An Image Quality Dataset with Triplet Comparisons for Multi-dimensional Scaling

Mohsen Jenadeleh*, Frederik L. Dennig*, Rene Cutura†, Quynh Quang Ngo†, Daniel A. Keim*, Michael Sedlmair†, Dietmar Saupe*
*Department of Computer and Information Science, University of Konstanz, Konstanz, Germany
†VISUS, University of Stuttgart, Stuttgart, Germany

VQEG 2024
University of Klagenfurt, Klagenfurt, Austria

July 2, 2024

Background

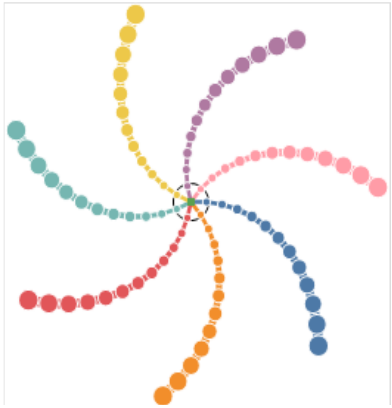
- State of the art for multi-dimensional quality assessment
  - Focus on assessment of predefined dimensions
  - A few works used difference scaling for visual and audio quality
  - Lack of annotated datasets for assessing image quality on multiple perceptual dimensions

Contributions

- Contributions
  - Image quality dataset with triplet comparisons
  - Analysis of the implicit dimensionality
  - Perceptually motivated triplet embedding method (LTE)

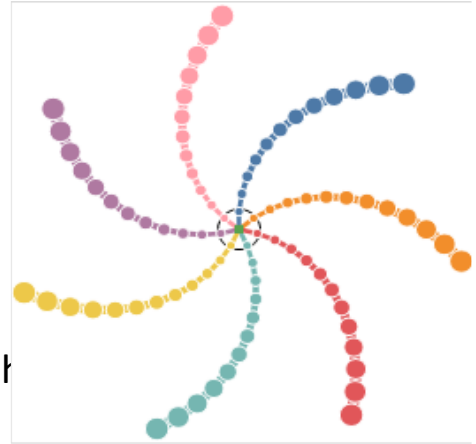
Simulations

- Simulated data with ground truth
  - Ground truth: 2D spiral pattern
  - Generating perceptual simulated responses for triplets
  - Multi-Dimensional Reconstruction of Perceptual Simulated Data:
    - We used 4 state-of-the-art methods
    - Our proposed method



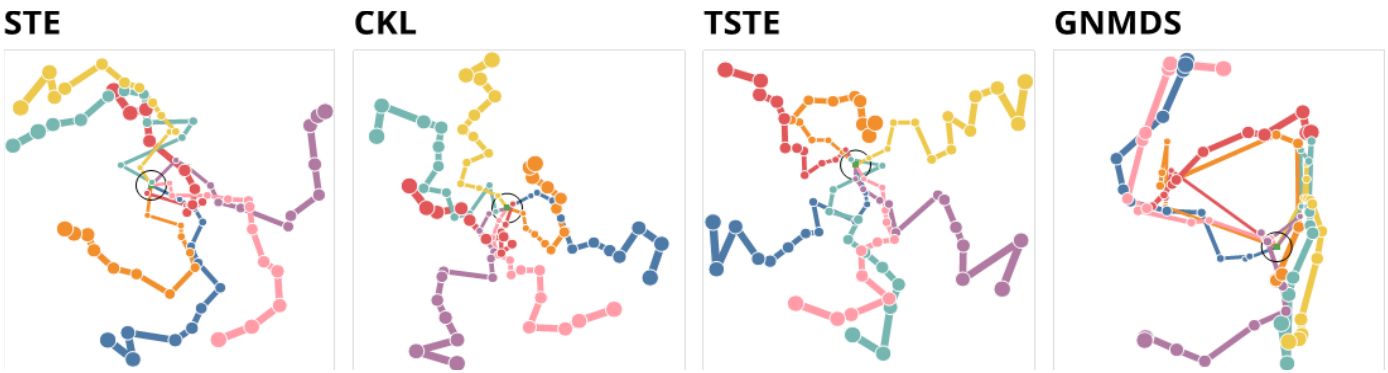
Simulations Results

- Results for simulated data with ground truth
  - Embedding using 90k triplets
  - All embedding methods reconstructed the ground truth



