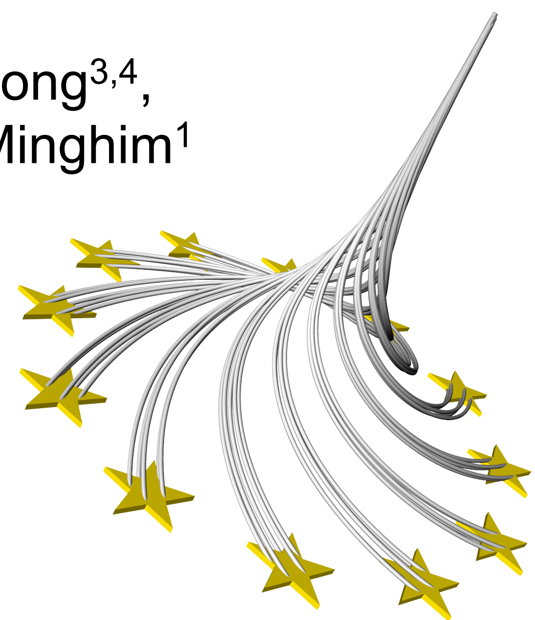
EuroVis 2011

A Framework for Exploring Multidimensional Data with 3D Projections

J. Poco[1,2], R. Etemadpour[3], F.V. Paulovich[1], T.V. Long[3,4], P. Rosenthal[3,5], M.C.F. Oliveira[1], L. Linsen[3] and R. Minghim[1]

1University of São Paulo, São Carlos, Brazil
2 University of Utah, USA
3 Jacobs University, Bremen, Germany
4 University of Transport and Communication Hanoi, Vietnam
5 Chemnitz University of Technology, Germany



Introduction

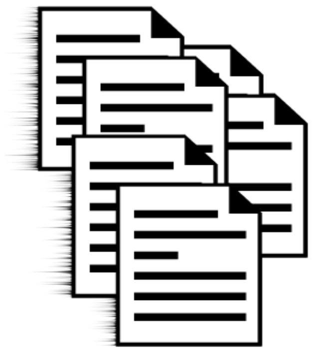
- MDS for visualization
- Intuitively 3D is better at group separation
- But difficult to interact with
- Framework to create and interact with 3D data projections

Outline

- Motivation
- Projections Algorithms
  - Quantitative Analysis
- 3D Cluster Visualization
- The Framework
- User Studies
  - 3D vs. 2D LSP
  - Surfaces Evaluation
- Conclusions

Motivation

Datasets



Document Collection

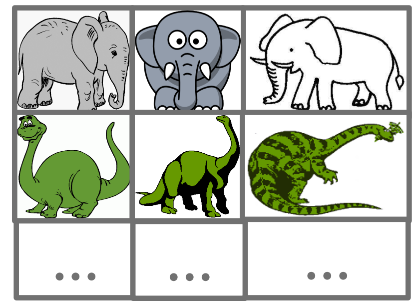
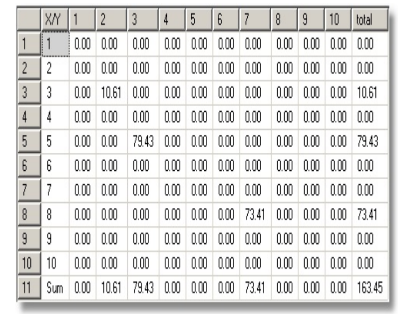


Image Collection



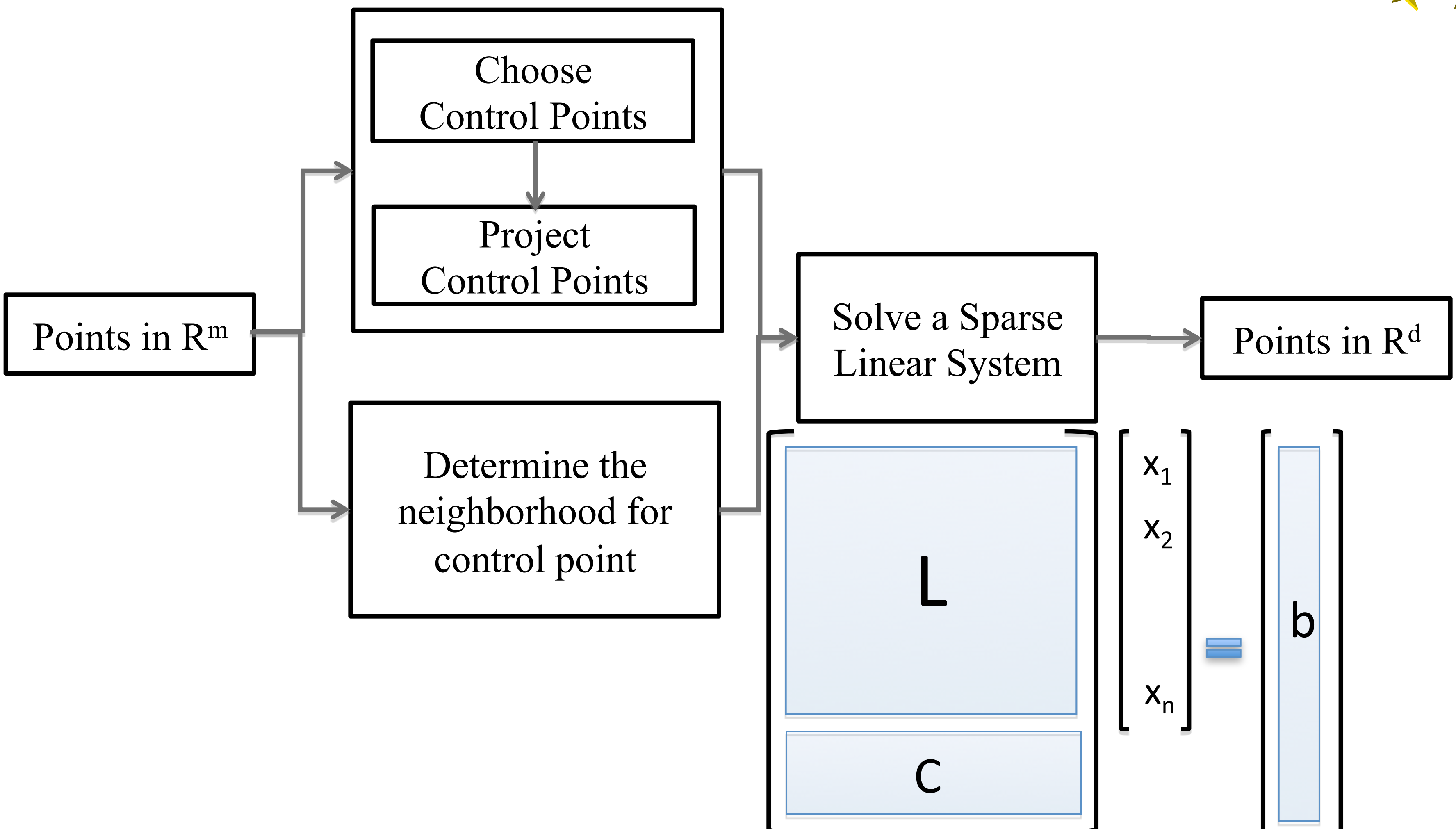
	X/Y	1	2	3	4	5	6	7	8	9	10	total
1	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	3	0.00	10.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.61
4	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	5	0.00	0.00	79.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	79.43
6	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	8	0.00	0.00	0.00	0.00	0.00	0.00	73.41	0.00	0.00	0.00	73.41
9	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	Sum	0.00	10.61	79.43	0.00	0.00	0.00	73.41	0.00	0.00	0.00	163.45

Multi Scalar Field

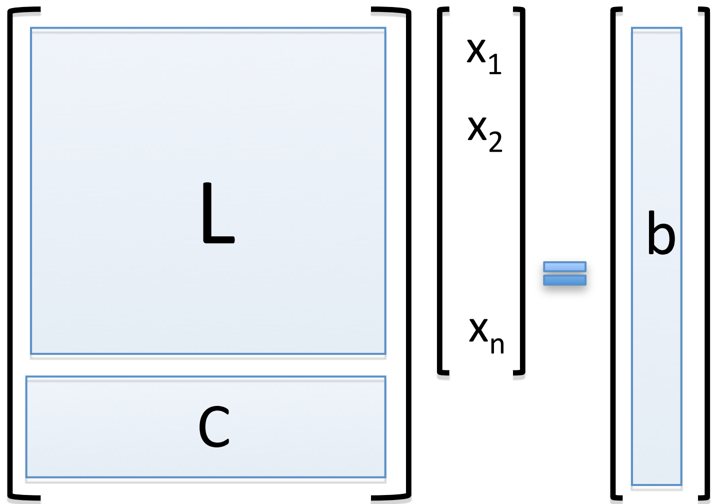
Questions

- Are there well-defined groups of similar objects?
- How are different groups related?
- What about instances within a group?
- Which data features determine the grouping?

Least Square Projection (LSP)



