Expanding the Universe – From Volume Rendering to High-Dimensional Data Visualization

Klaus Mueller

Visual Analytics and Imaging Lab
Computer Science Department
Stony Brook University and SUNY Korea
A Point
A Point
A Point
A Point in a Frame
Many Points in a Frame → An Image
A Point in a Cube
Many Points in a Cube → A Volumetric Image
Navigating the Volume
View Suggestion Framework
Features and Clustering

Feature – normal perturbation within a small neighborhood

\[ w(x_0, y_0, z_0) = \sum_{(x, y, z) \in N(x_0, y_0, z_0)} |\nabla f(x, y, z) - \nabla f(x_0, y_0, z_0)| \]

Example – cube with and without text
Interaction with Embedded Navigation Aids

Trackball with entropy map

presented at VIS 2011
Transfer Function Independence

May still need multiple transfer functions to see all features
Volume Generation
A Frame and Views
A Point Reconstructed from the Views
A Point in a Frame
A Point in a Cube
Many Points in a Cube
Computed Tomography
Backprojection

more:
Filtered Backprojection

First filter projection in the frequency domain

Then backproject
3D

Comprehension ✓

Navigation ✓

Reconstruction ✓
Consider the salient features of a car:

- miles per gallon (MPG)
- top speed
- acceleration
- number of cylinders
- horsepower
- weight
- year
- country origin
- brand
- number of seats
- number of doors
- reliability (average number of breakdowns)
- and so on...
What Does This Object Look Like?

Hard to imagine....
How Many 3D Cubes in N-D?

Combinatorial explosion

\[
\frac{N!}{3!(N-3)!}
\]
Scatterplot Matrices Don’t Scale

Can’t see multivariate relationships
- especially not when D is high
What You Need Is...

tripadvisor
Taking a Trip in Real Life

1. Identify the sights
   - use a map to identify the sights of interest and their location
Taking a Trip in Real Life

1. Identify the sights
   - use a map to identify the sights of interest and their location
Taking a Trip in Real Life

2. Plan the trip
   • connect the sights of interest along a path
Taking a Trip in Real Life

3. Go on the trip
   - travel along the route
Taking a Trip in Real Life

4. Hop off the bus
   • experience the location, look around, zoom into detail
5. Orient and localize
   • regain bearings in the map
Touring High-D Space...

**Exploration goals**
- find data configurations that best fit a personal preference in the presence of trade-offs
- find a data partitioning (a clustering) that best fits an exploratory domain model

**Initial Sights**
- key projections as obtained via PCA, projection pursuit, ..
- key clusterings as obtained via k-means, affinity propagation, ..

**In the tour...**
- refine outcome of these automated routines
- tune their parameters to better fit specific goals
User Interface: Sight Map

Sight map with sight glyphs of sub-space projections

Control panel

Glyph currently displayed in the local sight explorer

Vector component bar chart display
User Interface: Sight Map

Sight map with sight glyphs of sub-space projections

Control panel

Glyph currently displayed in the local sight explorer

Vector component bar chart display
User Interface: Sight Explorer

- Dynamic multivariate scatterplot display
- N-D touchpad polygon
- Touchpad configuration controls
- Data axes vector display
- Scatterplot controls
- Vector component bar chart display
Dynamic Scatterplots

Interaction to help ‘see’ N-D
  - user interface is key → N-D Navigator™

Motion parallax beats stereo for 3D shape perception
  - the same is true for N-D shape perception
  - help perception by illustrative motion blur
Navigation and Control

Elemental component is the polygonal touchpad
- allows navigation of projection plane in N-D space
- get axis vectors using generalized barycentric interpolation

\[
\vec{w}_3 = \frac{\cot(\alpha) + \cot(\beta)}{|| \vec{p} - \vec{v}_3 ||^2}
\]

\[
p = \sum_{i=1}^{N} a_i \vec{v}_i \quad \text{where} \quad a_i = \frac{w_i}{\sum_{k=1}^{N} w_k}
\]
Vertex order and presence defines the reachable subspace

- at least for $M > 3$

For a 4D space:

<table>
<thead>
<tr>
<th>dimensionality</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td># hyper spheres</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>hyper sphere dimension sets</td>
<td>1,2,3,4</td>
<td>12,13,14,23,24,34</td>
<td>123,124,134,234</td>
<td>1234, 1342, 1432</td>
</tr>
<tr>
<td>touchpad navigation</td>
<td>move to vertex</td>
<td>move along edge</td>
<td>move within triangle</td>
<td>move within poly, re-order</td>
</tr>
</tbody>
</table>

We control this via vertex weights

- setup by prior PCA
Alternative to Touchpad

Directly embed the touchpad into a trackball (under development)

presented at VIS 2012
The Curse of Dimensionality

High-D space tends to be rather sparse

Essentially hypercube is like a “hedgehog”
Relative Contrast

Points are all at about the same distance from one another

- concentration of distances
- fundamental equation: (Bellman, ’61)

\[ \lim_{n \to \infty} \frac{\text{Dist}_{\text{max}} - \text{Dist}_{\text{min}}}{\text{Dist}_{\text{min}}} \to 0 \]

- so as \( n \) increases, it is impossible to distinguish two points by (Euclidian) distance
  - points located in the same cluster of points are OK
  - but we need a better distance metric for far-away points
MDS Layout of 8 N-D Gaussian Clusters