- Embedding using 6947 triplets

2D embedding:

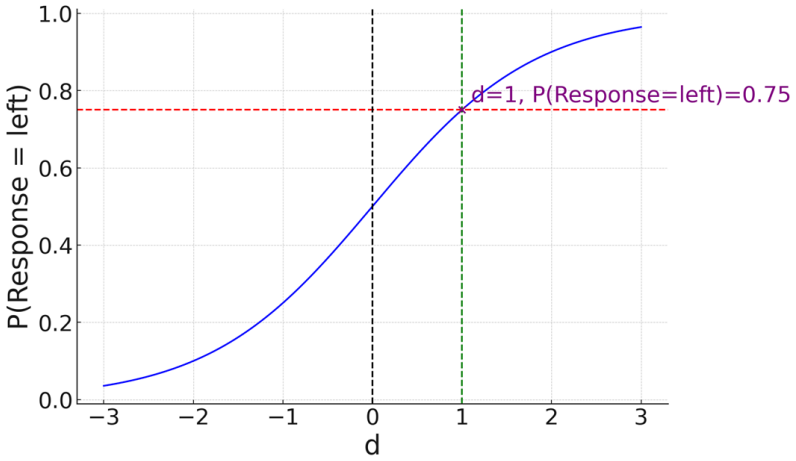


Proposed Logistic Triplet Embedding (LTE) with Tikhonov regularization

- Perceptually motivated (JND scale units)
- Equation

$$P(\text{Response} = \text{left}) = \frac{1}{1+e^{-\alpha d}}$$

$$d = \|x_r - x_p\| - \|x_l - x_p\|$$

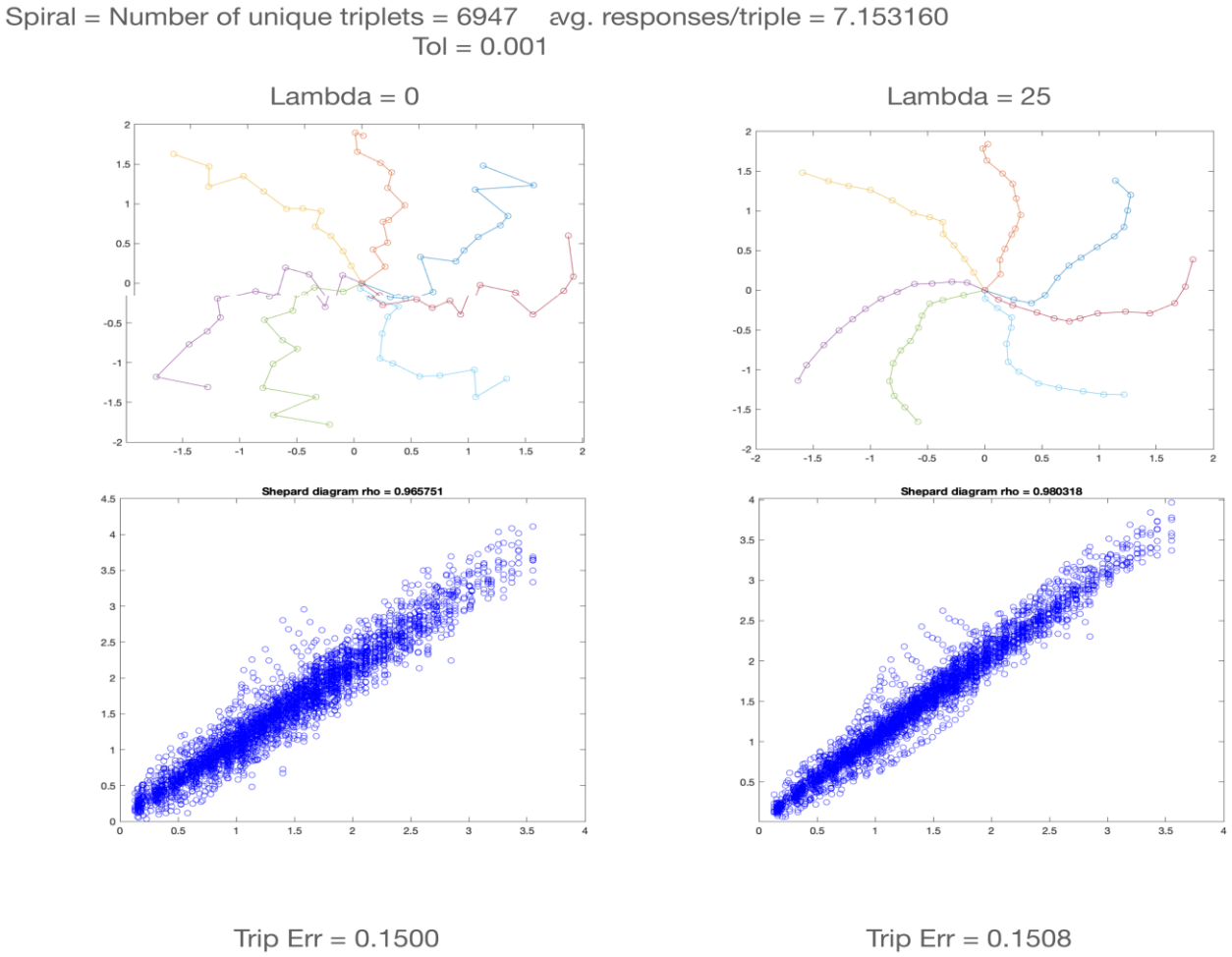


  - α=log3 scales the units in embe according to just-noticeable differences (JND).

- For a difference d = 1, we obtain P(Response = left) = 0.75, which corresponds to 1 JND.