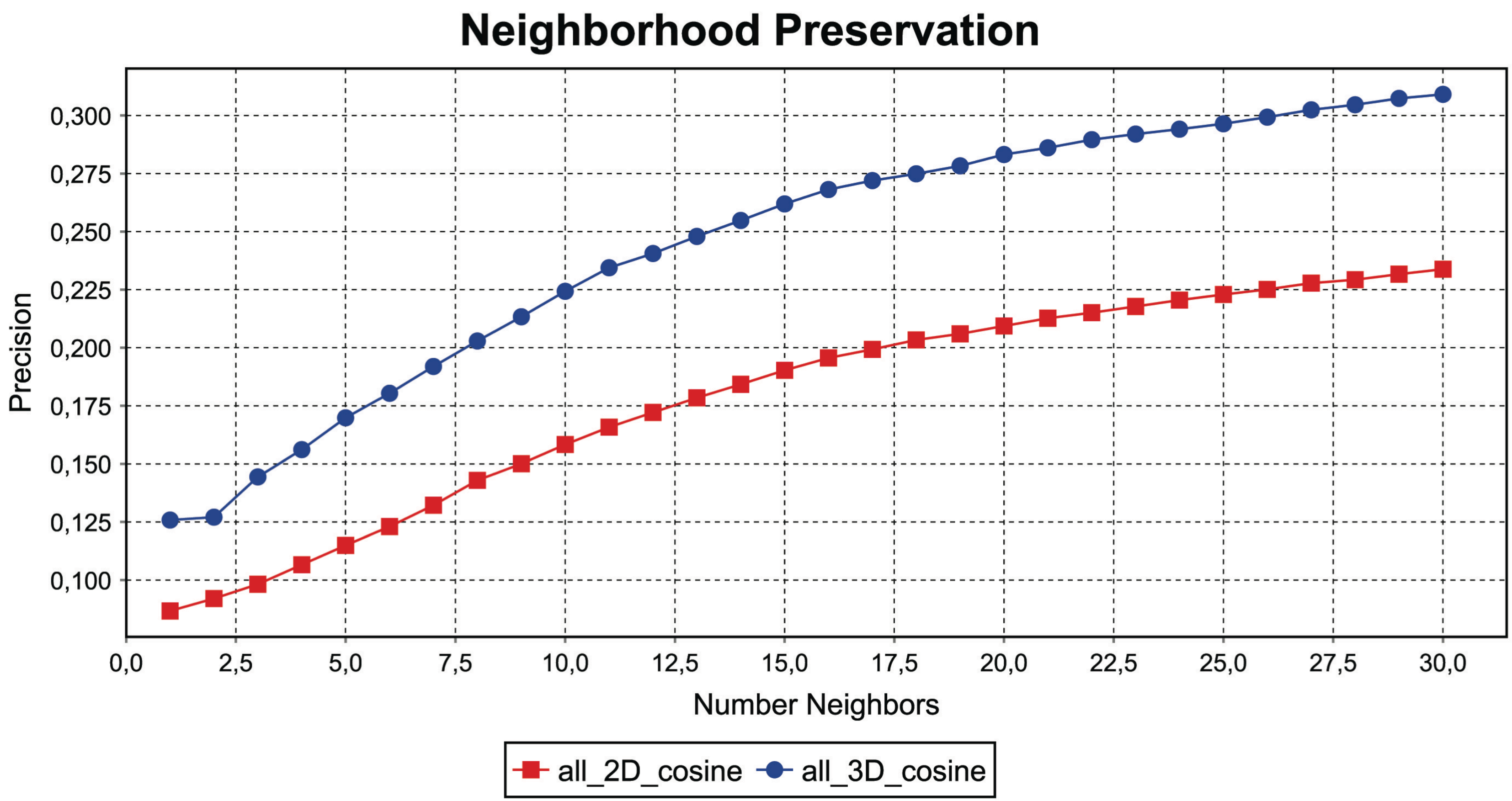
Extending LSP to 3D



- This system is solved for each dimension in visual space.
- In 3D case, we need to solve 3 linear systems.

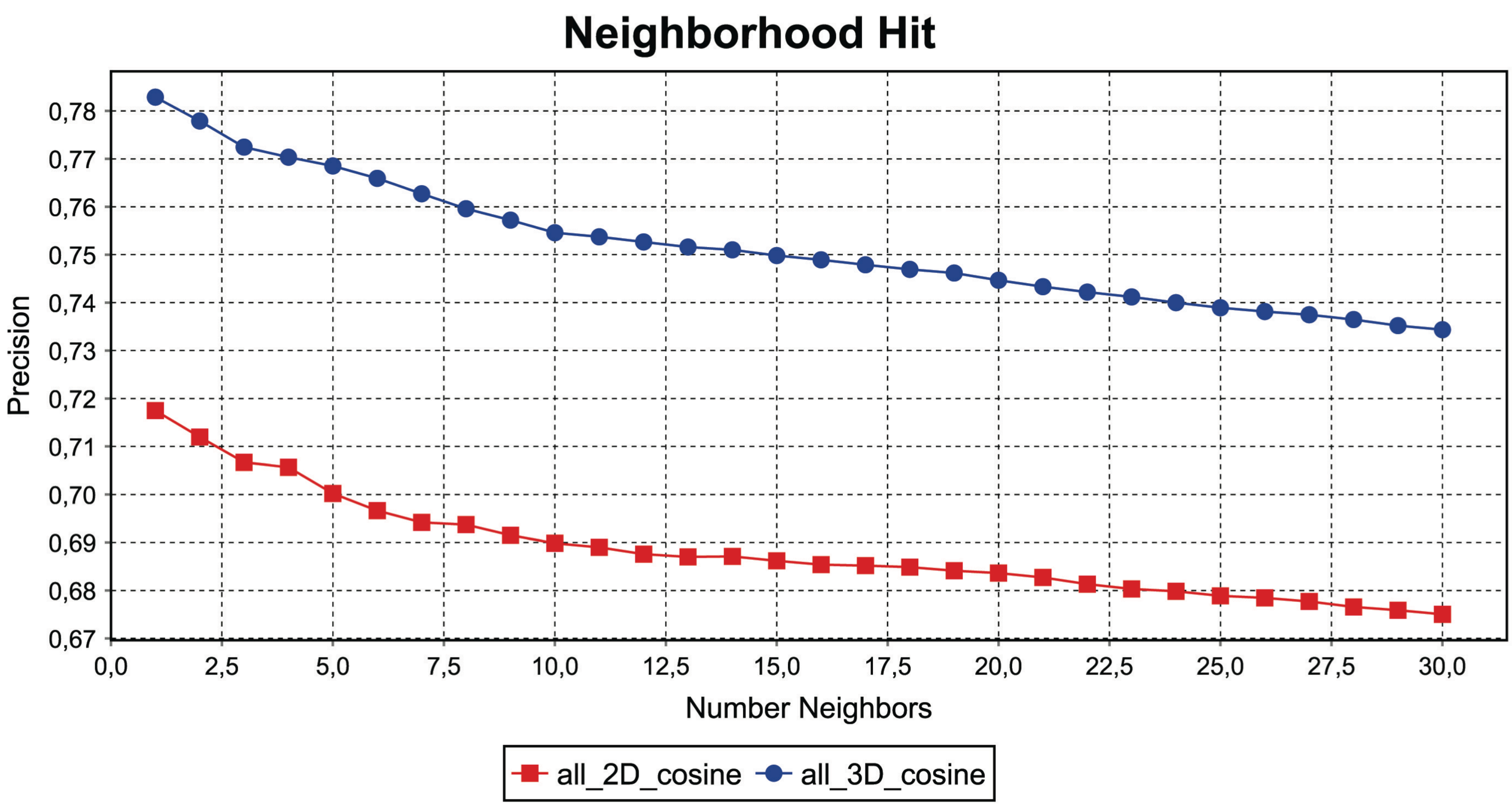
LSP 3D: Quantitative Evaluation

Neighborhood preservation - neighborhood after projection



LSP 3D: Quantitative Evaluation

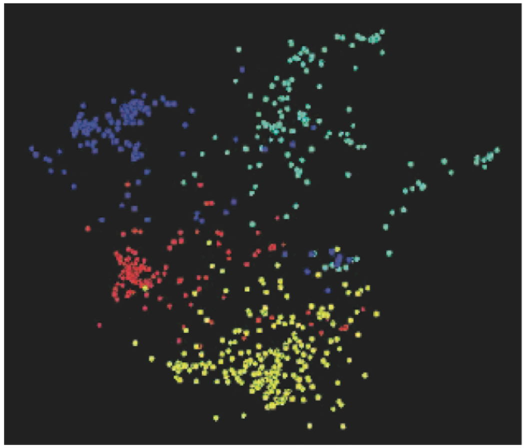
Neighborhood Hit – neighbors' classes after projection



3D Cluster Visualization

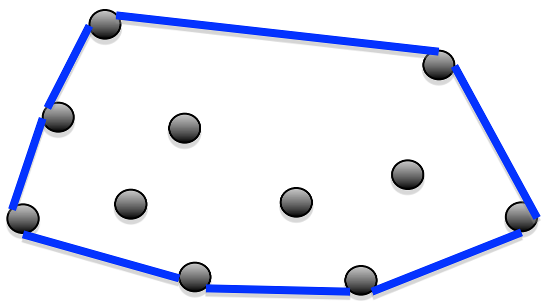
Color coded:

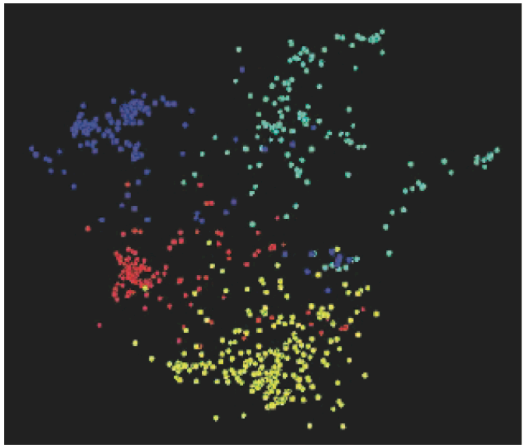
- ✓ Fast and robust.
- ✗ Loss in perception of depth .

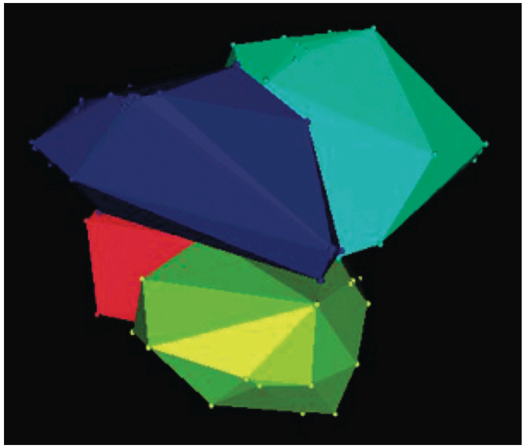


3D Cluster Visualization

Convex Hull:



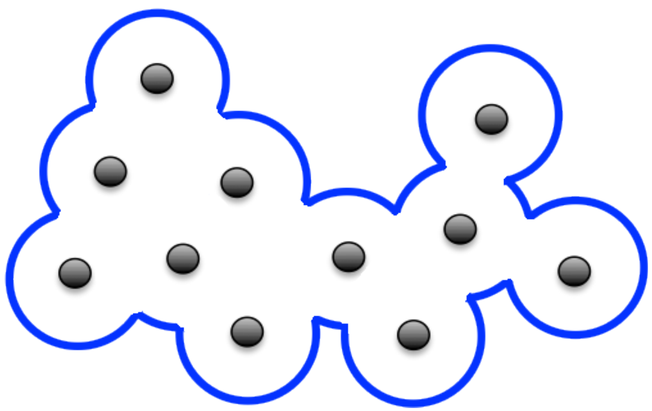


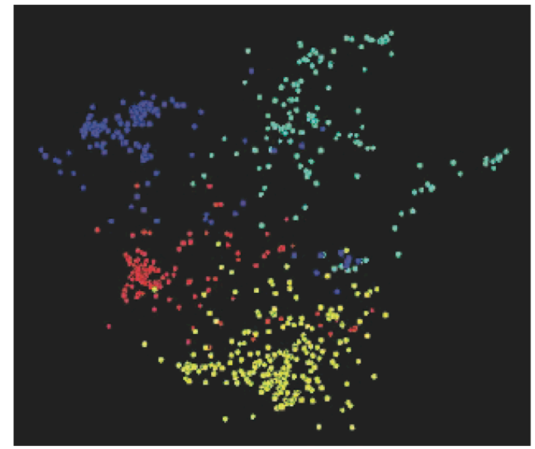


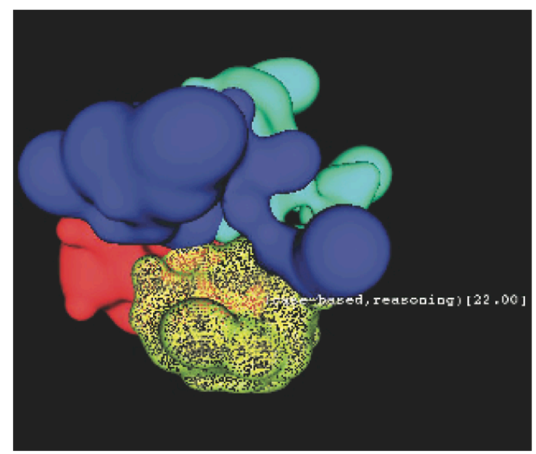
- ✓ Fast, robust and simple meshes.
- ✓ Effective in convex clusters.
- x Misinterpretation in non-convex clusters.

