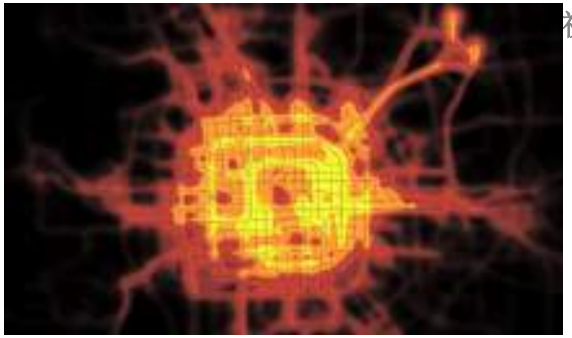
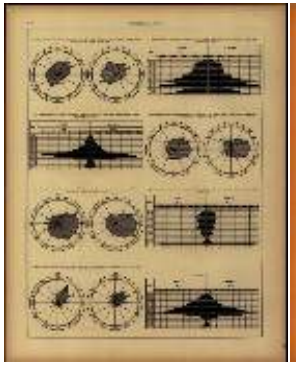
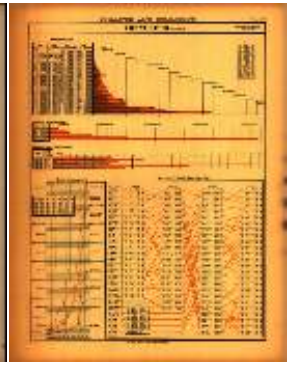
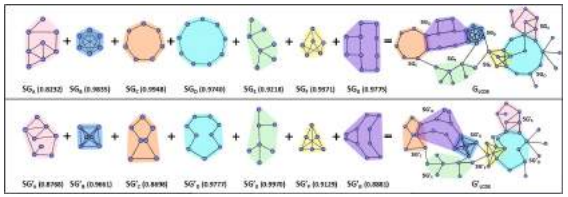
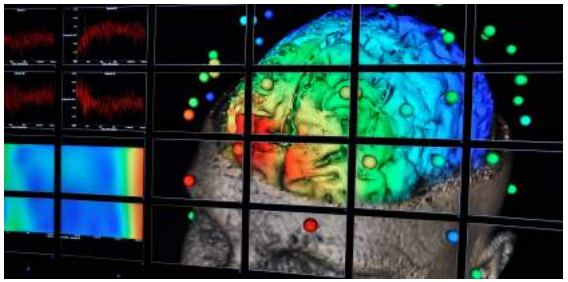










众里寻他千百度

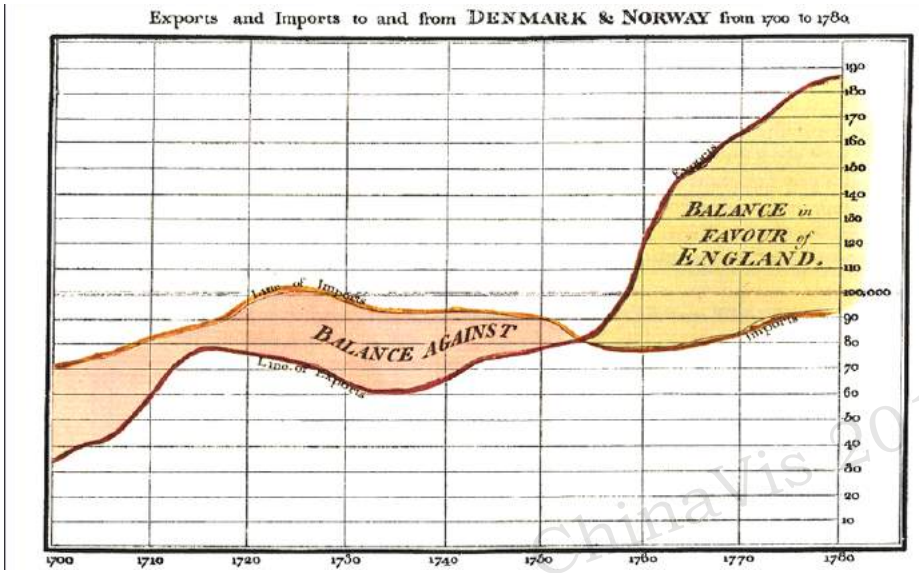
—从几篇可视化论文的起源谈起

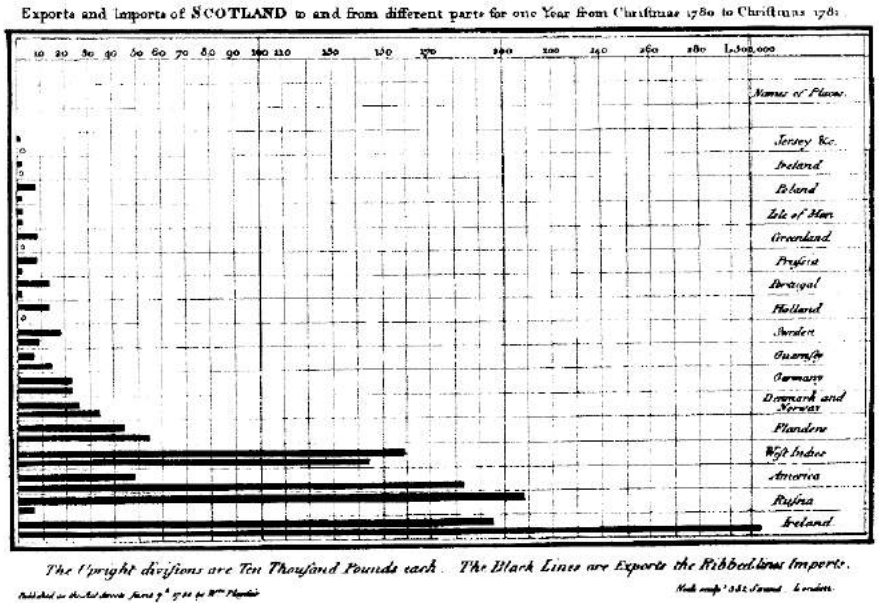
ChinaVis 2016 Tutorial

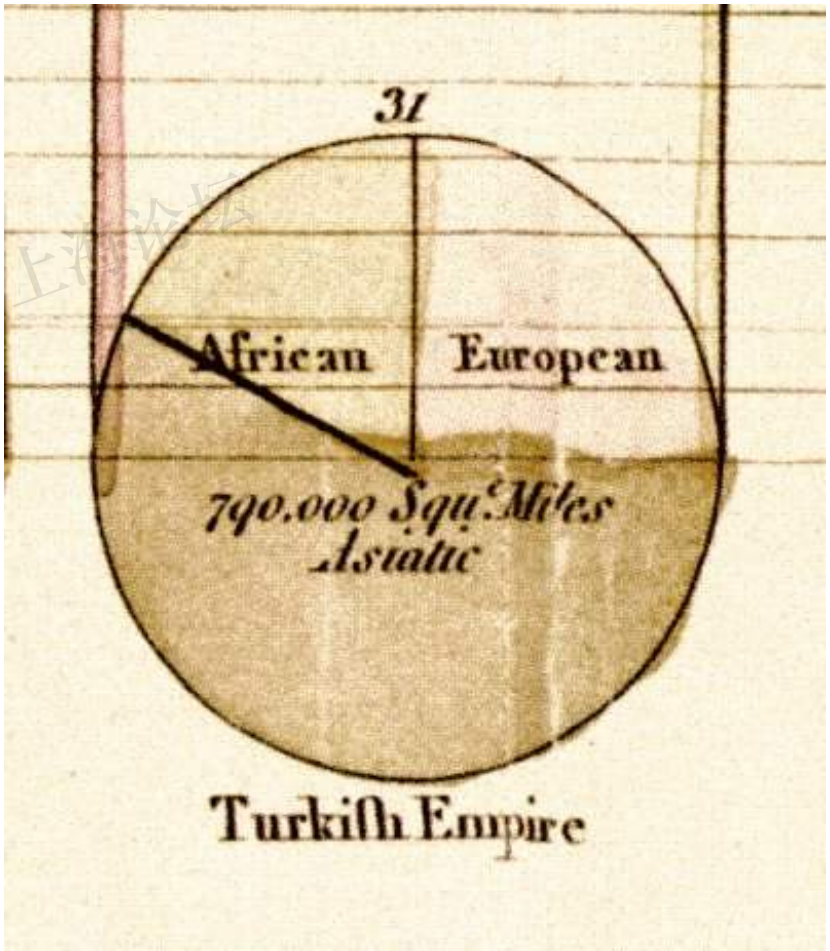
袁晓如 北京大学

2016.4.24

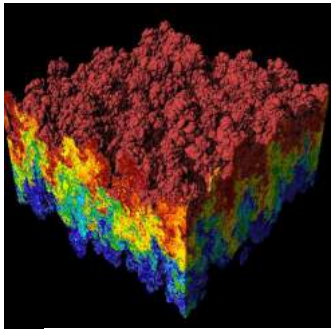
1786: William Playfair







ChinaVis 2016 上海论坛



1987 美国NSF会议

- 1987 report by the National Science Foundation's Advisory Panel on Graphics, Image Processing, and Workstations

Visualization in Scientific Computing

Editors

Bruce H. McCormick
Texas A&M University
College Station, TX

Thomas A. DeFanti
University of Illinois at Chicago
Chicago, IL

Maxine D. Brown
University of Illinois at Chicago
Chicago, IL

This material is based upon work supported by the National Science Foundation under Grant ASC-8712231.