Euclidian

\( m = 6 \) \hspace{1cm} \( m = 40 \) \hspace{1cm} \( m = 100 \) \hspace{1cm} \( m = 800 \)

(a) eMDS
Similarity of N-D Points

Same pattern, with offset

Same pattern, with scaling
A Pattern-Based Distance Metric

... with respect to high dimensional signatures

The Structural Similarity

\[
SSIM = \frac{2\mu_1\mu_2}{\mu_1^2 + \mu_2^2} \cdot \frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} \cdot \frac{\sigma_{12}}{\sigma_1\sigma_2}
\]

- maximized when the two points coincide

- similar mean
- similar variance
- good correlation
SSIM Cases

- Perceptual similarity
- Resulting MDS layout good match with parallel coordinate visualization
Improves Relative Contrast

SSIM pushes the limit of the curse of dimensionality

non-clustered  clustered
SSIM vs. LDA

(c) Parallel Coordinates of the clusters in C2

(d) Parallel Coordinates of the clusters not in C2

(a) sMDS

(b) MDS-LDA
Problem with SSIM layout

- good for far-distances (inter-cluster)
- no notion to gauge near-distances (intra-cluster)
- Euclidian distance is more appropriate here
Problem with SSIM layout

- good for far-distances (inter-cluster)
- no notion to gauge near-distances (intra-cluster)
- Euclidian distance is more appropriate here

Euclidian  |  SSIM  |  Bi-Scale
Bi-Scale Layout Example

with overlap removal

N-D Volume Generation
Why Generate N-D Datasets

Testing of new algorithms

Dataset editing and correcting

Dataset refinement
Interface

- Scatterplot display
- N-D touchpad Polygon
- Touchpad Polygon control panel
- Sketching Controls

Data axes vector display
- Brushing Controls
- Scatterplot display controls
- Vector component bar chart display
Interaction with Axis-aligned Scatterplots

- Point Generation
- Distribution painting backprojection
Interaction with Axis-aligned Scatterplots

- **Point Generation**
  - Distribution Backprojection
    - 2D distribution $\rightarrow$ probability map
    - Randomize the values of the other $N-2$ dimensions
Interaction with Axis-aligned Scatterplots

- Point Sculpting
- Distribution Carving
- Repair
Interaction with Axis-aligned Scatterplots

- **Point Sculpting**
  - Distribution Repair
    - Randomize undefined dimensions’ values
    - Joint probability map of defined dimensions
Interaction with non Axis-aligned Scatterplots

- Point Generation
Interaction with non Axis-aligned Scatterplots

- **Point Generation**
  - Distribution Backprojection

  - **Gram-Schmidt** new coordinate system

\[
\begin{align*}
  y_1 &= x_1, & e_1 &= y_1 / ||y_1|| \\
  y_2 &= x_2 - \text{proj}_{y_1}(x_2), & e_2 &= y_2 / ||y_2|| \\
  y_3 &= x_3 - \text{proj}_{y_1}(x_3) - \text{proj}_{y_2}(x_3), & e_3 &= y_3 / ||y_3|| \\
  & \vdots \\
  y_N &= x_N - \sum_{j=1}^{N-1} \text{proj}_{y_j}(x_N), & e_N &= y_N / ||y_N||
\end{align*}
\]
Interaction with non Axis-aligned Scatterplots

- **Point Sculpting**
  - Distribution Carving

presented at VAST 2013
The ND-Scope

- Parallel Coordinate Plot
- Sight Map
- Network Display
- Geospatial Display
- Spreadsheet
- Dynamic Scatterplot
Recap

Extending 2D/3D paradigms into N-D has good merit
- makes interactions with N-D data more natural and intuitive
- turns design task into paradigm extension and not invention
- curse of dimensionality can (most likely, will) cause headache
- but challenges are opportunities and the essence of research

Some examples I’ve demonstrated
- interaction and navigation via ND touchpad and trackball
- view suggestion and selection via subspace map
- ND distance metric motivated by image perception
- data generation via CT backprojection techniques extended to ND

The good news is...
- there are still many 2D/3D mechanisms and paradigm to port
- rotation matrices, illustrative rendering, shape understanding, ...
Support and Collaborators

Support from NSF, NIH, DOE, BNL, PNL, CEWIT, MKE

Faculty:
- Kevin T. McDonnell (Dowling College), IV Ramakrishnan, Erez Zadok, Tamara Berg (SBU Computer Science Department),

Domain scientists:
- Dr. Alla Zelenyuk, Dr. Dan Imre (PNL), Yangan Liu (BNL)

PhD students (at Stony Brook and SUNY Korea):
- Zhiyuan Zhang, Hyunjung Lee, Nafees Ahmed, Bing Wang, Puripant (Joe) Ruchikachorn, Sungsoo Ha,, Jisung Kim, Eric Papenhausen, Salman Mahmood, Shenghui Cheng, Ziyi Zheng (PhD, now at Amazon). Julia EunJu Nam (PhD, now at Microsoft), Wei Xu (PhD, now at BNL)
Relevant References


More Detail? Visit my Webpage...

Klaus Mueller Turns Data Into Artwork At Rapid Speeds

http://www.cs.stonybrook.edu/~mueller
(for videos see dedicated paper web pages)

http://nd-scope.net

Any questions?