3D Cluster Visualization

Enclosing Surface iso-distance to cluster points:



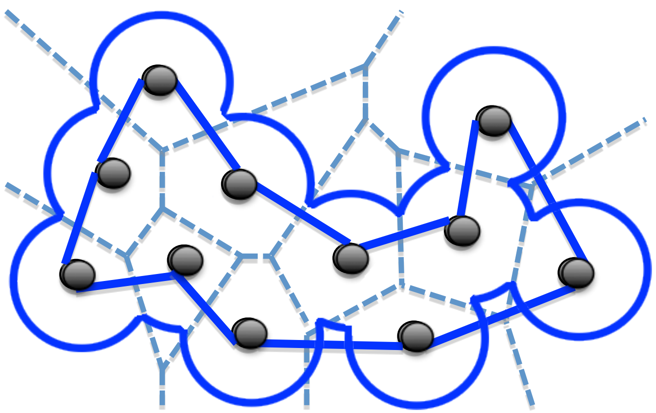


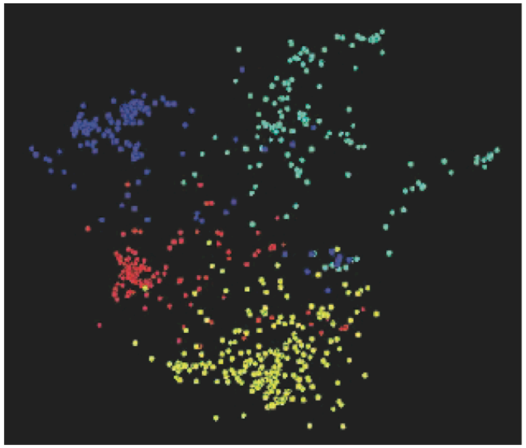


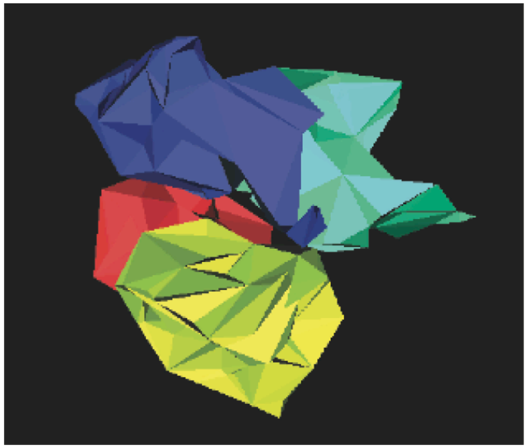
- ✓ Works for non-convex clusters.
- ✓ Visually pleasing rendering.
- x Adaptive adjustment of the radius of influence when clusters come close to each other.

3D Cluster Visualization

Non-Convex Hull:



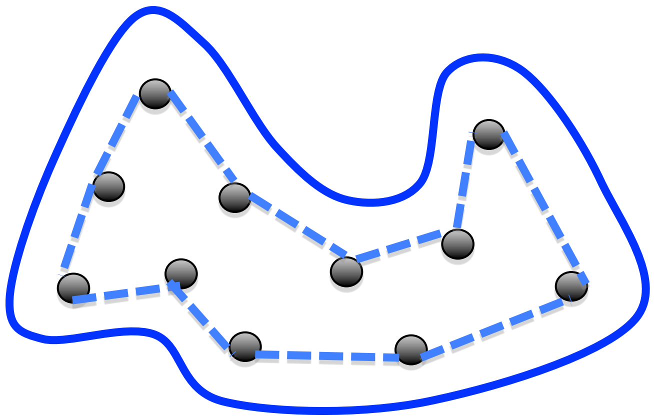


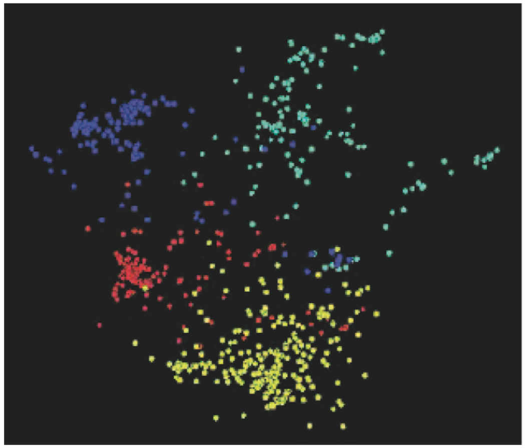


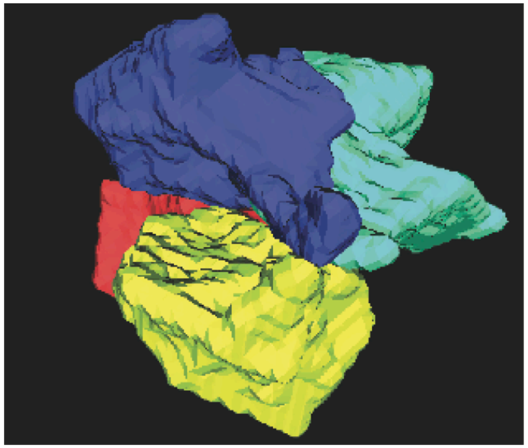
- ✓ Doesn't need radius adjustment.
- ✓ A distance field is created computing Voronoi diagram in GPU
- ✓ Works for non-convex cluster.

3D Cluster Visualization

Enclosing surfaces iso-distance to non-convex Hull:

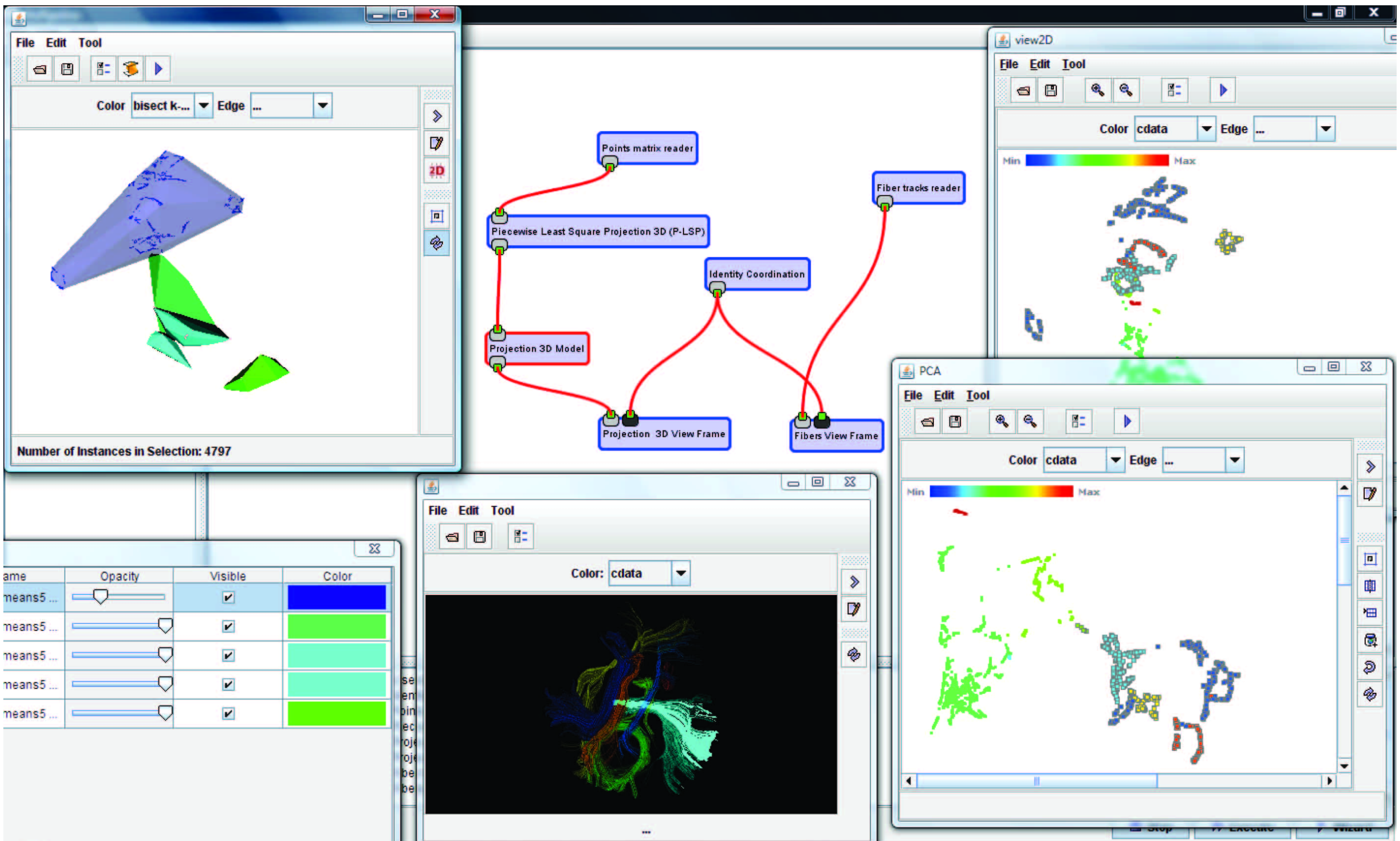






- ✓ Doesn't need radius adjustment.
- ✓ A distance field is created computing Voronoi diagrams in GPU
- ✓ Works for non-convex cluster.