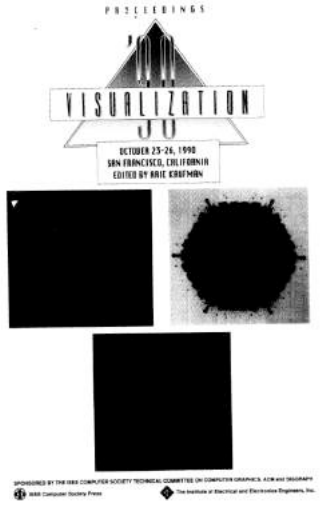
Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Published July,1987.

ChinaVis 2016 上海论坛

"Visualization in scientific computing"

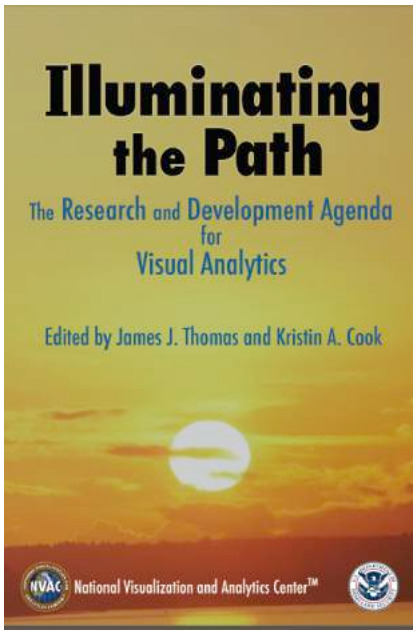
"Visualization in Scientific Computing," B.H. McCormick, T.A. DeFanti, and M.D. Brown, eds., Computer Graphics, **Vol.** 21, **No.** 6, **Nov.** 1987.

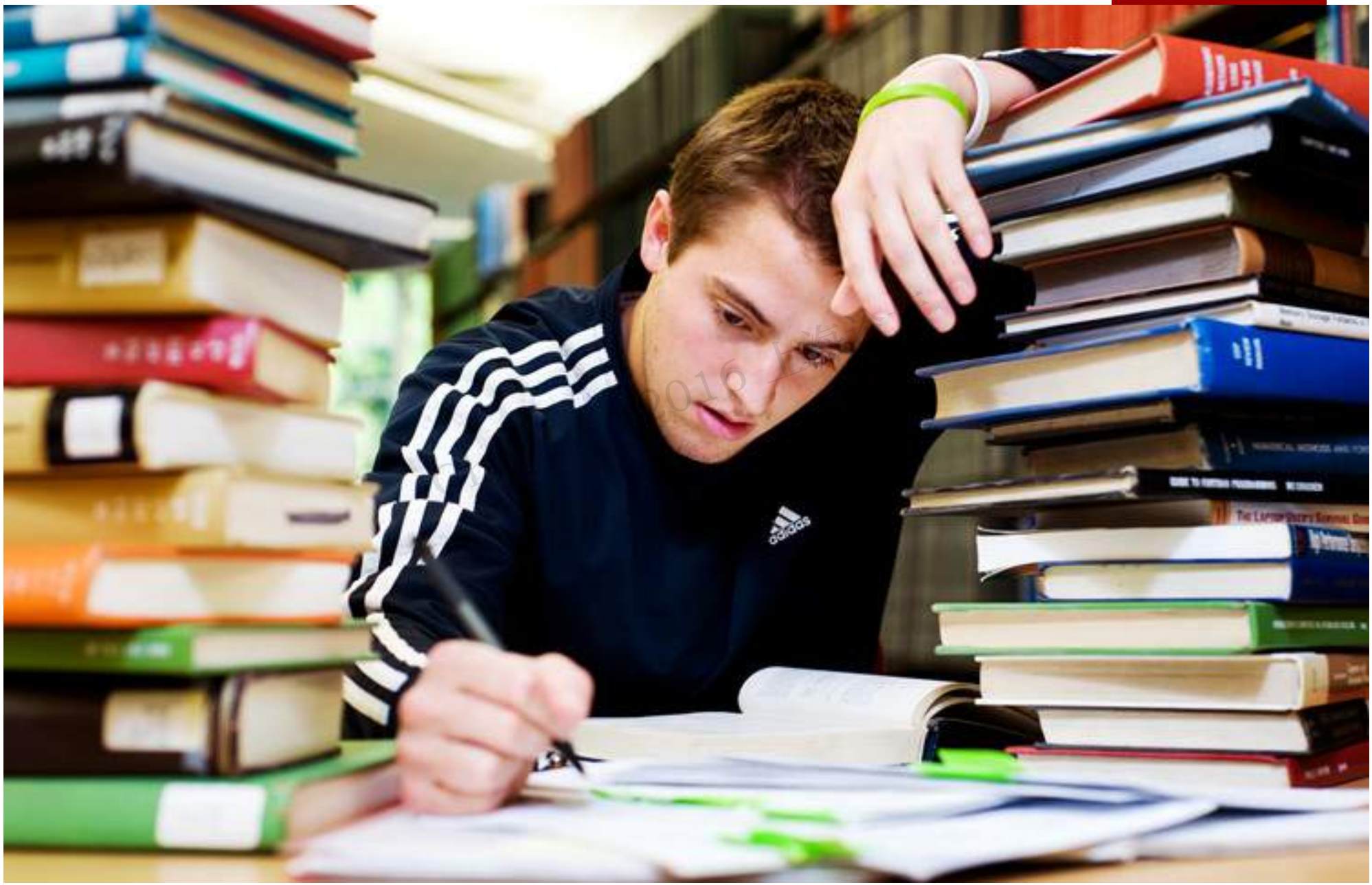


1990 首届IEEE Visualizatio

- 1990 首届IEEE Visualization
  - San Francisco, California October 23-26, 1990
- 1995 IEEE Information Visualization
  - Atlanta Airport Hilton and Towers, Atlanta, GA October 30-31, 1995
- 2006 IEEE VAST
  - Baltimore, MD October 31- November 2, 2006

