The Generalized Sensitivity Scatterplot

Yu-Hsuan Chan, Carlos D. Correa, and Kwan-Liu Ma
VIDI research group, University of California at Davis
Outlines

- Review of scatterplots and Flow-based Scatterplots (VAST10)
- The pilot study on perception of the high dimensional trends on 2D
- The generalized sensitivity augmentation
Scatterplot

- Scatterplot is frequently used to reveal correlations between two variables.
  - (+) intuitive
  - (-) limited number of variables
  - (-) projection
- Scatterplot can potentially show global trends.

Courtesy of Wikipedia
Flow-Based Scatterplot (FBS)

Limitations
- Analysis after projection
- Only one pair of variables.

Questions
- Is it difficult for people to understand highD trends on 2D plots?
- Can we reveal the highD trends on 2D plots by the sensitivity augmentation?
Pilot Study

- How complex the trend one can perceive?

1. Prepare synthetic 3D trends from functions.
2. Show the 2D proj of the 3D trend on web page. Ask the participant has to identify the 2 functions and draw curves to describe them.
3. Measure the perception by the distances between the drawing and the function.

- 226 participants; each worked on 20 trends;
- 6534 effective drawings for 300 3D trends
Factors in interpreting the 3D trends

• # of patterns.
• Noise in sampling points for the 3D trend.
• Interpolation function.
Generalized Sensitivity

- Key: decouple sensitivity analysis from projection
  - Flow-based sensitivity: computed *after* the projection.
  - Generalized sensitivity: differentiated *before* the projection.
    - FBS relies on a 2D projection (1st limitation)

- The new sequence of approximating the sensitivity

![Diagram of input data, subspace selection, differentiation, and projection with Sensitivity Line]
Generalized Sensitivity Scatterplots (GSS)

- sensitivity lines reveal three different trends btw Y and X
- the shape of the high dimensional relationship
Generalized Sensitivity Glyphs

• Sensitivity Fans and Stars compare multiple pairs of sensitivity from different 3D subspaces formed by various Z variables.
Generalized Sensitivity Glyphs

• Large fans indicate where the high dimensional feature may be at.
• The size of stars suggests a classification
Selection in GSS

- Generalized sensitivity support non-linear transformation.
- Selecting a trend: data-aligned, feature-aligned

FBS: 2D trend

GSS: 3D trend.

3D view of GSS
Different sensitivity scatterplots

- Wine recognition dataset. Three classes and 178 wine samples.
- Scatterplots (Proline, Color)

**FBS:** a trend in (Proline, Color)

**GSS:** two trends in (Proline, Color, Alcohol)

**Stars:** three classes in (Proline, Color, Zi)
Conclusion and Future Work

• The pilot study on the perception of 3D trend on 2D projection.
• Generalized sensitivity augmentation
  • Sensitivity lines, fans and stars
  • Novel operations such as selection and clustering.
  • Comparison between FBS, GSS, and Fans in classification task
• Future Work
  • GSS for the classification of complex data
  • User study on the generalized sensitivity
Thank you.

- This research was supported in part by HP Labs, Northrop Grumman Corporation, and the U.S. National Science Foundation through grants CCF-1025269, CCF-0811422, and CCF-0808896.


- [netzen](https://code.google.com/p/netzen/) v1.0 [https://code.google.com/p/netzen/](https://code.google.com/p/netzen/)

- Contact me via [pany@cs.ucdavis.edu](mailto:pany@cs.ucdavis.edu)