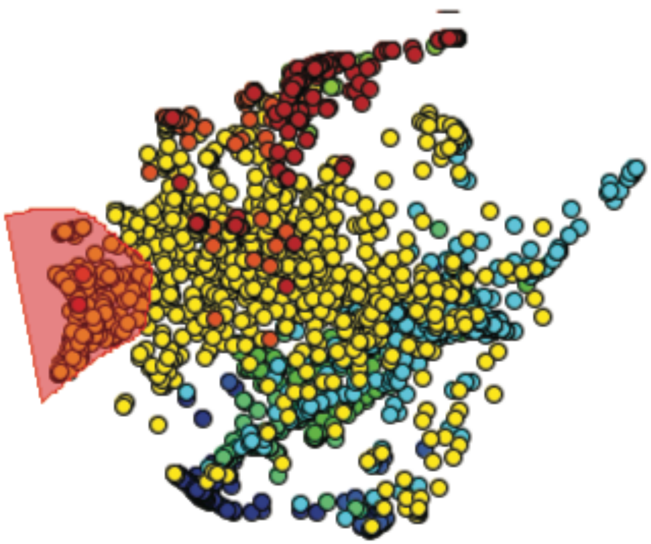
The System

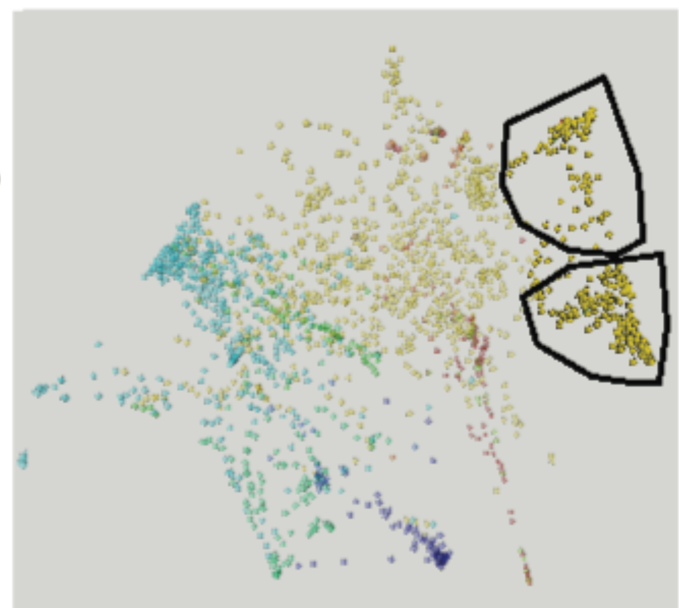


File Edit Tool
Color bisect k-... Edge ...
Number of Instances in Selection: 4797
Points matrix reader
Piecewise Least Square Projection 3D (P-LSP)
Identity Coordination
Fiber tracks reader
Projection 3D Model
Projection 3D View Frame
Fibers View Frame
view2D
Color cdata Edge ...
Min Max
PCA
Color: cdata
Opacity Visible Color

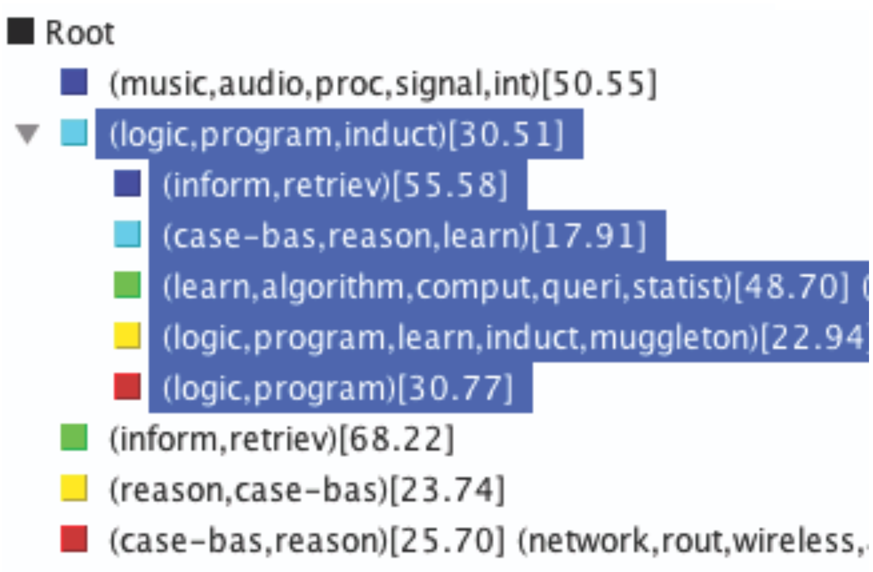
Coordination of 2D and 3D views.

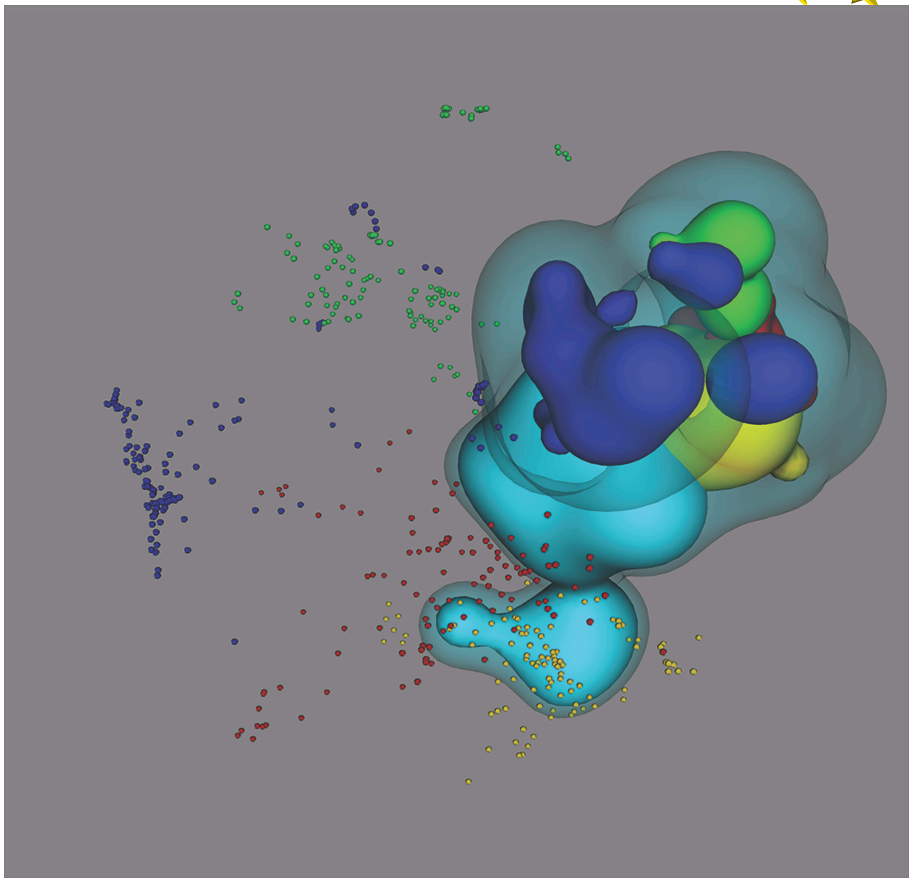
- From 3D to 2D: selecting enclosing surface.
- From 2D to 3D: selecting well "resolved" groups that are preserved in both spaces.
- From 2D to 3D: selecting hidden groups



