similar objects appear closely together, dissimilar ones far apart.

MDS has proven remarkably robust and insightful. The arrangements of objects often reveal aspects of the structure of the collection that weren't noticed before.

In database visualisation MDS has been used by MacCuish et.al. [2]. They displayed all objects in the collection as a cloud of points on a display, allowing users to gather information on an object by clicking these points. The MDS arrangement was static, calculated previously because of the large size of the database.

Making it Interactive
The second question deals with searching the database.The robustness of the MDS method is such that if we remove a single sample, the arrangement of the whole does not change much. If we turn this reasoning around, this means that we are able to make the removed sample reappear based on the information of its position relative to the set. In this interpretation, any position in the MDS layout specifies a query for the database to find a sample which relates to the ones shown on the screen as indicated by its position.

Unwanted samples are removed by dragging them out of view. Together, the actions of adding and removing form the dialogue for exploring the collection. The dialogue can be initiated with a few samples chosen at random, or with some representative anchors (prototypical samples).

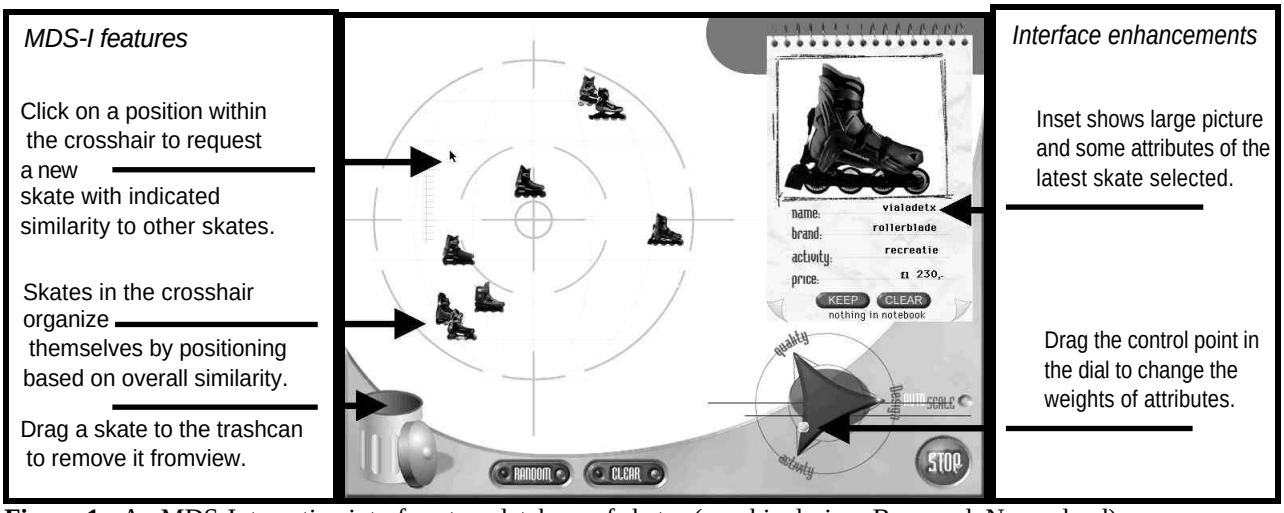

Figure 1. An MDS-Interactive interface to a database of skates (graphic design: Bram v.d. Nouweland).

DEVELOPMENT AND TESTING

These ideas have been implemented with Macromedia Director™ in a number of simple, interactive prototypes. An example of such a prototype is depicted in Figure 1. It shows a database of (roller) skates, rated on attribute dimensions of quality (e.g., price, type of bearings), design (e.g., colour, laces), and application area (e.g., leasure, speed). Other prototypes used a database of colours, organized by similarity on the basis of attributes of red, green, blue, hue, value, and saturation, and a collection of whiskys, rated on price and taste.

All prototypes have a single MDS-I overview (in Figure 1 depicted as a crosshair), and a means for advanced user to adapt the similarity criterion. Such advanced users would need to learn the meaning of the attribute dimensions, but not necessarily the attribute values.

Software feasibility

The prototypes were tested in a rough way in order to establish the fruitfulness of the concept. It turned out that an efficient iterative algorithm for the scaling is fast enough to make the samples smoothly arrange themselves in real-time. Also, for small numbers of objects, convergence is fast so that the layout only moves gradually and settles down quickly to a new position. Other database visualisations, such as ThinkMap [3], do not converge and keep moving. Although this continual motion reduces the problem of suggested 'definitiveness' mentioned in the introduction, our experience suggests it distracts from the visual comparisons, and smooth convergence is more suited to the type of queries we address.

User testing

In first tests about a dozen people tried searching tasks with the Whiskey samples or the skates sampler. Their dialogue with the system was videotaped, and they filled in questionnaires. In this phase we looked for signs of confusion in the interface design.

Preliminary results indicate that the dialogue by addition and removal works very well, but that the specific design needed to be improved. For example, some people thought the center of the display represented the focus of their query; some expected only high-quality skates to appear when they shifted the attribute weight toward 'quality', whereas this means that similarity scaling separates high-quality skates from low-quality skates. For non-experts it is better to choose established perceptual dimensions as attributes. For instance, colour comparisons weighting only the 'red' colour channel are confusing, as they group bright red and yellow together, while they are perceptually different.

CONCLUSION

The results suggest that the information in a database can serve users who do not know the jargon, as long as the samples shown to them have a meaning.

An advantage of the MDS-Interactive dialogue is its simplicity: users need little instruction, and can learn its operation by peeking over the shoulder of another user. As with traditional samples mentioned in the introduction, MDS-I can be used by more than one user, and needs only a touchscreen for interaction, enabling a low-threshold interface for consumer settings.

We are currently preparing for such consumer tests, and will apply the principle in a visual database supporting creative use of existing products by designers. Also, we continue to analyse the dialogue, with the aim of building a model of how people use the dimensional structure specified by MDS-scaled samples to navigate through large collections of objects.

ACKNOWLEDGMENTS

We thank our students Arvid Smit and Bram van de Nouweland for designing and user-testing the first prototypes of the dialogue style.

REFERENCES

1. Kimball Romney, A., Shepard, R., Nerlove, S., *Multidimensional Scaling; Theory and applications in the behavioural sciences*, Seminar, London, 1972.
2. MacCuish, J., McPherson, A., Barros, J., & Kelly, P. Interactive Layout Mechanisms for Image Database Retrieval. *Proc. SPIE Vol 2656*, 104-114, 1996.
3. Plumb Design, The Thinkmap White Paper. Available at http://www.thinkmap.com/.