OBJECT-CENTERED INTERACTIVE MULTI-DIMENSIONAL SCALING: ASK THE EXPERT

Joost Broekens, Tim Cocx, Walter A. Kosters.
Leiden Institute of Advanced Computer Science (*LIACS*), Leiden University, The Netherlands
Email: {broekens, tcocx, kosters}@liacs.nl

Abstract: Multi-dimensional scaling (MDS) is a widely used technique to show, in a low dimensional space, relations between objects—such as humans, documents, soil samples—that are defined by a large set of features. Key benefit is that it enables visual inspection of object relations in an intuitive way. One of the limitations is that different projections exist, leading to different graphical representations and therefore different interpretations of the data. This problem is made more significant in case of noisy data or heuristic approaches to MDS. We propose Object-Centered Interactive Multi-Dimensional Scaling (OCI-MDS), a technique that allows a data expert to try alternative positions for objects by moving them around the space in real time. The expert is assisted by visual feedback, such as the proportional error contribution of the expert-controlled object. Here we show the potential of this technique in two different domains: investigation of computer experiment configurations and biomedical data.

Keywords: MDS, Interactive Data Visualization, Clustering.

Hypothesis: Expert Interaction Helps Heuristic Projection

The use of computers has enabled people to create large amounts of data.

Multi-dimensional scaling (MDS) is a widely used technique to show, in a low dimensional space, relations between objects—such as human subjects, documents, soil samples—that are defined in a higher dimensional space.

Key motivation to use MDS for visualization of high dimensional data is its ability to give overview over a complete dataset. This is important in the exploratory phase of data analysis. For example, in the criminal investigation area.

Heuristic approaches to MDS are often used but these do have some problems:

- Different, quantitatively equivalent projections (especially in noisy data)
- Suboptimality (due to search-like behavior of heuristic)

Proposed solution: Object-Centered Interactive MDS

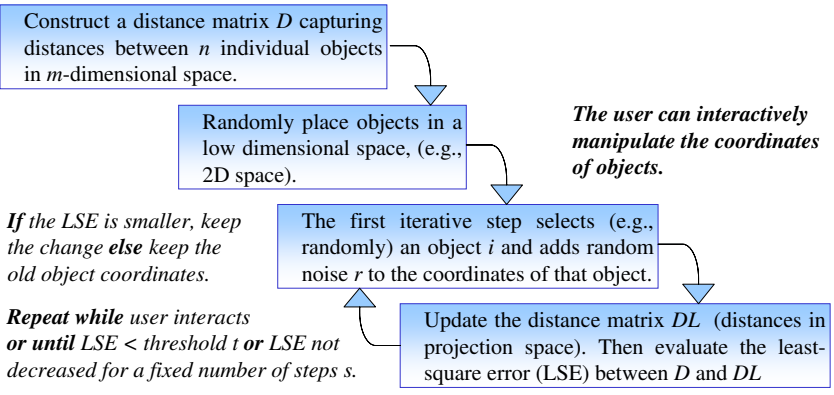
We propose Object-Centered Interactive Multi-Dimensional Scaling (OCI-MDS), a technique that allows a data expert to propose alternative positions for objects by moving them around the 2D space in real time.

Object-based interaction with MDS is useful, provided that users get feedback information so that they can

- select objects to move (*object selection criterion*), and
- evaluate the result of the move (*evaluation criterion*).

Object-Centered Interactive MDS: object manipulation by data experts in a heuristic MDS approach

We propose a four step algorithm, two preparatory steps, two iterative steps with user interaction during the last two:



Objects are drawn in colour so the user receives the following feedback:

- When objects are released, the projection mechanism restarts iteration so the user can directly observe the effect of the moved objects on the total distribution of objects.
- Objects are drawn in a color that represents the relative error contribution of the object.

Experimental results: clustering biomedical data

Figure 1: Using OCI-MDS in visualizing simulation configurations. Dataset: 44 experimental configurations for reinforcement learning experiments. ***Color scheme: relative error contribution of the object***.

Figure 2: Presents biomedical data visualization. Data consists of 400 objects (patients) with 10 features. Every feature represents the severity of a disease in a different body part. ***Color scheme: total disease severity calculated by averaging over the different features***.

Figure 3: Presents same biomedical data. ***Color scheme: relative error contribution of the object***.

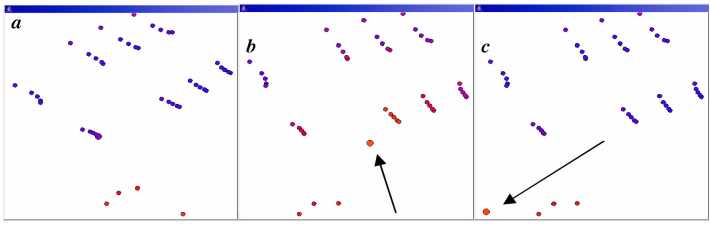


Figure 1. Manipulating experiment configuration clusters (local error color).

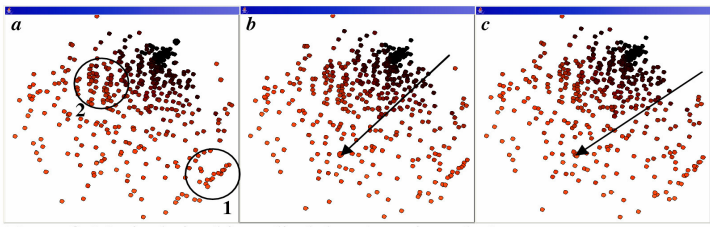


Figure 2. Manipulating biomedical data (severity color).

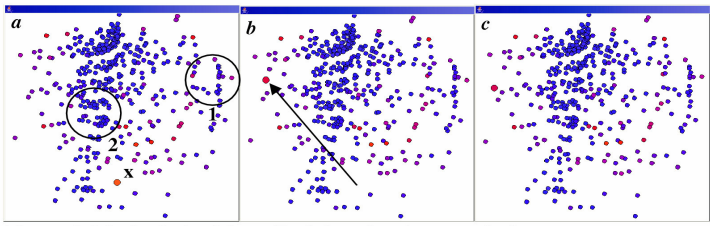


Figure 3. Manipulating biomedical data (local error color).

Conclusion: Object-Centered Interactive MDS has potential. It can be used for direct manipulation of clustering results based on a heuristic MDS approximation (**ASK FOR DEMO!**).

- It helps verifying the MDS result.
- It helps to generate hypotheses about alternative object relations, that were not found, for example, because the MDS converged to a local optimum.
- Currently its usefulness is somewhat limited on highly unstructured data.
- Future work: multiple object selection, more efficient MDS heuristic.

References and Related Work

J. Broekens & F. J. Verbeek. In: Proc. MNAS, pp.166-173, 2005.
T. K. Cocx & W. A. Kosters. In: Proc. ICDM, pp. 511-525, 2006.
M. L. Davison. Multidimensional Scaling, 1983.
W. A. Kosters & M.C. van Wezel. In: Proc. E-Commerce and Intelligent Methods, Studies in Fuzziness and Soft Computing, pp. 41–60, 2002.
P. J. Stappers, G. Pasman & P. J. F. Groenen. In: Proc. IEA2000, pp. 3-575–3-578, 2000.
E. Tejada & R. Minghim. In: Proc. IV'05, pp. 818–825, 2005.
W. S. Torgerson. Psychometrika 17(4), pp. 401–419, 1952.
M. Williams & T. Munzner. In: Proc. INFOVIS'04, pp. 57–64, 2004.