Data Visualization and Data Mining

Jürgen Symanzik

Utah State University, Logan, UT

***e-mail: symanzik@sunfs.math.usu.edu**

WWW: http://www.math.usu.edu/~symanzik

Contents

- Data Visualization / Data Mining
  - Software & Tools
  - Applications
- Immersive Projection Technology
  - CAVE & C2
  - The MiniCAVE

Visual Data Mining

- Working Definition:
  - Find structure (cluster, unusual observations) in large and not necessarily homogeneous data sets using graphical methods and user interaction
  - Goal or expected outcome of exploration usually unknown in advance

The Software: XGobi

Swayne, Cook and Buja

- Interactive environment for exploring multivariate data
  - Linked views allow ``linked brushing''
  - Univariate, Bivariate and Multivariate views of the data
  - Grand tour
  - Wide variety of methods
  - Open source
  - Free

- Caveats
  - Only on UNIX and Linux platforms

The Software: ExplorN

Carr, Wegman, Luo

- Interactive environment for exploring multivariate data (similar to XGobi)
  - * Advanced Parallel Coordinates Displays
  - * 3D Surfaces
  - * Stereoscopic Displays

- Caveats
  - * Only on SGI platforms
  - * No interface

The Software: ArcView

ESRI™

- Desktop GIS with wide range of viewing and data manipulation functions
  - * Editing features
  - * Query operations
  - * Map display
  - * Interactive interface
  - * High level internal scripting language
- ArcView has a wide user base

- Caveats
  - * Poor statistical display facilities

The Software: XploRe

Härdle, Klinke, et al.

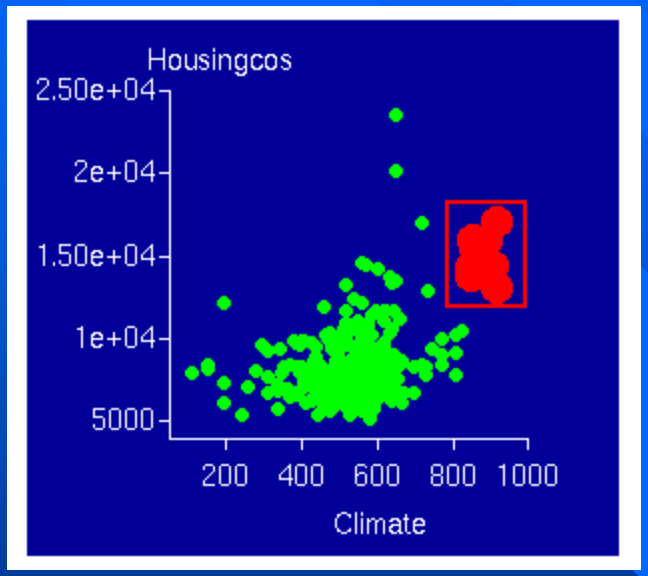
- Statistical Computing Environment
  - Built in set of statistical methods
  - High level programming language
  - Highly Interactive Graphics
  - Built in editors for data and programs
  - High level help system

- Caveats
  - Statistical views not as developed as in XGobi

The ArcView/XGobi/XploRe Link

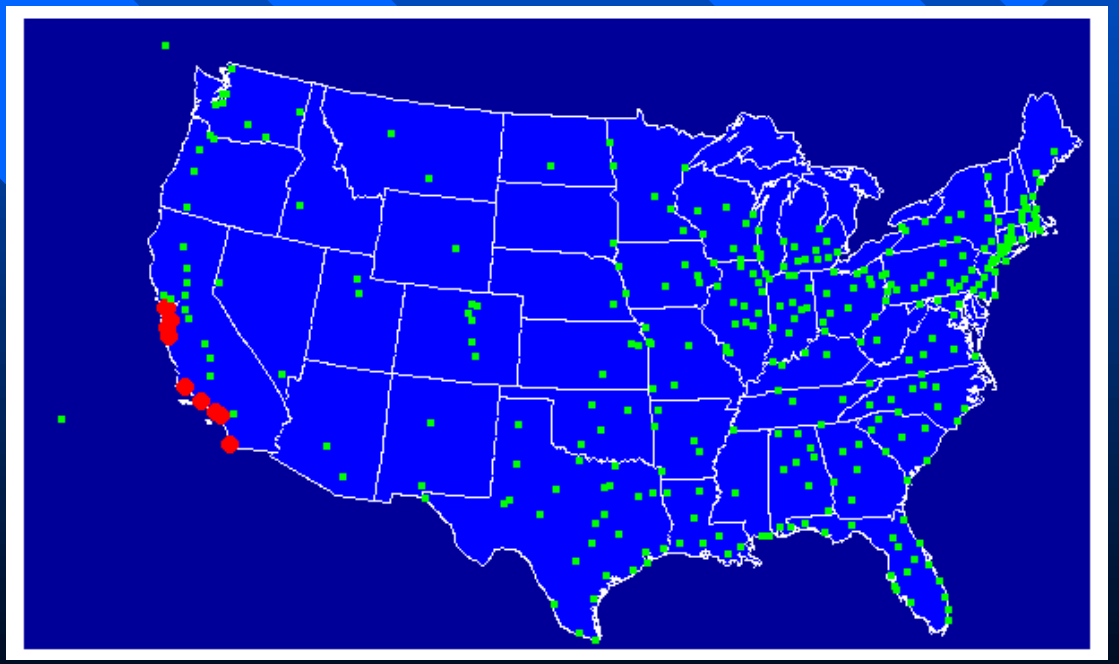
- Links these 3 packages using RPCs
- Remote Procedure Calls (RPCs):
  - Process on the local system (client) invokes a procedure on a remote system (server)
  - Request = client's desire to execute a particular remote procedure
  - Response = result produced by the remote procedure
- ArcView, XGobi, XploRe: server & client