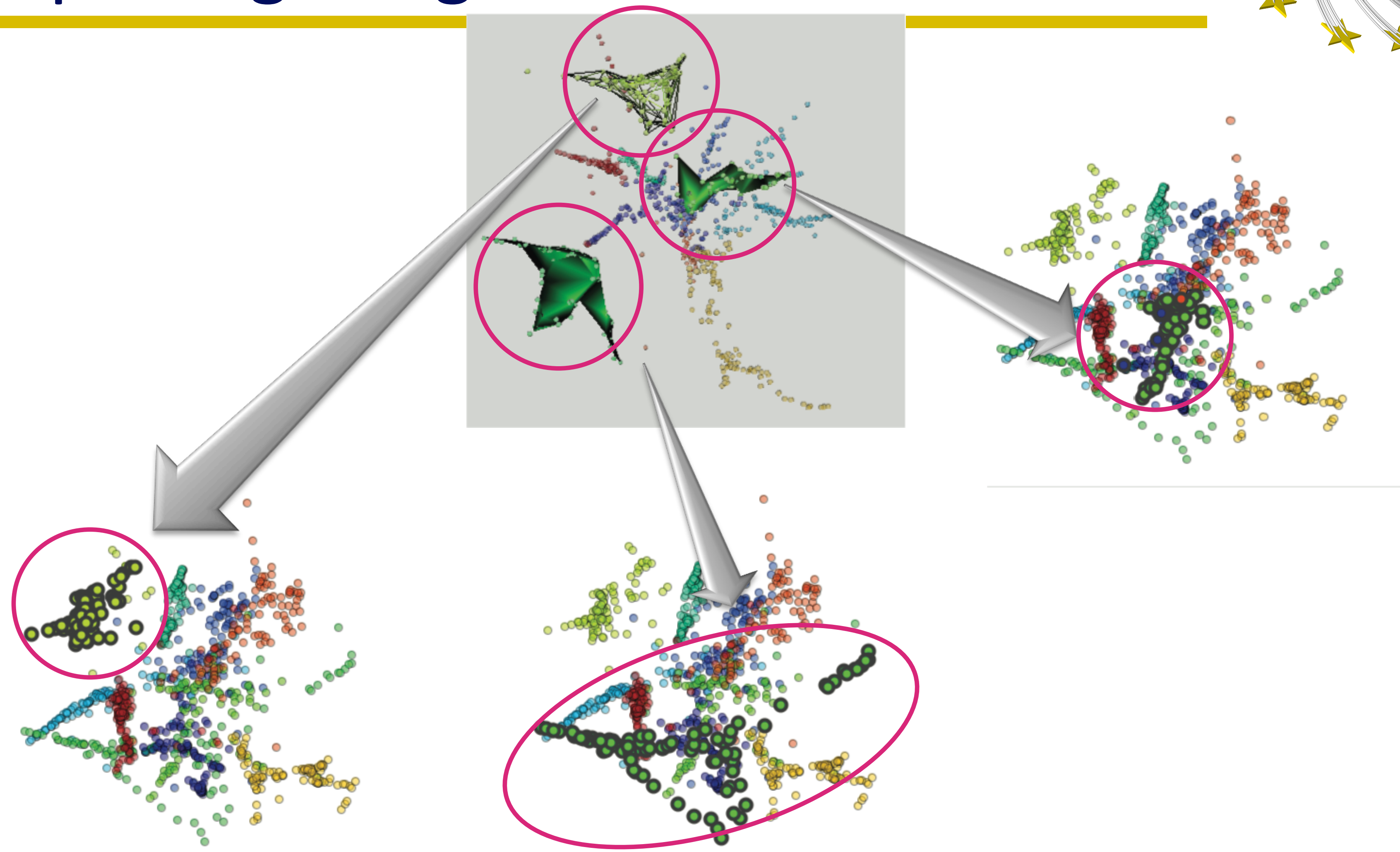


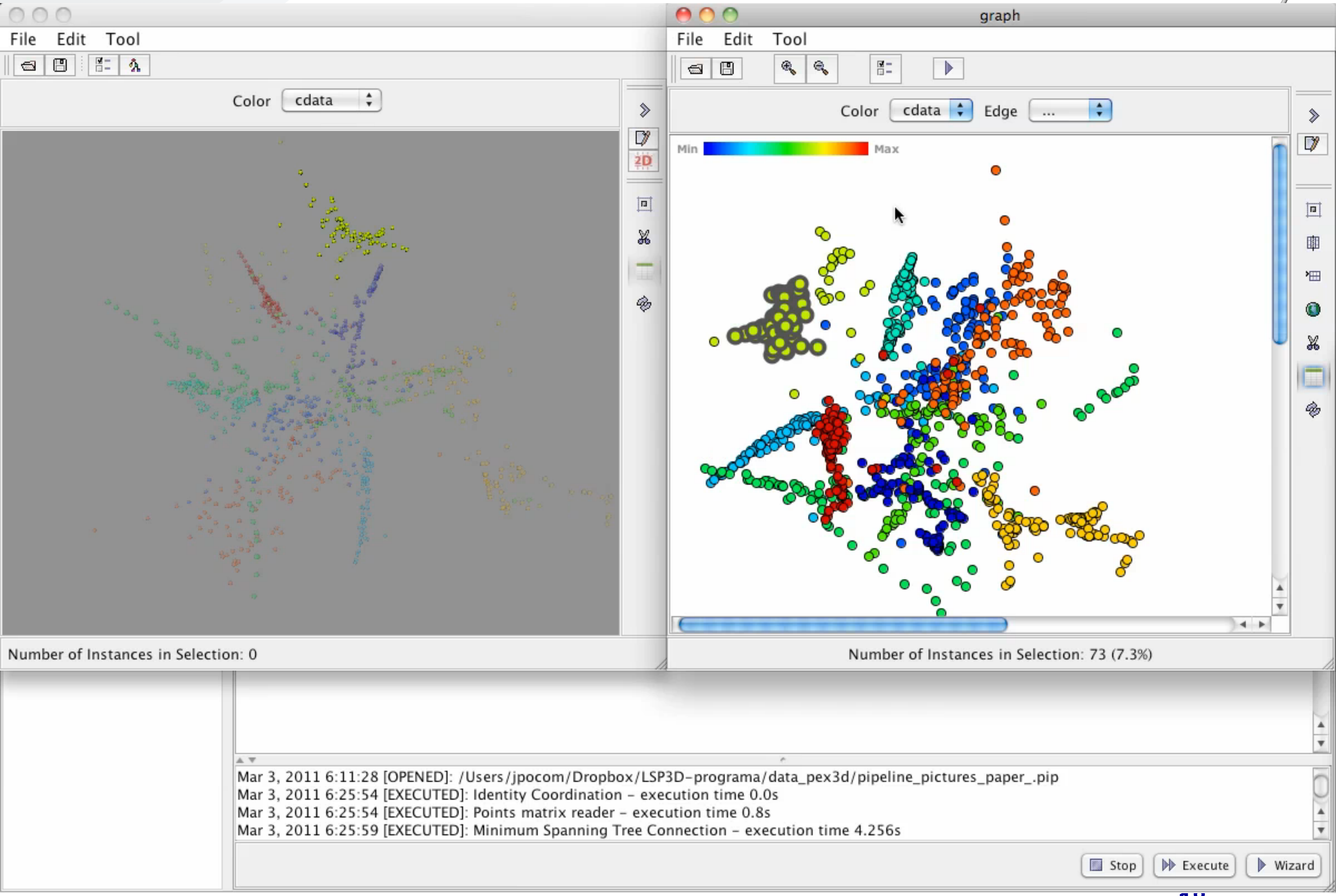
Hierarchy Clustering





Exploring image sets




File Edit Tool
Color cdata
Number of Instances in Selection: 0
graph
File Edit Tool
Color cdata Edge ...
Min Max
Number of Instances in Selection: 73 (7.3%)
Mar 3, 2011 6:11:28 [OPENED]: /Users/jpocom/Dropbox/LSP3D-programa/data_pex3d/pipeline_pictures_paper_.pip
Mar 3, 2011 6:25:54 [EXECUTED]: Identity Coordination - execution time 0.0s
Mar 3, 2011 6:25:54 [EXECUTED]: Points matrix reader - execution time 0.8s
Mar 3, 2011 6:25:59 [EXECUTED]: Minimum Spanning Tree Connection - execution time 4.256s
Stop
Execute
Wizard

Evaluation user study set up

- 12 participants
- Experience with computers
- Short training session
- 2 datasets
  - Document data (681 objects, 2993 dimensions)
  - Medical imaging data (1000 objects, 150 dimensions)

Quality Measures

- Reliability → Correctness
- Efficiency → Response time
- Learn ability → User satisfaction
- User-friendliness → User preferences

Set Up

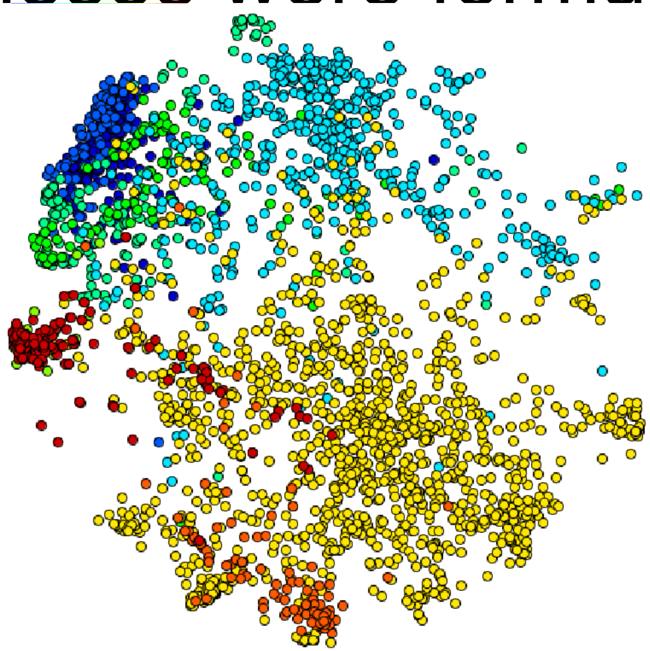
- Set of tasks
- Record time
- Participants confidence using Likert scale
- User preference

Statistical Significance

- Shapiro-Wilk test against normal distribution
- 2 projections
  - Wilcoxon Matched-Pairs Signed-ranks test
  - T-test
- 5 cluster visualizations
  - Friedman's χ^2-test or test for repeated measures
  - ANOVA test

2D vs. 3D Hypotheses

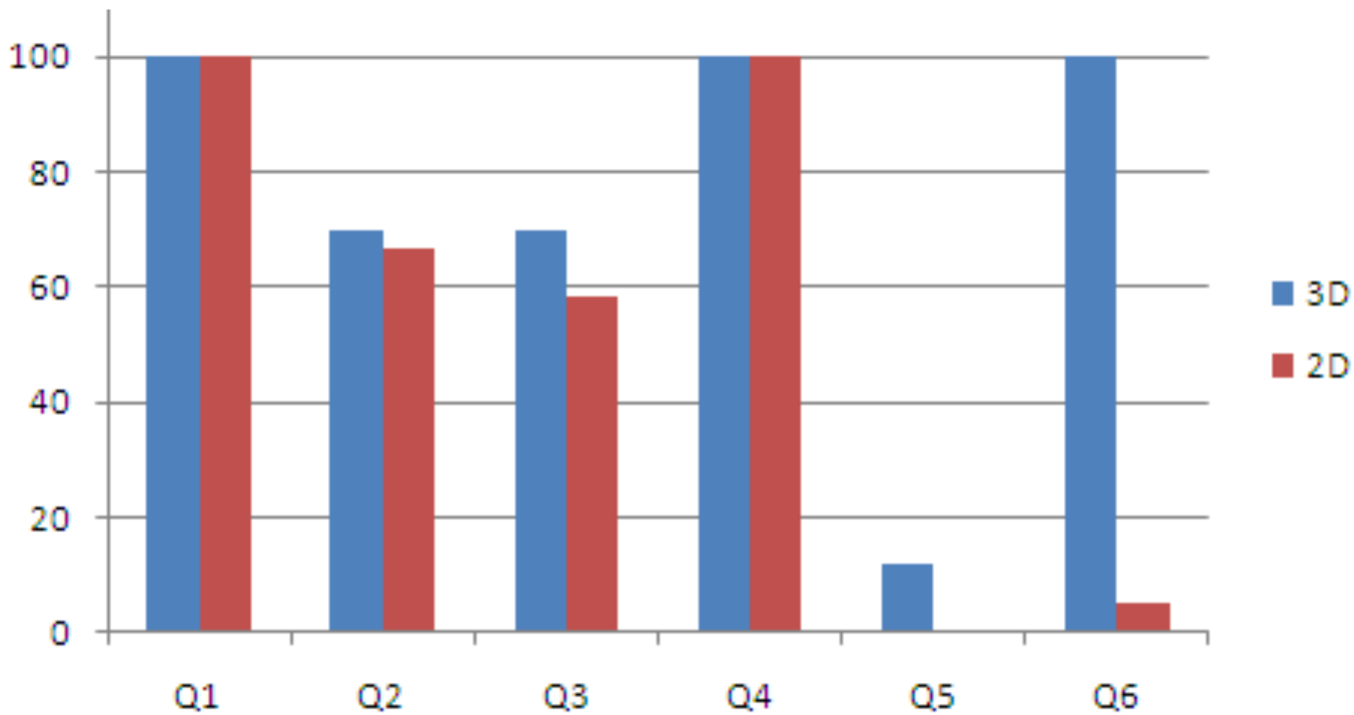
- Data were clustered before projection
- Two hypotheses were formulated
  - 2D projec… to find in 3D.
  - 3D projec… …ation leading to more corr…