ChinaVis 2016 上海论坛

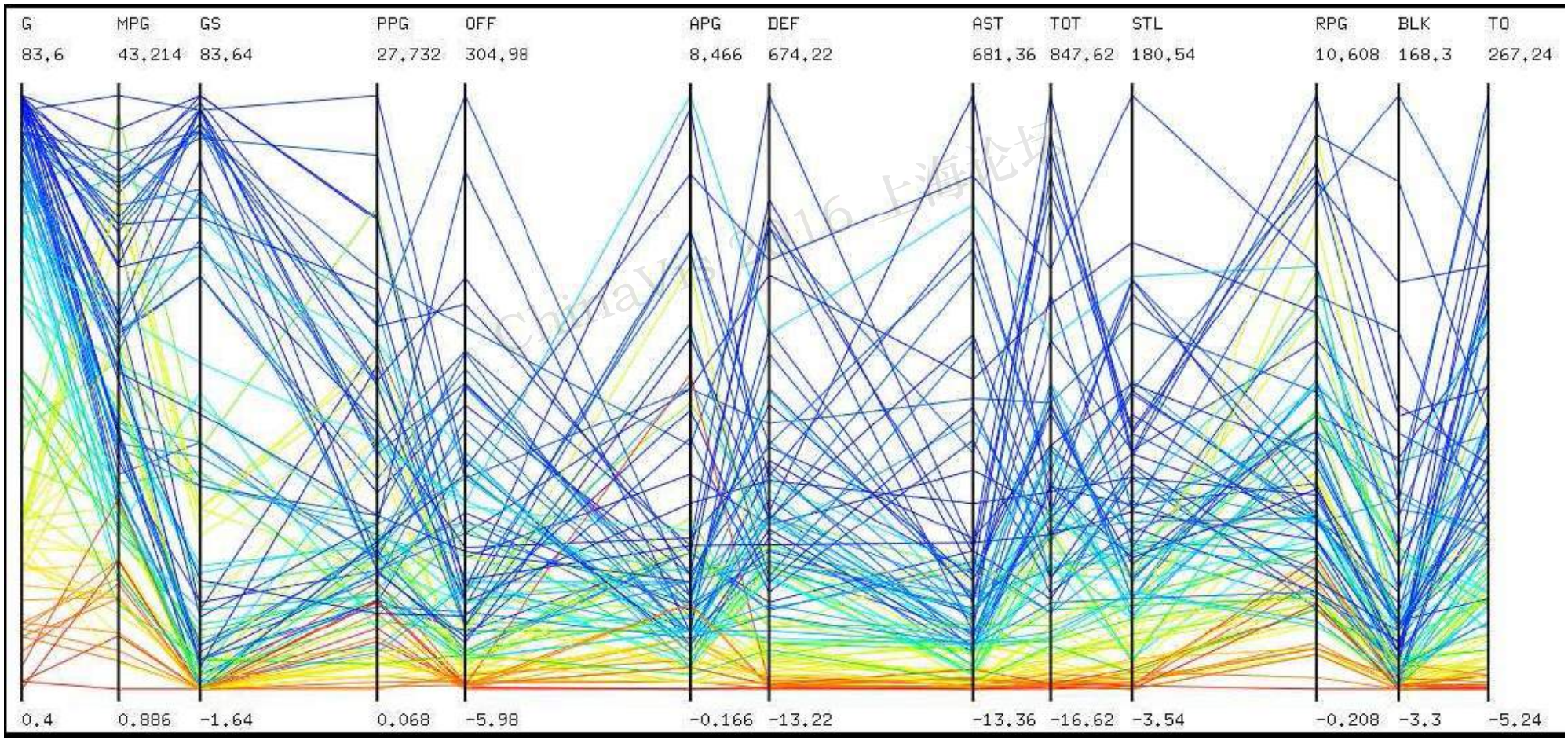




Scattering Points in Parallel Coordinates

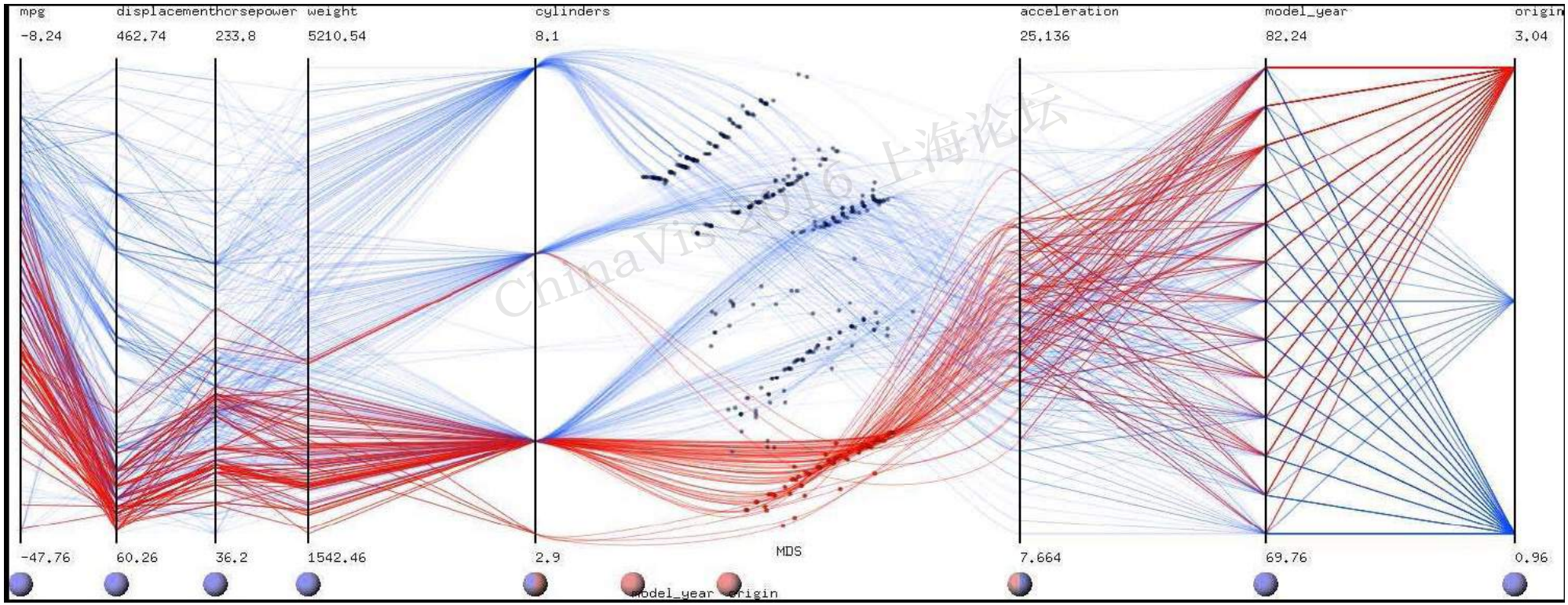
Submitted to IEEE Infovis 2009

Parallel Coordinates



[Yuan et al. TVCG 2009]

Data Exploration with SPPC



ChinaVis 2016 上海论坛

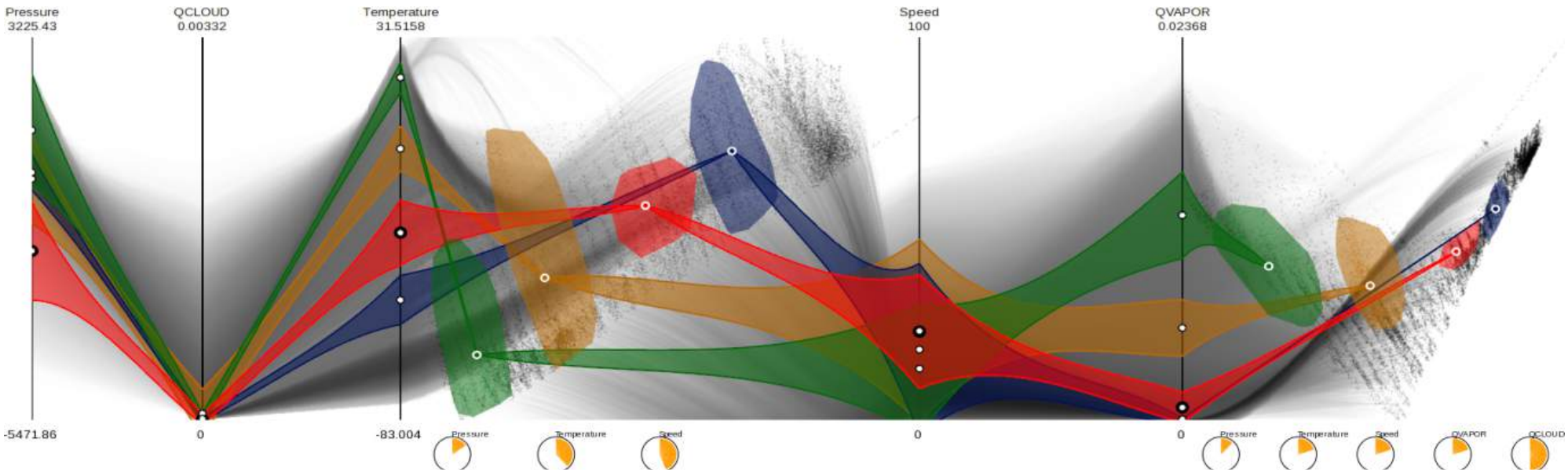
[Yuan et al. TVCG 2009]

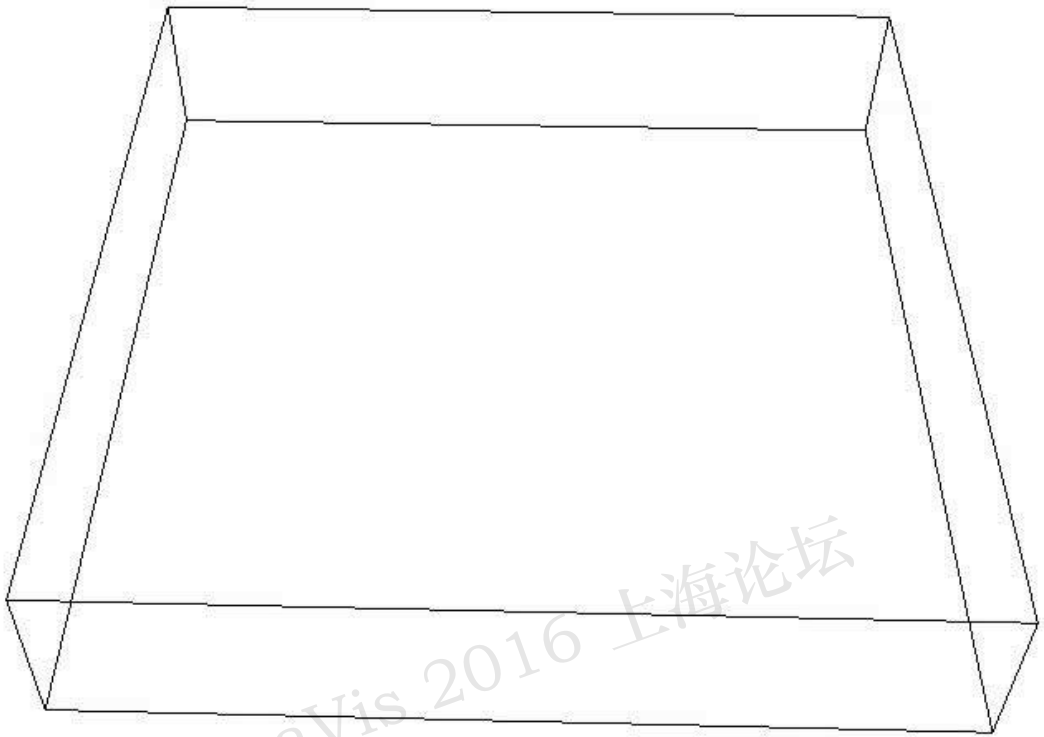
Multivariate Transfer Function Design

- Integrate MDS into PCP to facilitate multivariate TF design.
  - Avoids context jumps between polyline and point regions when exploring data clusters
  - Provides multiple perspective views upon the data, supporting linked queries