Tools: Linked Brushing 1

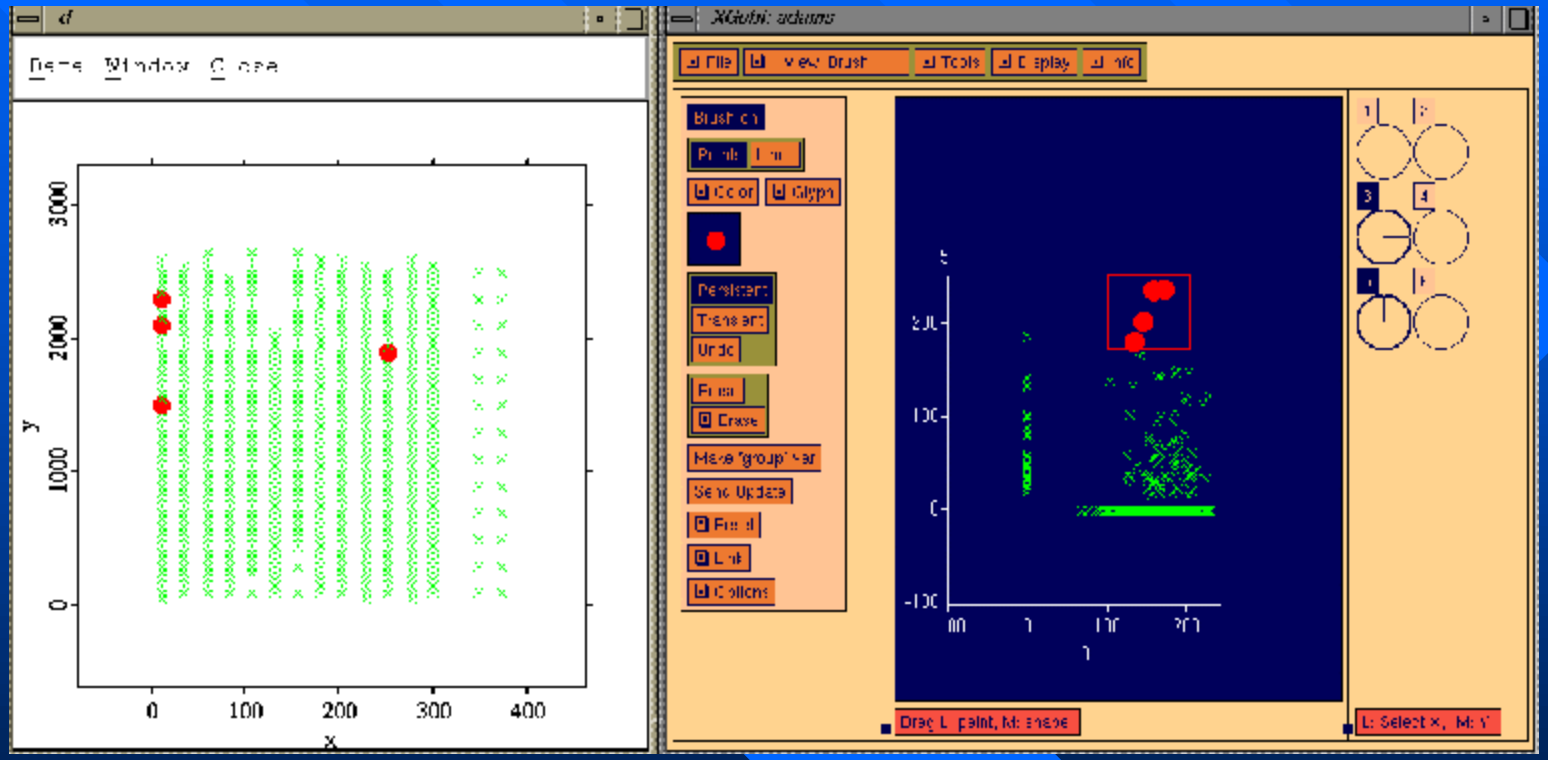


XGobi

ArcView