Tasks

1) Count the clusters
2) Order the clusters by their density
3) List all overlaps of clusters
4) Detect an object within a cluster
5) Find closest cluster to a specific point
6) Repeat Task 5 with a different point

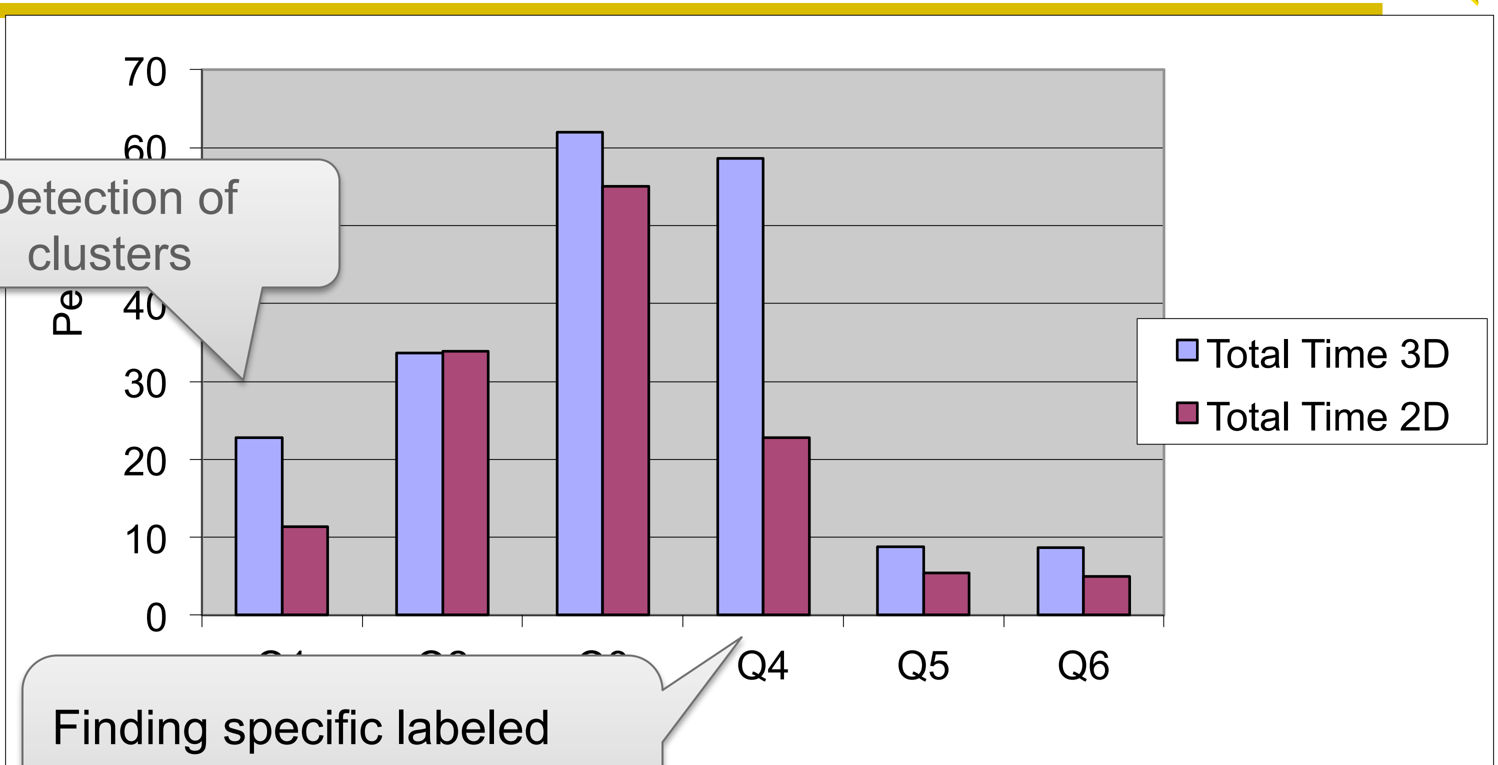
Correctness 2D vs. 3D



		Q3	Q5	Q6
		[illegible]	0.01562	0.01953

The user' s perception on detecting the similarities among clusters.

Response Times 2D vs. 3D



Detection of clusters

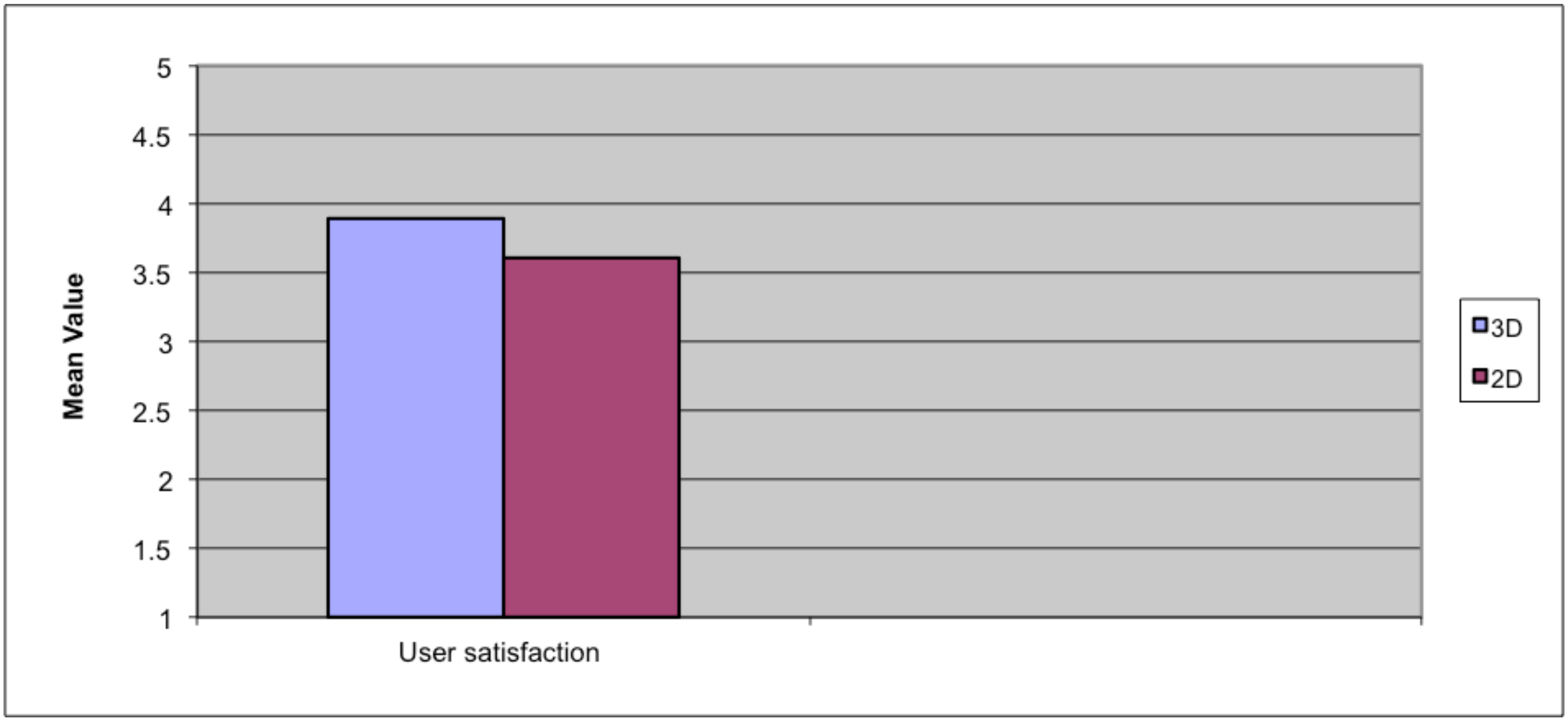
Finding specific labeled point

	Q1	Q2	Q3	Q5	Q6
Asymp. Sig.	0.01855	0.8501	0.5186	0.375	0.4922

Mean	**58.58**	**22.75**

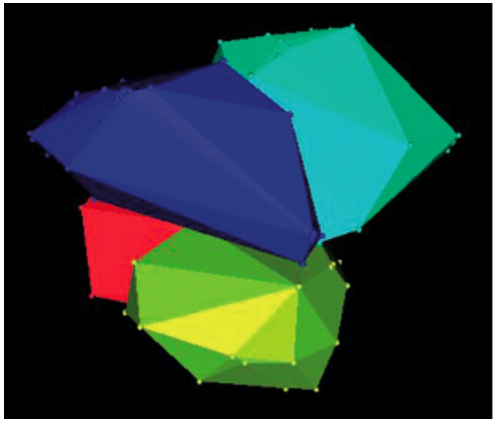
The two-tailed P value equals 0.0189

User Satisfaction

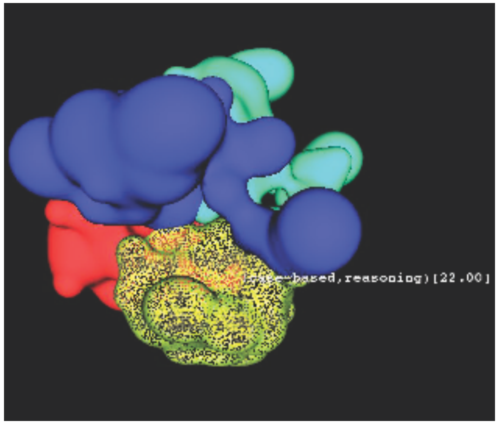


	3D vs. 2D
Asymp. Sig.	0,007812

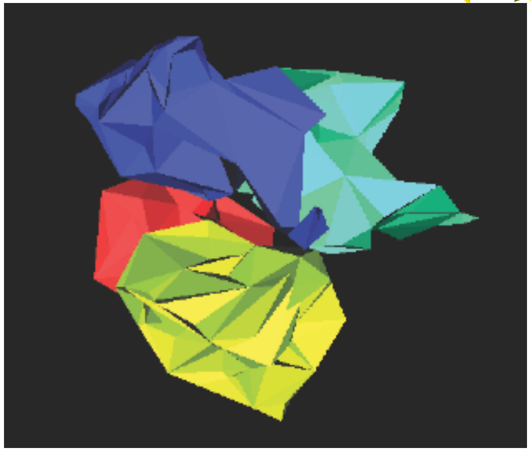
User Study Evaluation
Cluster Visualization