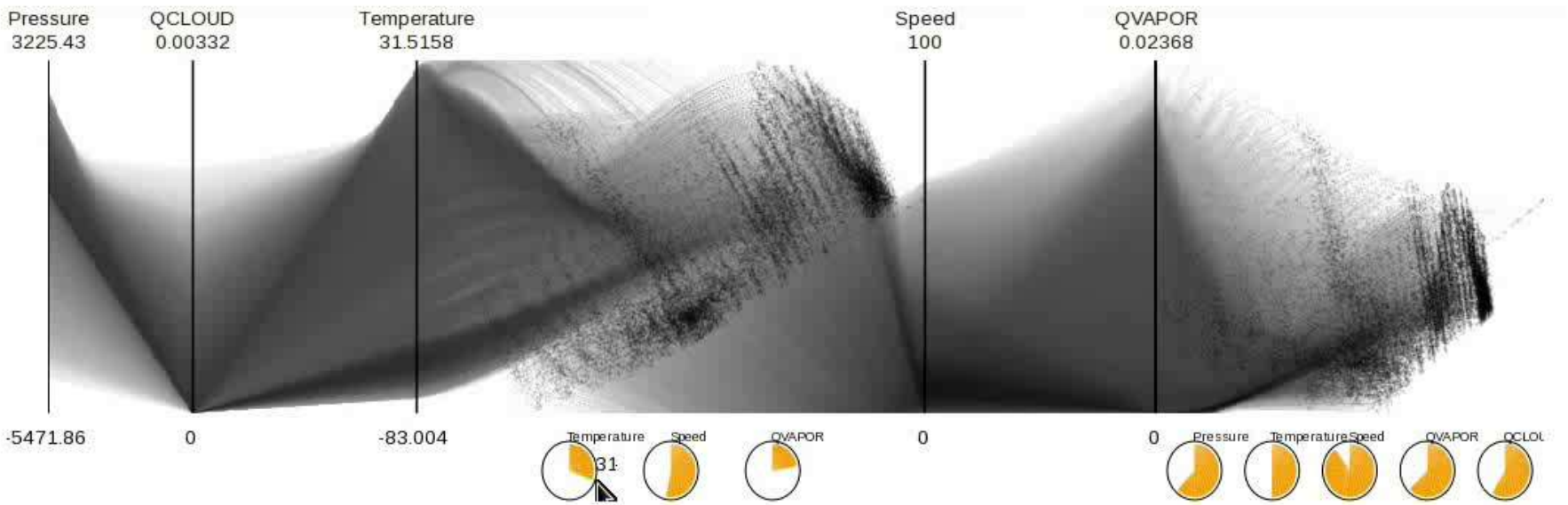
ChinaVis 2016 上海论坛

[Guo, Xiao, and Yuan. PacificVis 2011, TVCG2013]





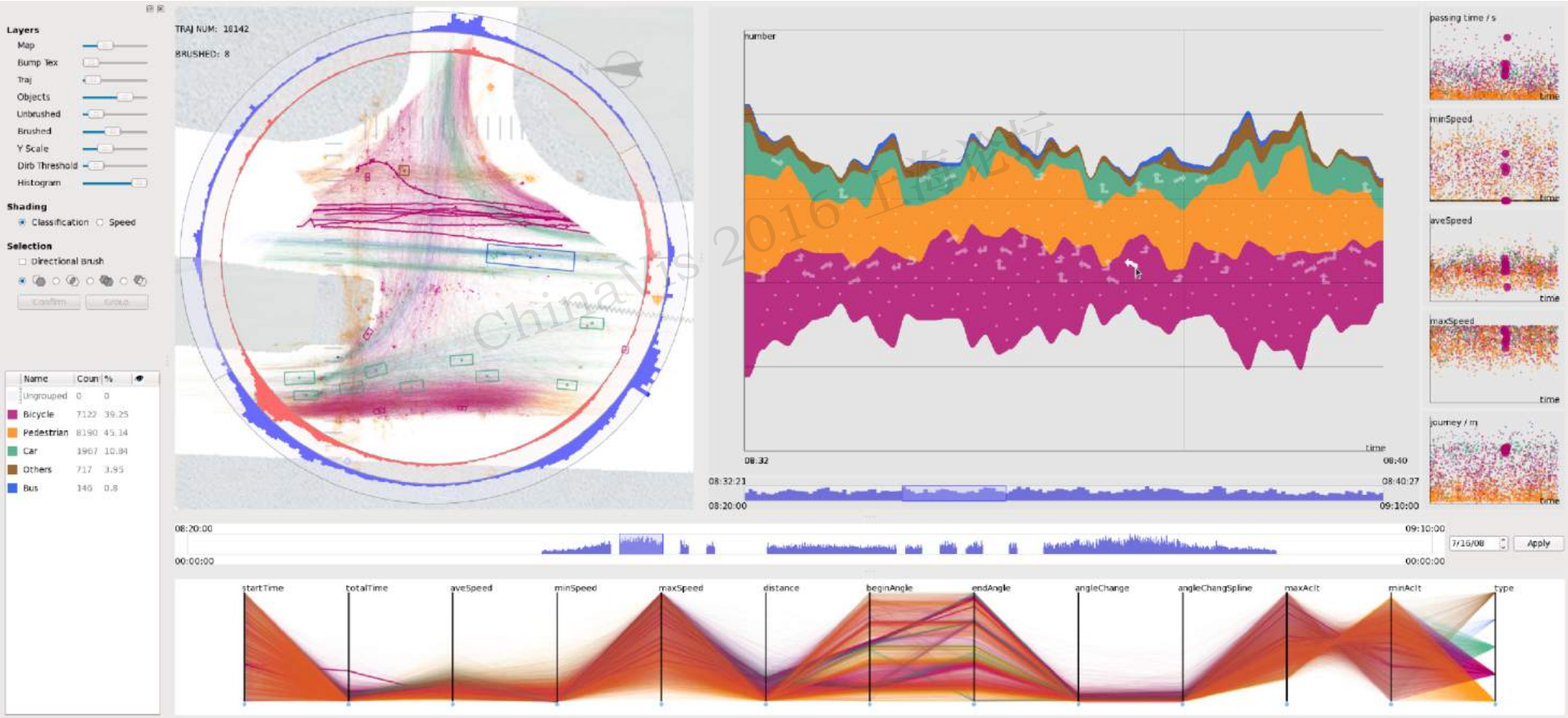
ChinaVis 2016 上海论坛



Scalable Multi-variate Analytics of Seismic and Satellite-based Observational Data

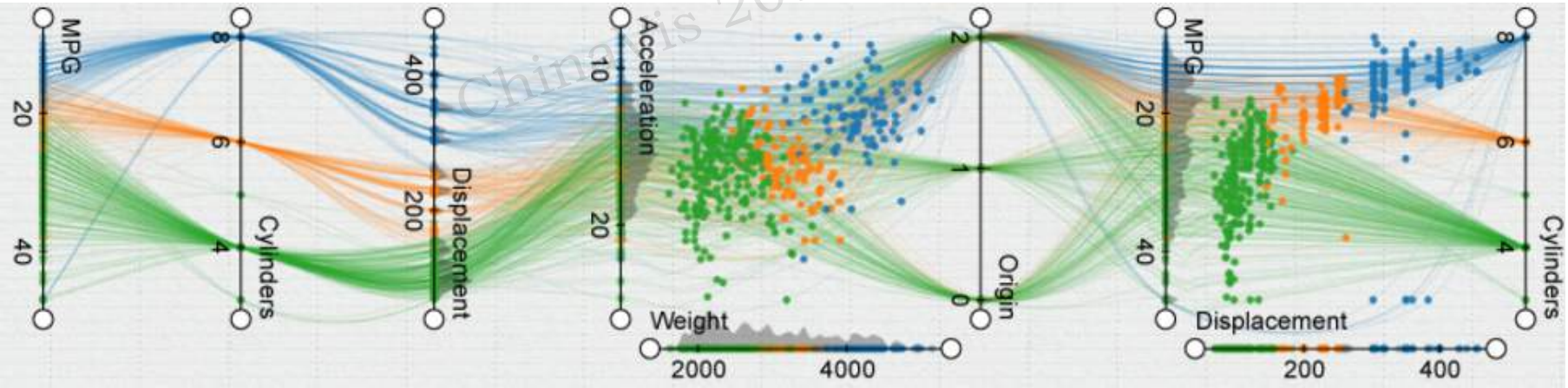
[Yuan et al. TVCG'10]

TripVista – Triple Perspective Visual Traffic Analytics



[Guo et al. PacificVis 2011]

