Data Visualization for Kansei Analysis

Takeshi Furuhashi
Dept. of Computational Science and Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
Email: furuhashi@cse.nagoya-u.ac.jp

ABSTRACT

Kansei data are multi-dimensional data. It is difficult for an analyzer to interpret data whose dimensionality is higher than three because his/her vision is used only to one -- three dimensions. Visualization by reducing the dimensionality of Kansei data to less than or equal to three dimensions could help the analyzer to understand the data.

For an effective visualization, definition of distances between data is important. For the definition, axes that form a space of Kansei data should be determined first. The choice could be questions or objects. A questionnaire uses several objects and many questions. Respondents are asked to answer each question one by one by marking on a rating scale. Questions are usually used as axes for multivariate analysis. Objects are another choice for the axes. By changing our viewpoint from different axes, new relationships between questions/objects could be found.

One of the most popular visualization methods is PCA (Principal Component Analysis). By applying PCA to questionnaire data on question axes, relationships between objects can be visualized. If the method is applied to the data on object axes, relationships between questions can be seen. The distance used in PCA is usually Euclidean distance. If some features can be extracted by a nonlinear transformation from original data, distances can be defined between the features. Kernel PCA[1] can be applied and kernel trick would work for dealing with multi-dimensional features.

MDS (Multi-Dimensional Scaling)[2][3] is also an effective visualization method. This method is to reduce the dimensionality of data while preserving the distances or dissimilarities between data in the original space. The weights on the differences between the distances in the visualized space and those in the original hyper-space make MDS a flexible tool for the visualization. These weights can be set to determine distances to be preserved or ignored. In exploratory visual data analysis, interactive manifold unfolding can be carried out by setting these weights.

ISOMAP[4] is an extension of MDS in which Euclidean distance is replaced with geodesic distance. Kansei data on manifold embedded in a hyper-dimensional space could be measured by their geodesic, or shortest path, distances and the manifold can be unfolded and visualized.

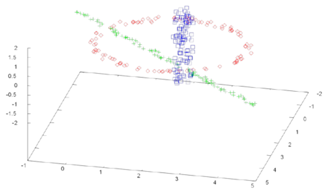

Fig.1 Clustering result by CGA

CGA (Conformal Geometric Algebra)[5] can deal with unique distances based on underlying geometry of data in multi-dimensional space. Distances between hyper-Lines, -circles, -spheres can be defined and if these entities are embedded in Kansei data, they could be clustered and thus be visualized. Fig. 1 shows a result where data were clustered in the forms of a line, a circle and a sphere.

Once data are visualized, exploratory Kansei data analysis, i.e. repetition of a process of choice of a distance, visualization, grouping, profiling, reading respondents' text data. An example of interactive grouping is shown in Fig. 2.

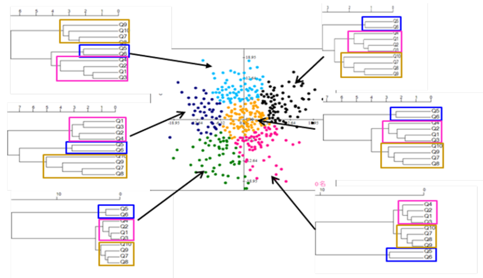

Fig.2 Interactive grouping of Kansei data

REFERENCES

[1] B. Schoelkopf, A. Smola, and K.-R. Mueller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, vol. 10, pp. 1299–1319, 1998
[2] J. A. Lee and M. Verleysen, Unsupervised Dimensionality Reduction: Overview and Recent Advances, Pr*oc. of WCCI 2010 IEEE World Congress on Computational Intelligence,* 2010.
[3] M. Ma, R. Gonet, R. Yu, G. C. Anagnostopoulos, Metric Representations of Data via the Kernel-based Sammon Mapping, http://phoenix.fit.edu/~georgio/pubs/Ma_ijcnn_2010.pdf
[4] J. B. Tenenbaum, V. de Silva, J. C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, SCIENCE, Vol. 290, No. 22, pp.2319-2323, 2000.
[5] C. Doran and A. Lasenby, Geometric algebra for physicists, *Cambridge University Press*, 2003.