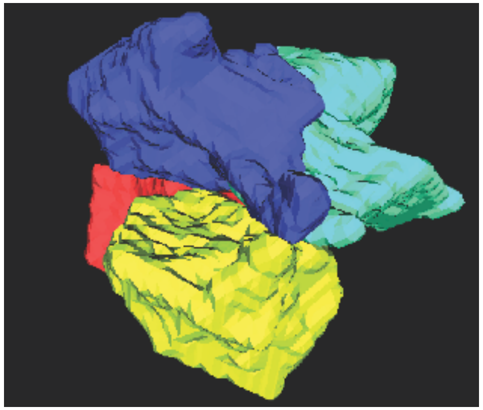
convex hull



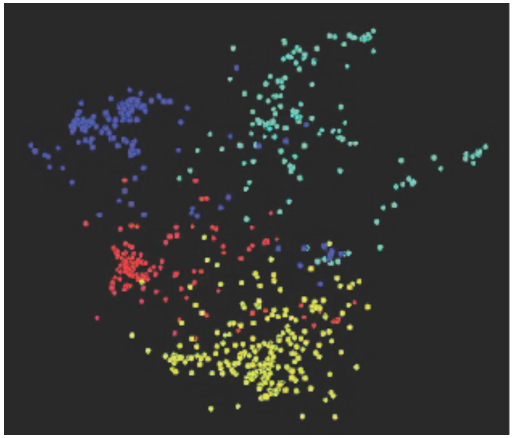
enclosing surfaces isodistant to points



non-convex hull



enclosing surfaces isodistant to non-convex hull



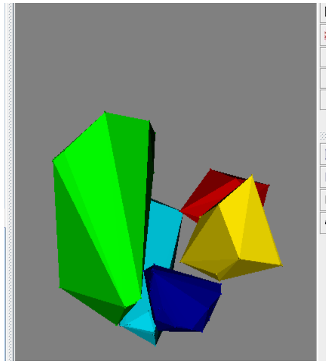
color-coded point cloud.

Tasks

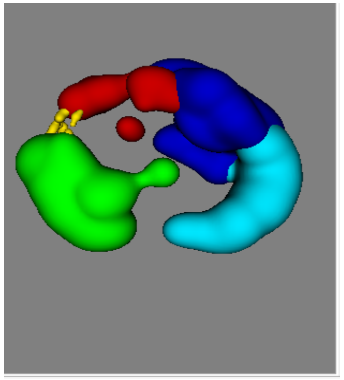
1) Count the clusters
2) List cluster overlaps
3) Identify most distant clusters

Set up

- Different techniques coupled with different datasets
- Half of the participants complementary combinations to the other half