可视化工具



China vis 2016 上海论坛

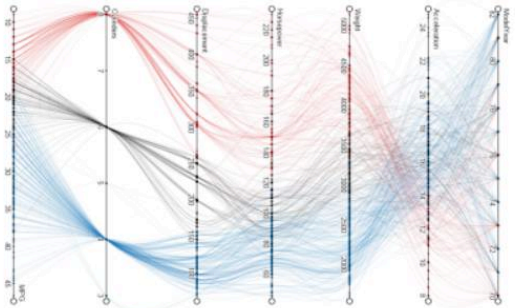
ChinaVis 2016 上海论坛

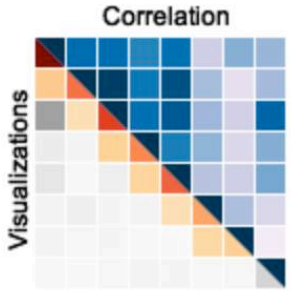
Visualizations of CarData

All Friends Related Trash Previous 1 Next

Visualizations of page 1

myp3uk2005 Feb 26 2016, 13:30



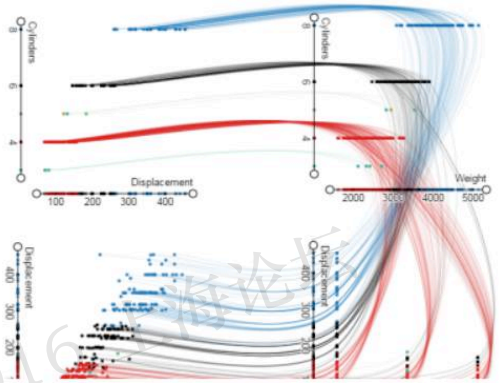


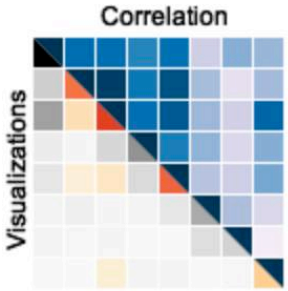
carvisflow42

Load Merge Delete

Comments: 0

myp3uk2005 Feb 26 2016, 13:26





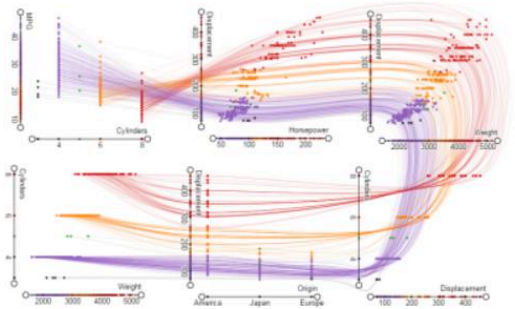
carvis42

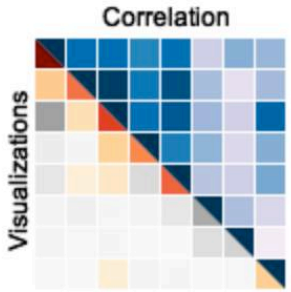
Load Merge Delete

Comments: 0

ChinaVis 2016 上海论坛

myp3uk2005 Feb 26 2016, 12:31

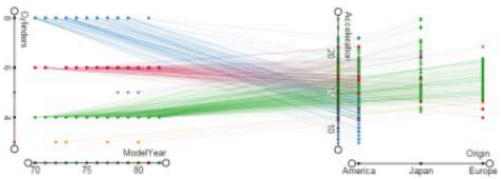


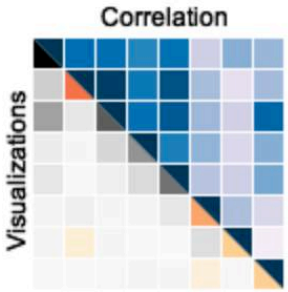


Car Cylinders 4939

Load Merge Delete

myp3uk2005 Feb 25 2016, 19:51

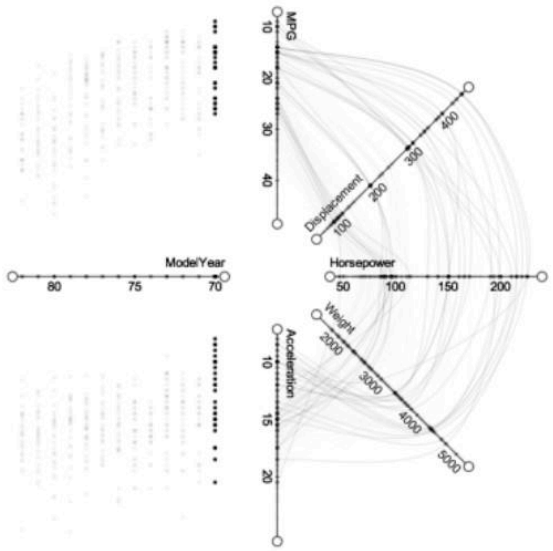




testing_102

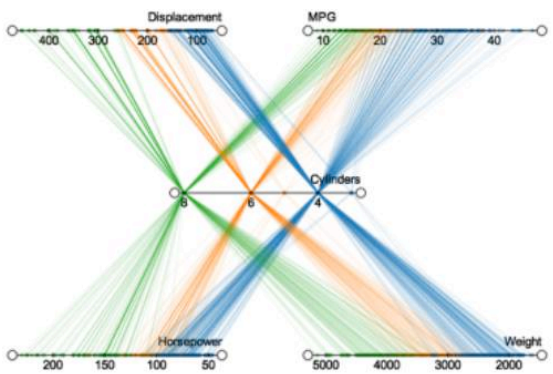
Load Merge Delete

Comments: 0



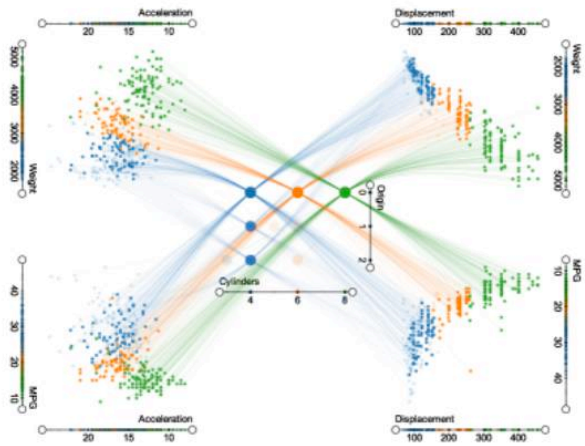
windwish: Star coordinates, and filtering.

Aug 27 2012, 15:51 Load



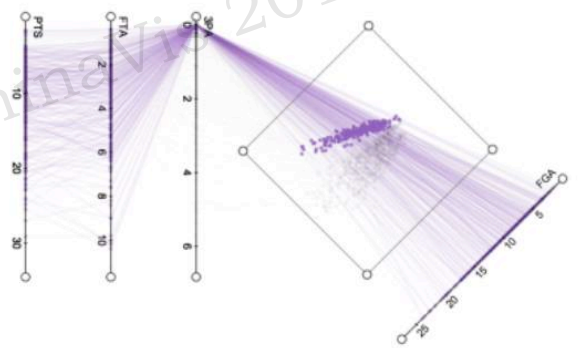
windwish: 一对多的连接

Aug 26 2012, 21:34 Load



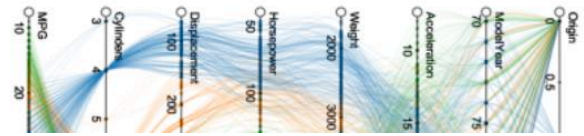
windwish: Multilink for scatterplots.

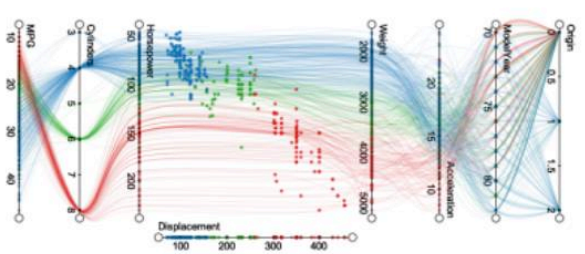
Aug 27 2012, 15:44 Load



windwish: NBA 中不投三分的球员

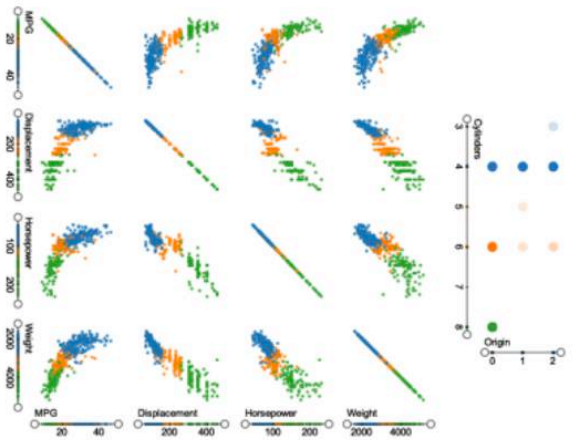
Aug 26 2012, 21:44 Load





windwish: Scatterplot in parallel coordinates.

Aug 27 2012, 11:15 Load



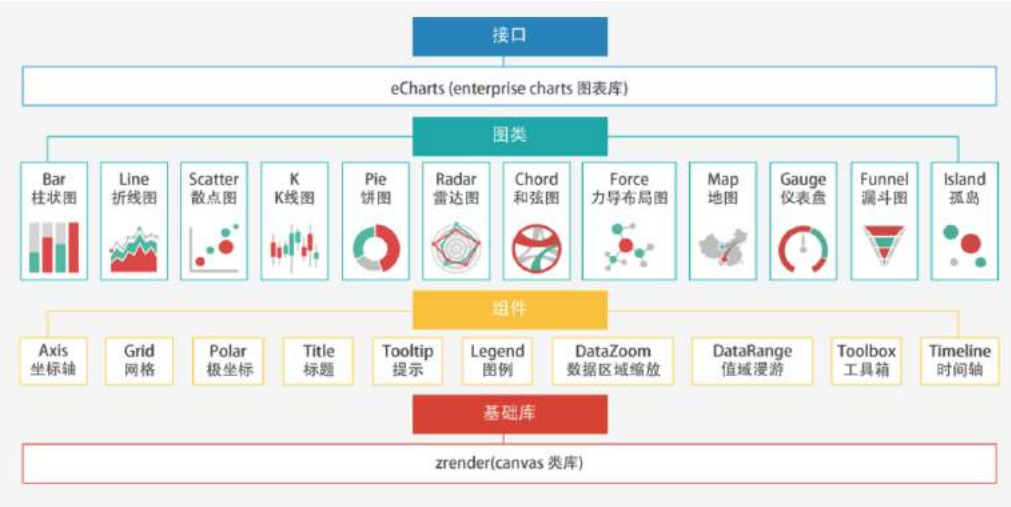
windwish: 经典可视化形式——散点图矩阵 Scatterplot Matrix

Aug 26 2012, 21:29 Load

ChinaVis 2016 上海论坛

可视化通用性讨论

- 可视化 – 解决数据分析的需求
- 通用数据探索 v.s. 特定领域分析
  - Echarts提供了较多且简单易用的视图，容易上手
  - 但对于专业问题可能缺乏交互与表达
- 更深层次的需求：所见即所得与可视化定制