LTE method

- Reconstruction of the simulated data in 2D



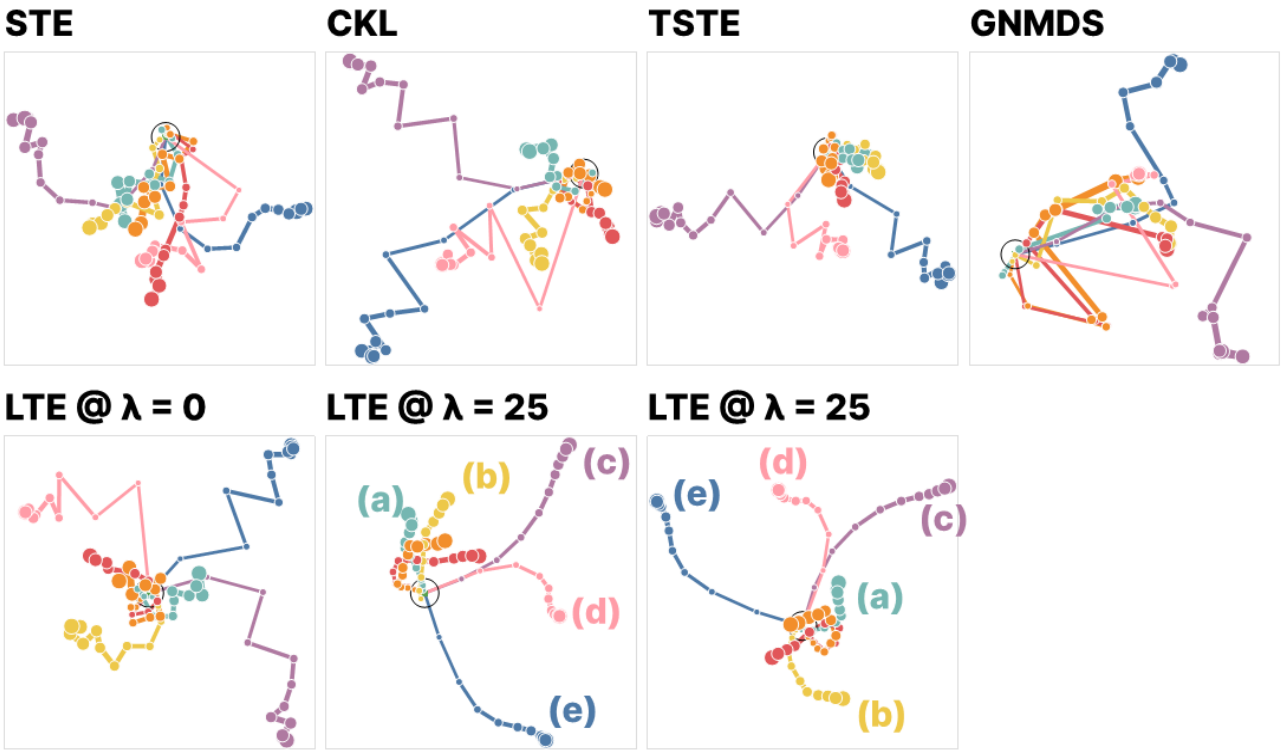
Subjective data

- Crowdsourced dataset for image quality assessment
  - Dataset with 85 images (1 source and 84 distorted versions)
  - 7 distortion types at 12 distortion levels
  - 49,693 responses to 6,947 triplet comparisons

Question: Which image (L/R) is more similar to the middle one?

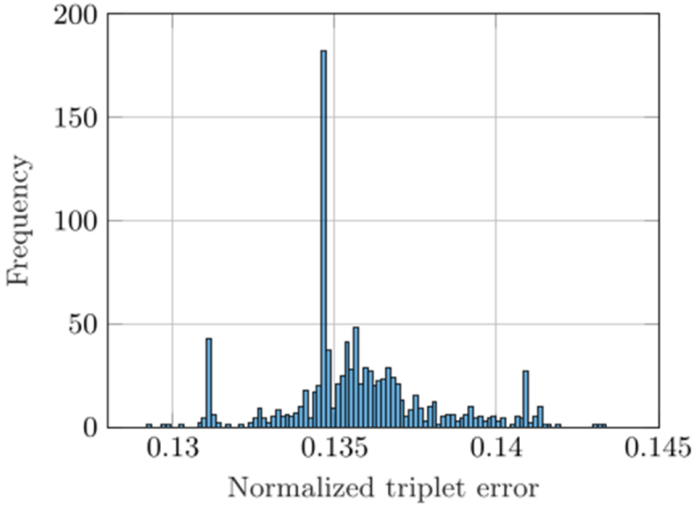
Multi-dimensional reconstruction of subjective data

- Embeddings for the dataset show how different distortion types branch out in 2D space



ASSESSMENT OF TRIPLET EMBEDDING METHODS

- Triplet embedding methods show distinct properties:
  - different global topologies may emerge for the same dataset
  - Sensitive dependence on initialization
  - Tikhonov regularization helps reveal the curvilinear strands



GNMDS

FUTURE WORK

- Extending the current study:
  - Extended analysis and visualizations.
- New large subjective study:
  - Source Images: 10; distortion types: 3; distortion levels: 5
- Which triplet subsets and which reconstruction methods are best for extracting hidden multi-dim. structures from perceptual response data?
- Design space:
  - Triplet subset selection methods
  - Triplet subset sizes
  - Number of dimensions
  - Selection of triplet reconstruction algorithms

Access the dataset:

- https://github.com/jenadeleh/multidimensional-IQA-dataset/tree/main

Funding:
DFG Project-ID 251654672 – TRR 161 projects A03 and A08;
DFG Project ID 496858717

Deutsche Forschungsgemeinschaft
German Research Foundation

SFB-TRR 161
Quantitative Methods for Visual Computing