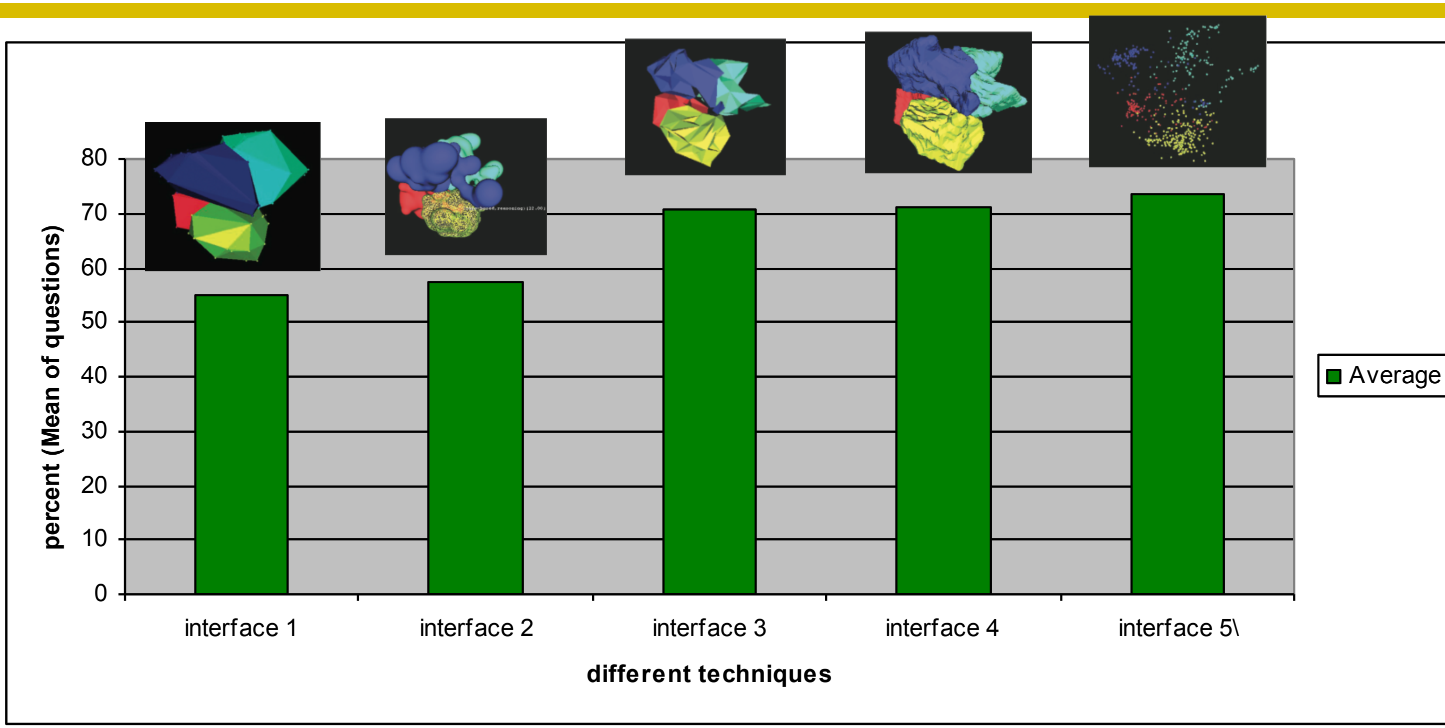


Document data, 5 classes



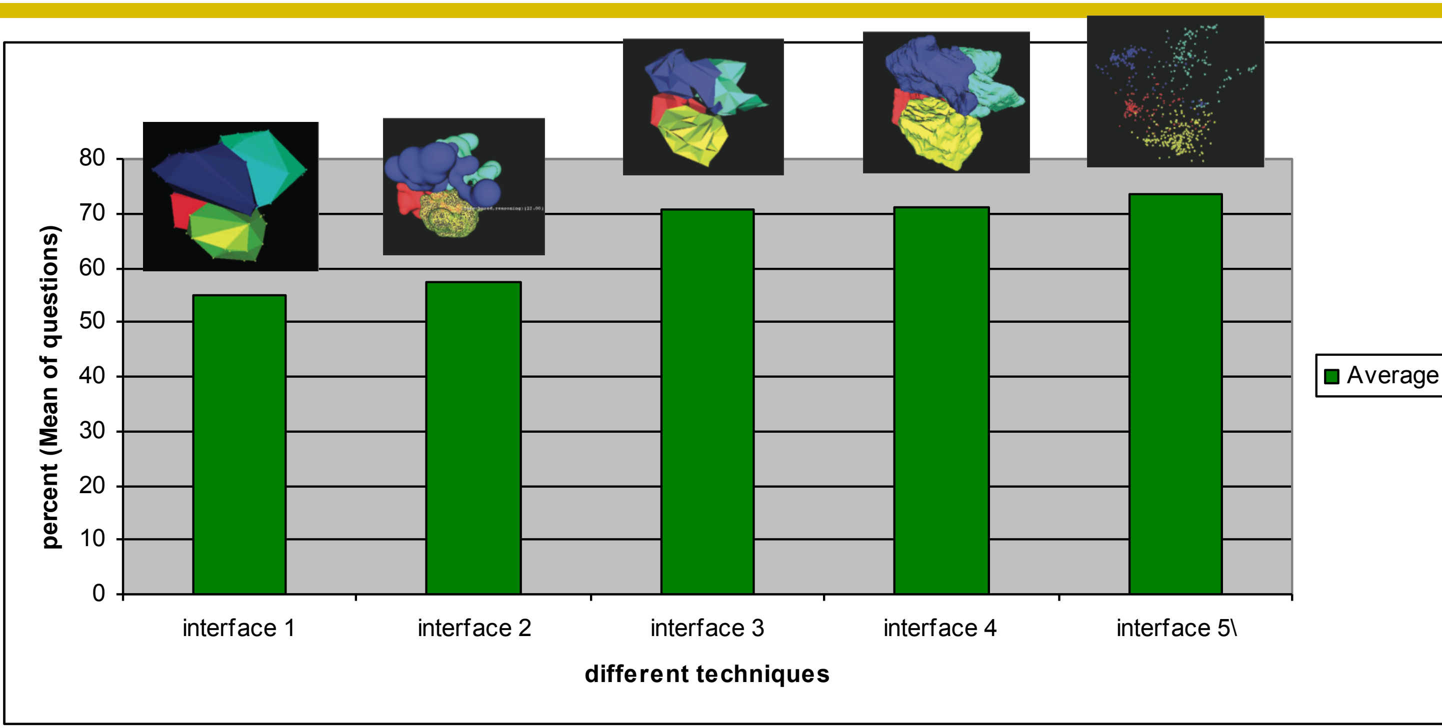
Medical data, 12 classes

Correctness



No significant difference

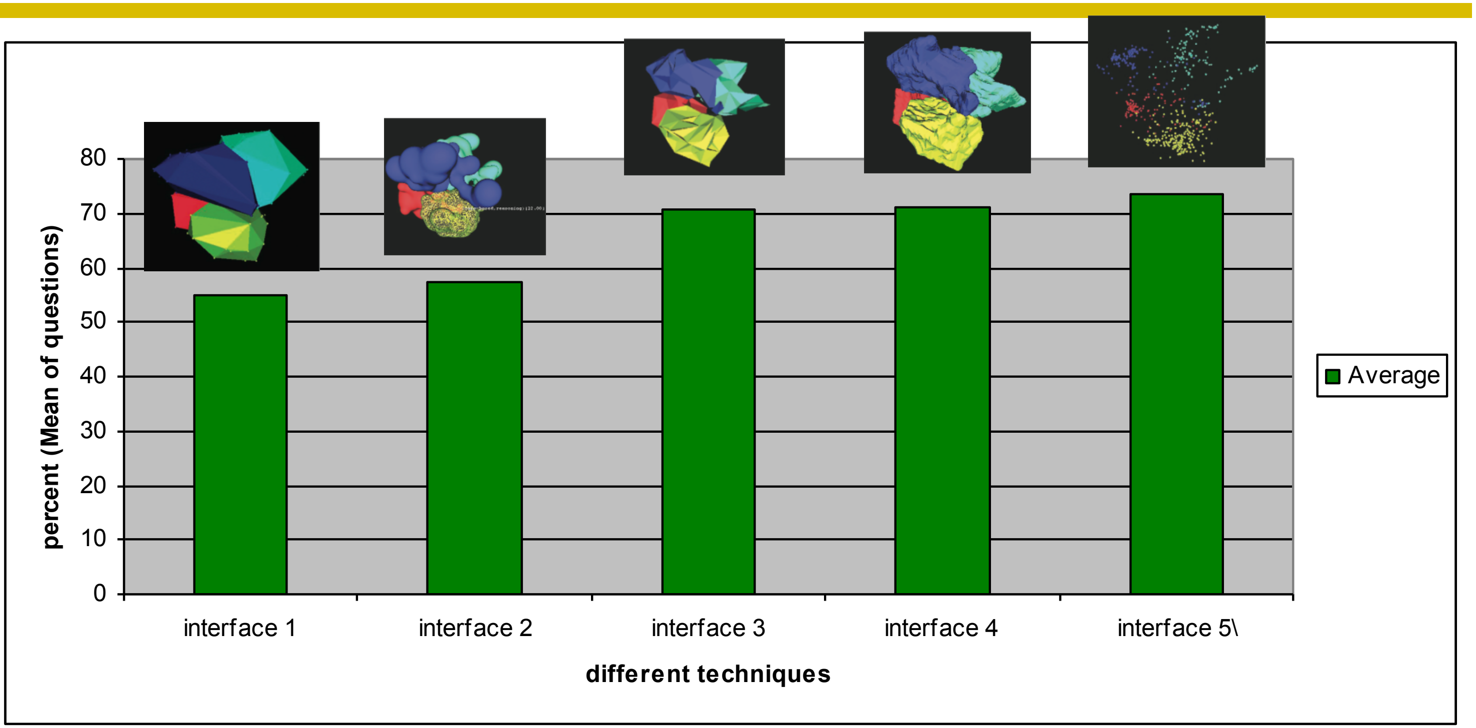
Correctness



	Q2	Q3
Friedman's Chi-Square	11.24	6.8
Asymp. Sig.	**0,025**	0,149

Document Dataset

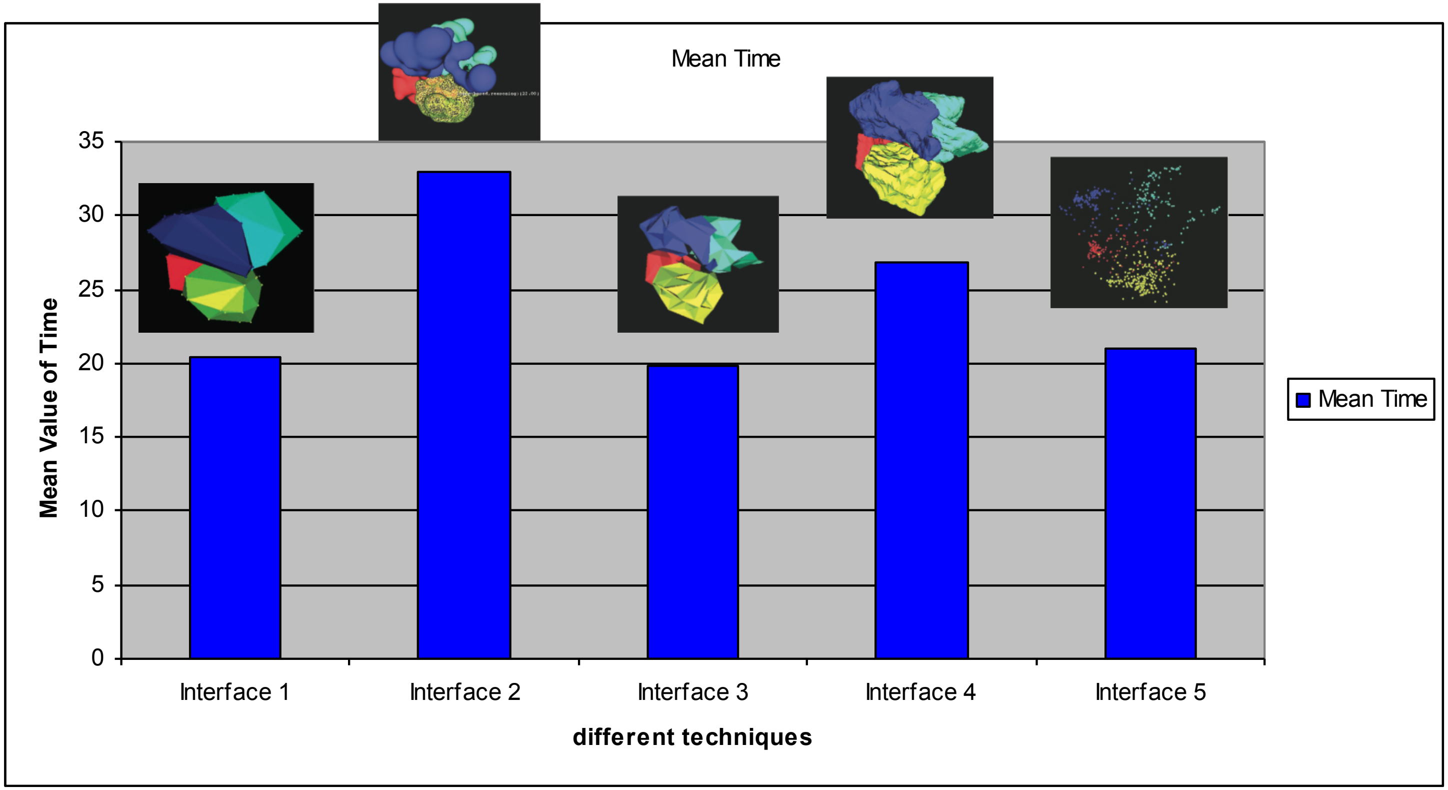
Correctness



	Q2	Q3
Friedman's Chi-Square	9	2,5
Asymp. Sig.	0,063	0,645

Medical Image Dataset

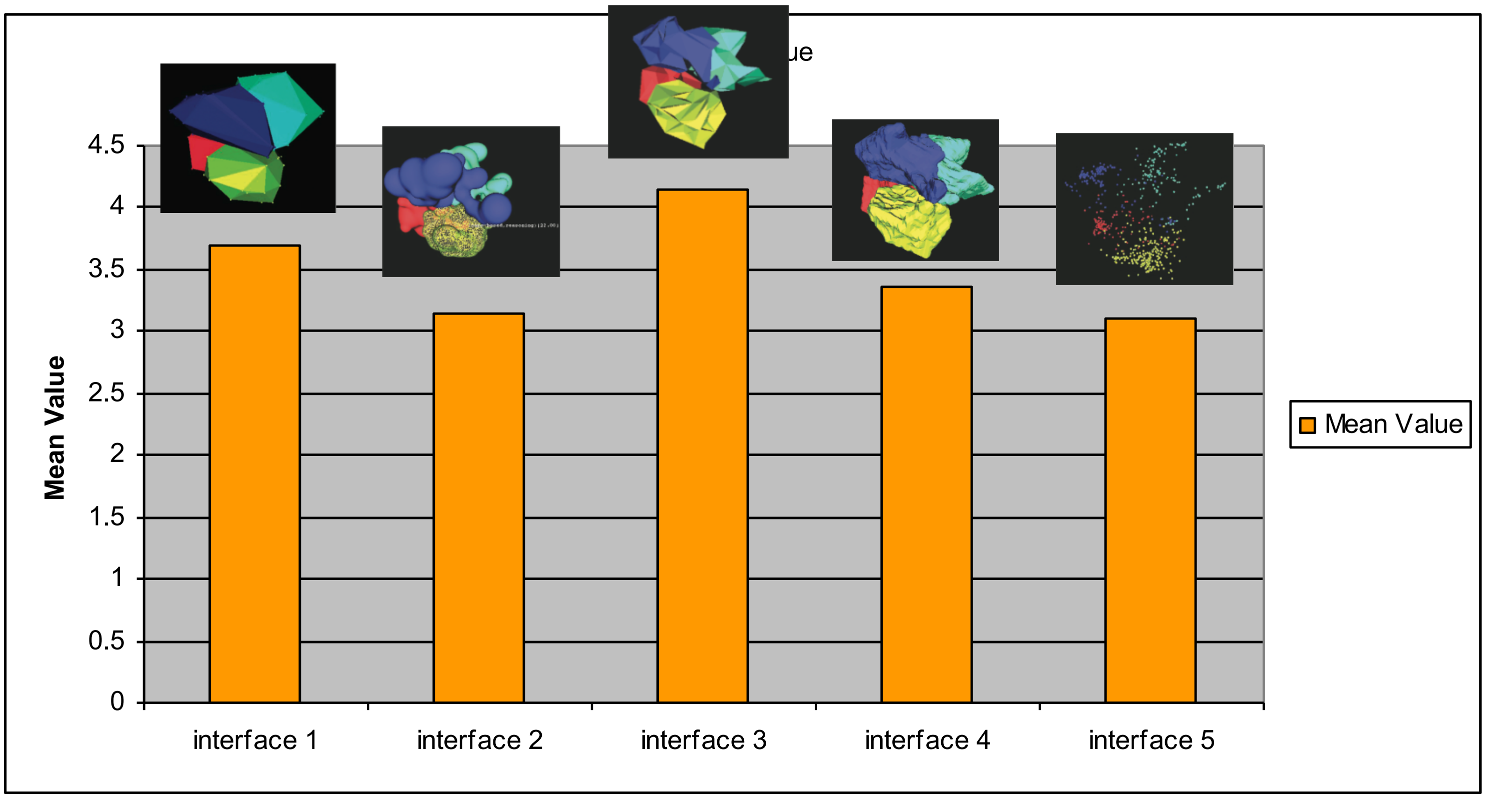
Response Time



Anova Probability = 0.022
Less Accurate
Longer Time

EuroVis 2011

User Satisfaction



Lowest User Satisfaction

Highest rate of satisfaction

No significant difference

Conclusion

- Framework for multidimensional data visualization
- 3D projection
- 3D cluster visualization
- Quantitative evaluation for 2D vs. 3D projection
- Higher precision in 3D
- User study for 2D vs. 3D
- Higher correctness & user satisfaction in 3D
- User study different cluster visualization

Acknowledgements

- FAPESP, CNPq, CAPES
  - Brasil
- DAAD, DFG
  - Germany
- Available at
  http://infoserver.lcad.icmc.usp.br/infovis2/Tools