ChinaVis 2016 上海论坛

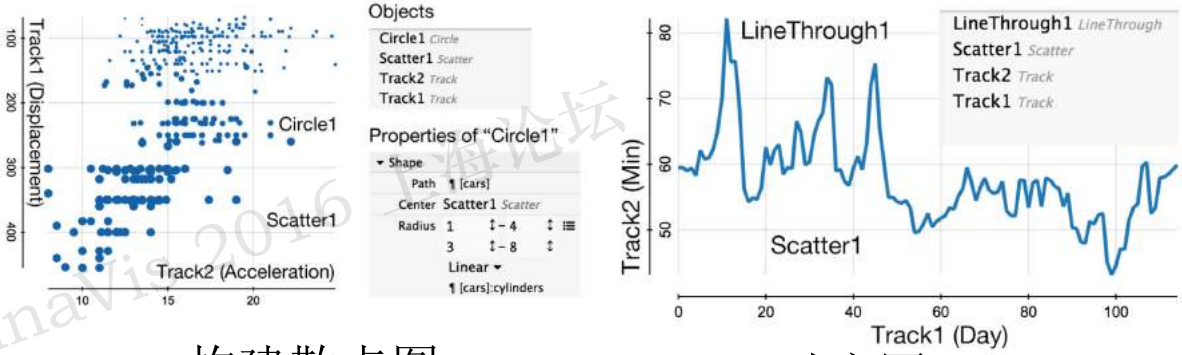
可视化通用性讨论

- 可视化 – 解决数据分析的需求
- 通用数据探索 v.s. 特定领域分析
  - Echarts提供了较多且简单易用的视图，容易上手
  - 但对于专业问题可能缺乏交互与表达
- 更深层次的需求：所见即所得与可视化定制

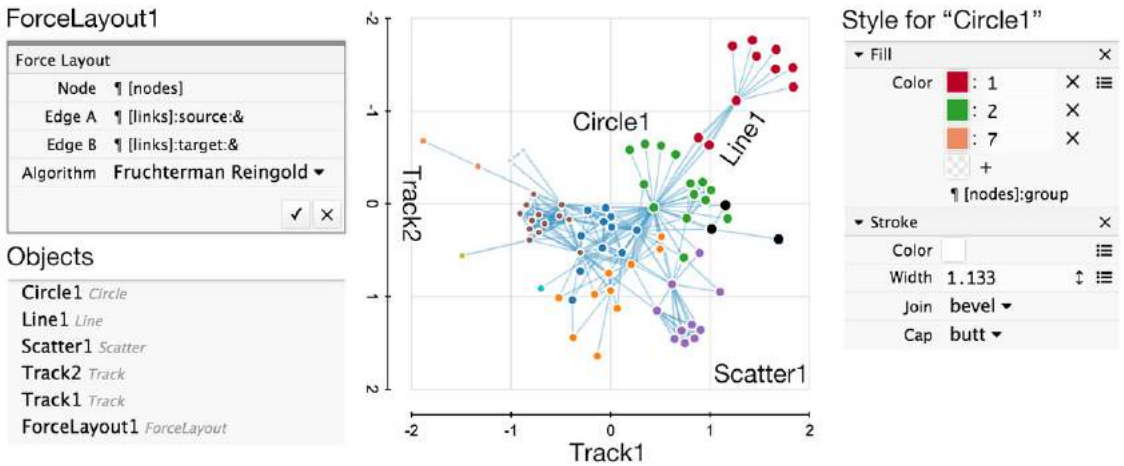
ChinaVis 2016 上海论坛

可视化通用性讨论

- 用数据创造属于自己的可视化 iVisDesigner
  - 支持多种基础视图的创建
  - 不仅仅是将数据绘制出来，而是对数据进行解构、探索与关联
  - 颜色、形状、大小的可定制映射，表达数据分析的需求



构建散点图　　时序图



点边图

ChinaVis 2016 上海论坛

可视化通用性讨论

自定义的空气污染可视化 – 无需编程、动态链接

北京大学微博可视分析系列

北京大学PKUVIS微博可视分析系列，利用新颖的可视分析技术为您提供强大的视觉盛宴与交互享受的同时，帮助您更深地挖掘微博事件、关键词、用户等等的关系。下面是我们目前已有的主要工具：

WeiboEvents

这是我们开发的挖掘微博里故事的在线工具。它利用新颖的可视化方式呈现微博中的事件，能够快速浏览和分析微博事件中的人和事，把握微博传播的脉络，挖掘深层的人人、人事关系。

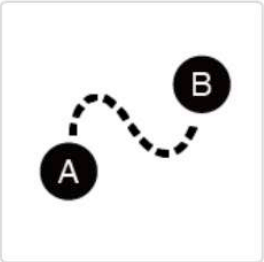
点击这里查看更多介绍

WeiboFootprint

这是我们开发的针对带有地理信息的微博，来分析用户的行为时空特征的可视化系统。它展示了用户的移动轨迹，并提供多种交互方式，让用户对移动轨迹的特征进行分析，把握时空变化的关系。

点击这里查看更多介绍。

ChinaVis 2016 上海论坛



WeiboGeo

这是我们开发的针对大量带有地理信息的微博，来分析群体用户行为时空特征的可视分析系统。用户可以选择不同层次的区域，筛选符合条件的移动轨迹，展现区域间群体移动特征，例如不同城市的旅游线路、时间花费等。

点击这里查看更多介绍。

ThemeMap

Thememap是我们推出的最新微博可视分析工具，支持用户创建自己感兴趣的主题地图，通过广大用户的标注，集合众人的力量快速生成一张完整的主题地图，快速掌握主题的区域热度分布和变化发展。

Polaris

- Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases
- Chris Stolte and Pat Hanrahan
- *Proceedings of the Sixth IEEE Symposium on Information Visualization*, October 2000.

ChinaVis 2016 上海论坛

Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases

Chris Stolte and Pat Hanrahan
Computer Science Department
Stanford University

Abstract

In the last several years, large multi-dimensional databases have become common in a variety of applications such as data warehousing and scientific computing. Analysis and exploration tasks place significant demands on the interfaces to these databases. Because of the size of the data sets, dense graphical representations are more effective for exploration than spreadsheets and charts. Furthermore, because of the exploratory nature of the analysis, it must be possible for the analysts to change visualizations rapidly as they pursue a cycle involving first hypothesis and then experimentation.

In this paper we present Polaris, an interface for exploring large multi-dimensional databases that extends the well-known Pivot Table interface. The novel features of Polaris include an interface for constructing visual specifications of table-based graphical displays and the ability to generate a precise set of relational queries from the visual specifications. The visual specifications can be rapidly and incrementally developed, giving the analyst visual feedback as they construct complex queries and visualizations.

1. Introduction

In the last several years, large databases have become common in a variety of applications. Corporations are creating large data warehouses of historical data on key aspects of their operations. International research projects, such as the Human Genome Project [15] and Digital Sky Survey [23], are generating massive databases of scientific data.

A major challenge with these databases is to extract meaning from the data they contain: to discover structure, find patterns, and derive causal relationships. The analysis and exploration necessary to uncover this hidden information places significant demands on the human-computer interfaces to these databases. The exploratory analysis process is one of hypothesis, experiment and discovery. The path of exploration is unpredictable, and the analysts or scientists need to be able to rapidly change both what data they are viewing and how they are viewing that data.

The current trend is to treat multi-dimensional databases as n-dimensional data cubes [13]. Each dimension in these data cubes corresponds to one dimension in the relational schema. Perhaps the most popular interface to multi-dimensional databases is the Pivot Table [11]. Pivot Tables allow the data cube to be rotated, or pivoted, so that different dimensions of the dataset may be encoded as rows or columns of the table. The remaining dimensions are aggregated and displayed as numbers in the cells of the table. Cross-tabulations and summaries are then added to the resulting table of numbers. Finally, graphs may be generated from the resulting tables. Visual Insights recently released a new interface for visually exploring projections of data cubes using linked views of bar charts, scatterplots, and parallel coordinate displays [10].

In this paper we present Polaris, an interface for the exploration of multi-dimensional databases that extends the Pivot Table interface to directly generate a rich, expressive set of graphical displays. Polaris builds tables using an algebraic formalism involving the fields of the database. Each table consists of layers and panes, and each pane may be a different graphic. The use of tables to organize multiple graphs on a display is a technique often used by statisticians in their analysis of data [3][7][28].

The Polaris interface is simple and expressive because it is built upon a formalism for constructing graphs and building data transformations. We interpret the state of the interface as a visual specification of the analysis task and automatically compile it into data and graphical transformations. This allows us to combine statistical analysis and visualization. Furthermore, all intermediate specifications that can be created in the visual language are valid and can be interpreted to create visualizations. Therefore, analysts can incrementally construct complex queries, receiving visual feedback as they assemble and alter the specifications.

2. Related Work

The related work to Polaris can be divided into three categories: formal graphical specifications, table-based data displays and database exploration tools.

2.1. Formal Graphical Specifications

Bertin's *Semiology of Graphics* [4] is one of the earliest attempts at formalizing graphing techniques. Bertin developed a vocabulary for describing data and the techniques for encoding data in a graphic. One of his important contributions is the identification of the retinal variables (position, color, size, etc.) in which data can be encoded. Cleveland [7][8] used theoretical and experimental results to determine how well people can use these different retinal properties to compare quantitative variations.

Mackinlay's APT system [18] is one of the first applications of formal graphical specifications to computer-generated displays. APT uses a set of graphical languages and composition rules to automatically generate 2D displays of relational data. The Sage system [21] extends the concepts of APT, providing a richer set of data characterizations and generating a wider range of displays.

Livny et al. [17] describe a visualization model that provides a foundation for database-style processing of visual queries. Within this model, the relational queries and graphical mappings necessary to generate visualizations are defined by a set of relational operators. The Rivet visualization environment [6] applies similar concepts to provide a flexible database visualization tool.

Database Schema:
The user drags fields from the database schema to shelves to define the visual specification.

Layer Tabs:
Each layer has its own tab; different transformations and mappings can be specified for each layer.

Axis Shelves:
The fields placed here determine the structure of the table and the types of graphs in each table pane.

Context Menu:
The context menu provides access to the data transformation and interaction capabilities of Polaris such as sorting, filtering, and aggregation.

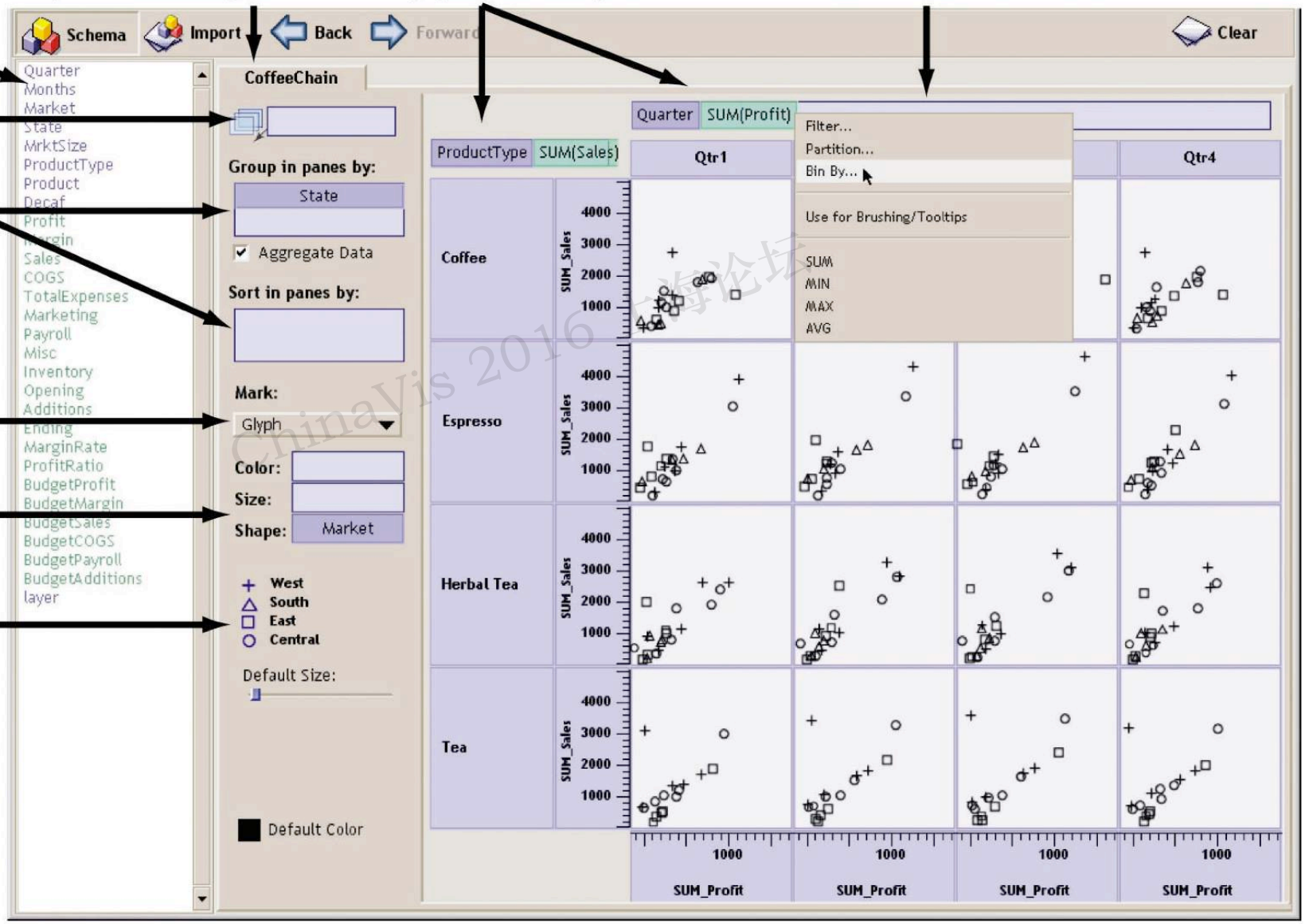
Layer Shelf:
The fields placed here determine how records are partitioned into layers.

Grouping and Sorting Shelves:
The fields placed here determine how records are grouped and sorted within the table panes.

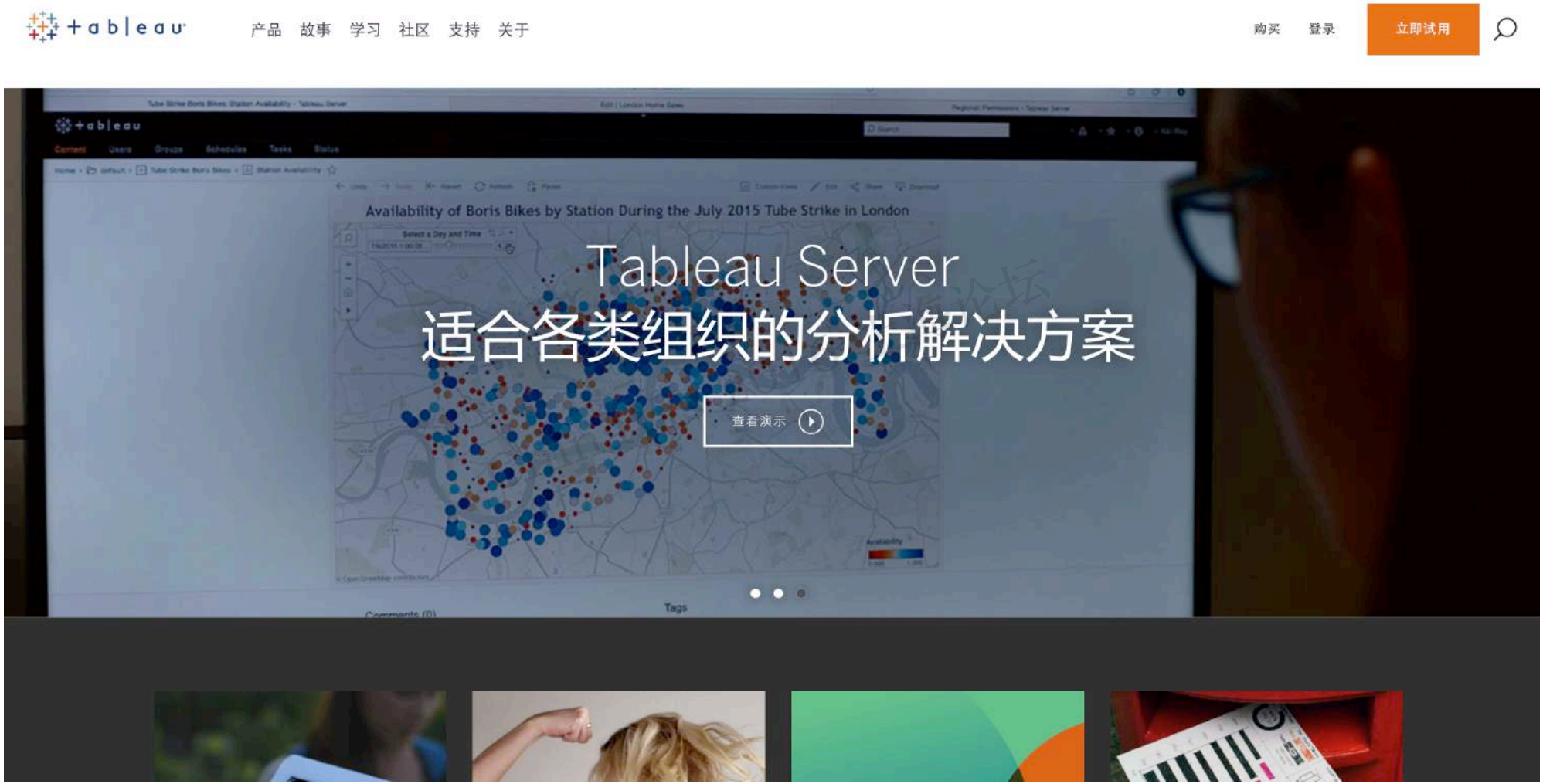
Mark Pulldown:
Relations in each pane are mapped to marks of the selected type.

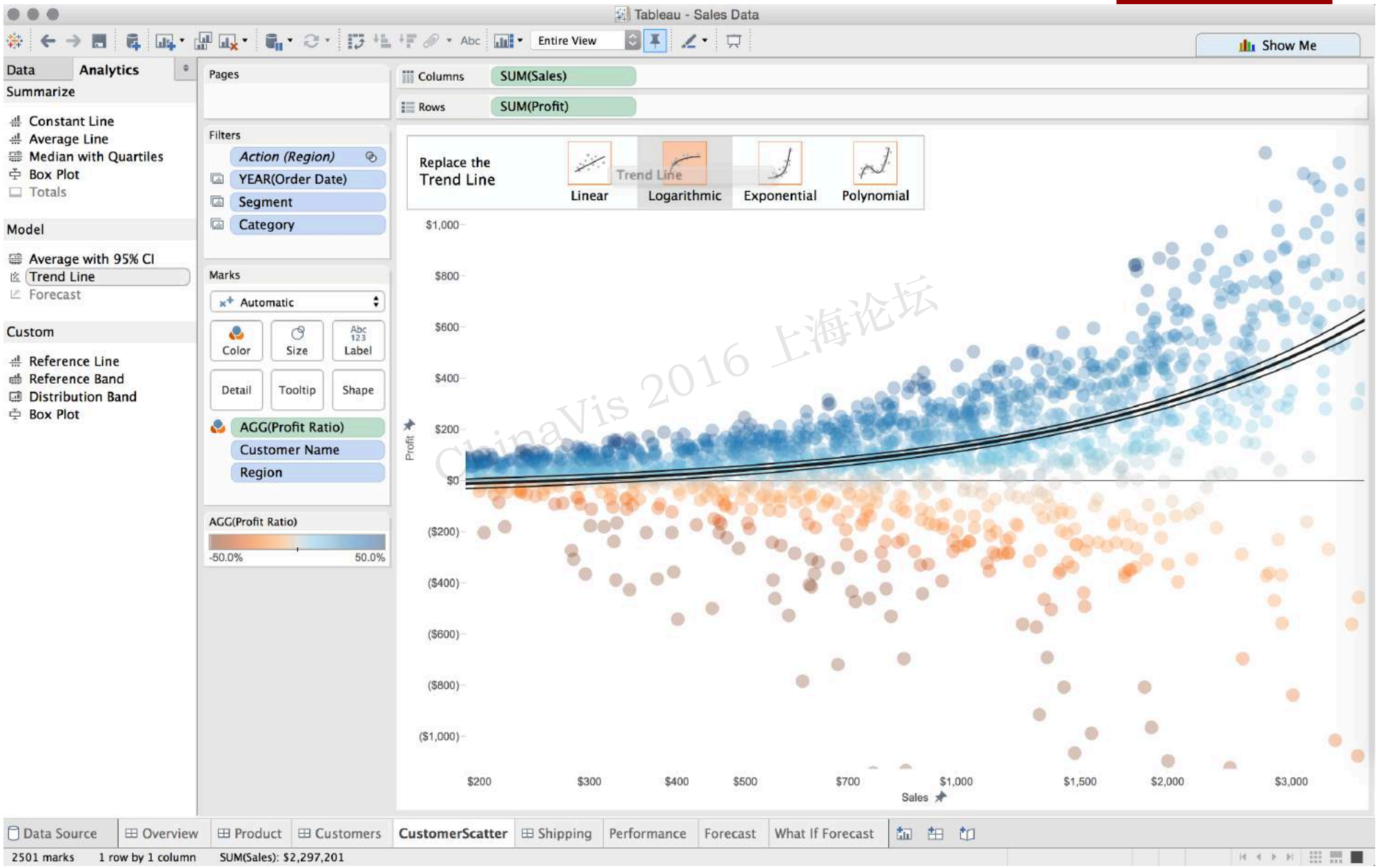
Retinal Property Shelves:
The fields placed here determine how data is encoded in the retinal properties of the marks.

Legends:
Legends enable the user to see and modify the mappings from data to retinal properties.



ChinaVis 2016 上海论坛


+ableau
产品 故事 学习 社区 支持 关于
购买 登录 立即试用
Availability of Boris Bikes by Station During the July 2015 Tube Strike in London
Tableau Server
适合各类组织的分析解决方案
查看演示



ChinaVis 2016 上海论坛

王國維—《人間詞話》

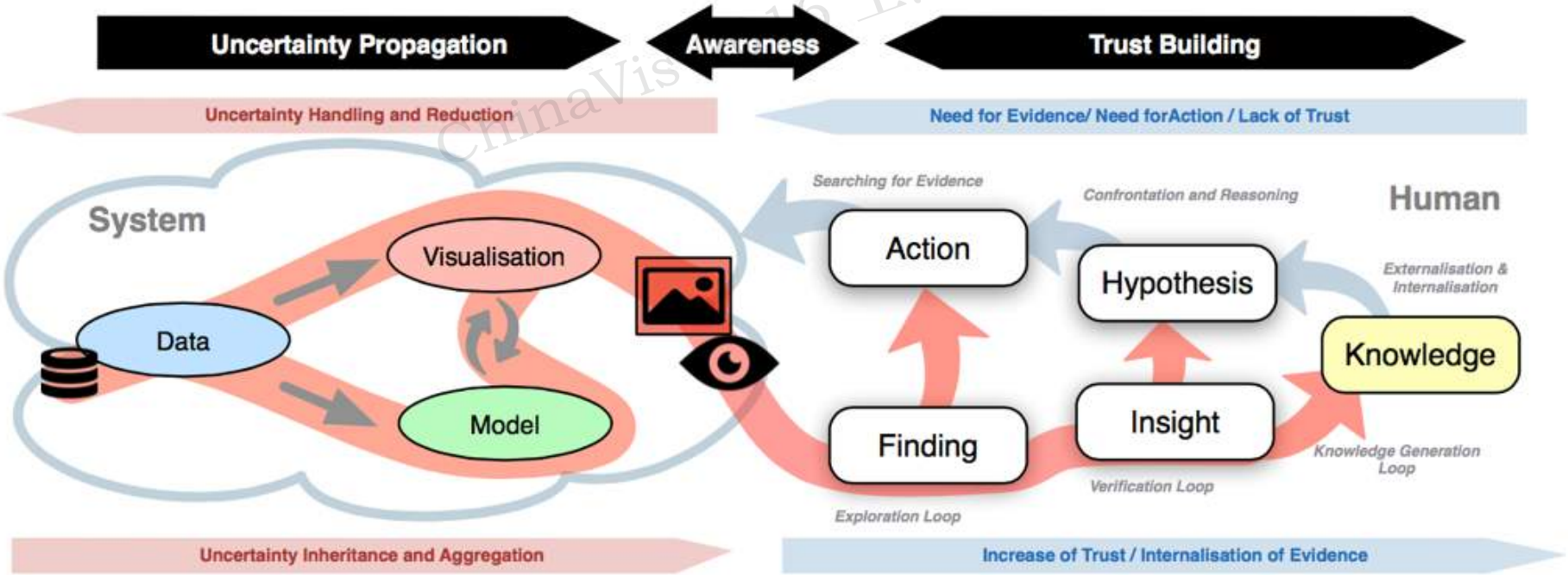
- 「昨夜西風凋碧樹。獨上高樓，望盡天涯路」，此第一境也；
- 「衣帶漸寬終不悔，為伊消得人憔悴」此第二境也；
- 「眾裏尋他千百度，回頭驀見，那人正在燈火闌珊處」，此第三境也。

ChinaVis 2016 上海论坛

可视化研究进展

- 从单一的可视化方法的研究，转变为针对具体特征数据的可视分析系统研究
- 结合更多的异构数据，针对高维、时空、网络等方面挑战，构建综合的可视分析系统

ChinaVis 2016 上海论坛



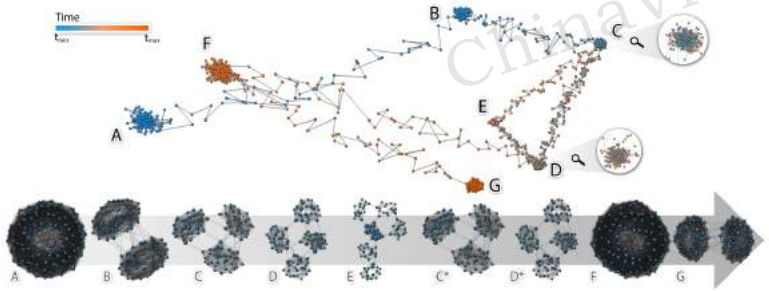
可视分析模型，包含数据、模型、可视化的关系，以及背后蕴含的事态感知与知识构建

Sacha, D.; Senaratne, H.; Kwon, B.C.; Ellis, G.; Keim, D.A., "The Role of Uncertainty, Awareness, and Trust in Visual Analytics," in *IEEE Transactions on Visualization and Computer Graphics*, (*VAST2015*),

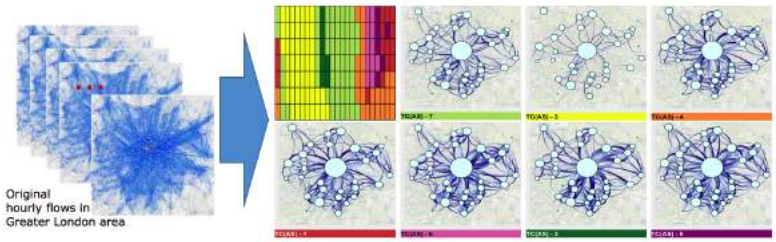
可视化前沿技术

- 针对海量、时变的科学可视化交互分析技术
- 结合先进数据挖掘、机器学习与可视化技术的面向领域的可视分析系统
- 针对异构时空数据、动态网络结构的可视分析

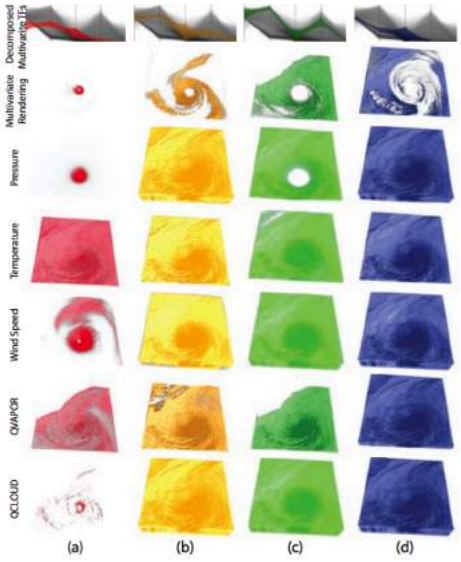
ChinaVis 2016 上海论坛



基于动态网络的聚类投影的可视分析



时空轨迹可视分析



科学数据可视化

■

“可视化推动大数据平民化”
位列CCF 2016年大数据发展趋势首位

ChinaVis 2016 上海论坛

Acknowledgement

- Students
- Funds
  - NSFC
  - 863
  - PKU
  - Beijing NSF
  - Company funds from Qihu 360, IBM, Nokia, Microsoft
- Collaborators

http://vis.pku.edu.cn/wiki
Xiaoru.yuan@pku.edu.cn

ChinaVis 2016 上海论坛