Tools: Linked Brushing 2

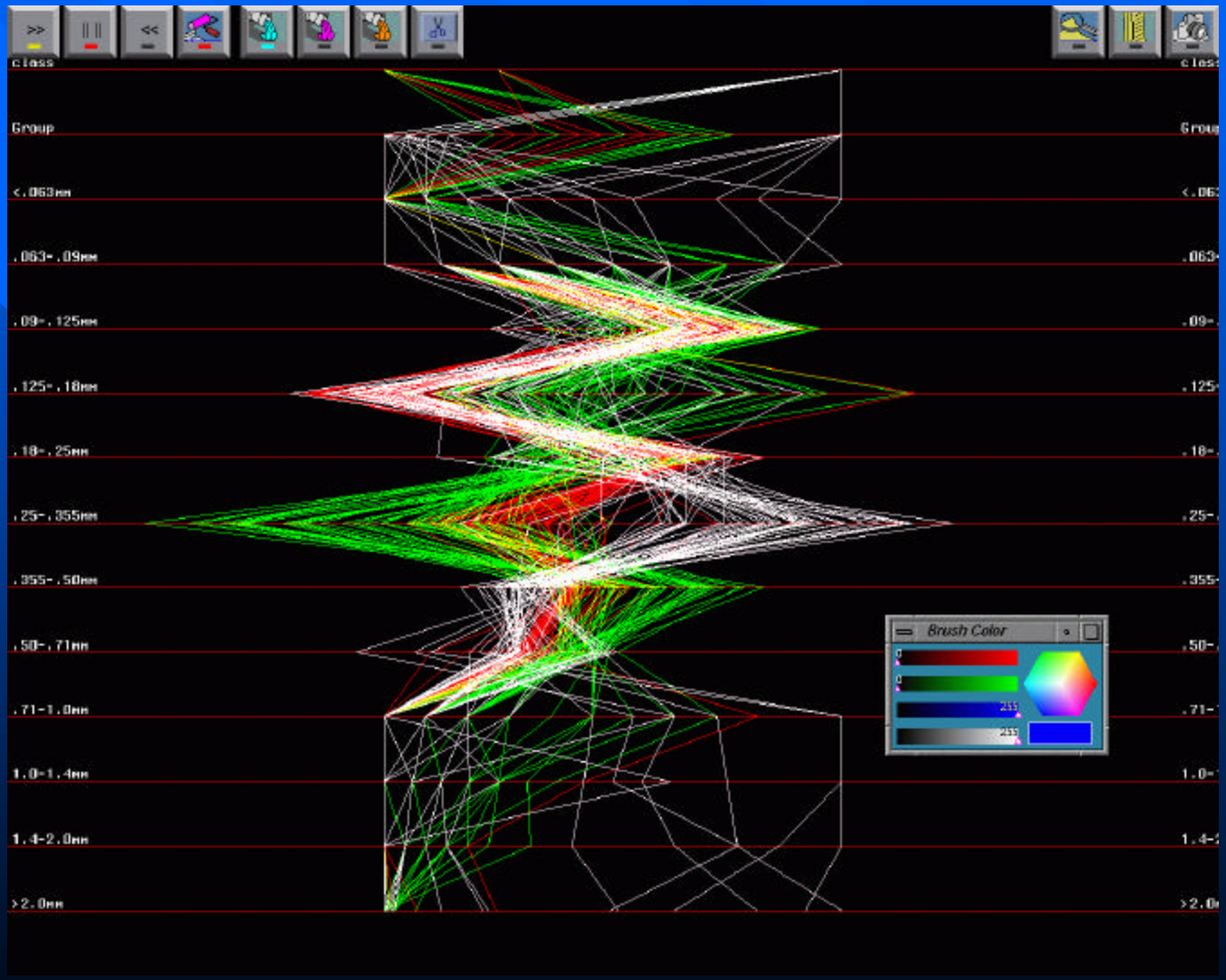


XploRe

XGobi

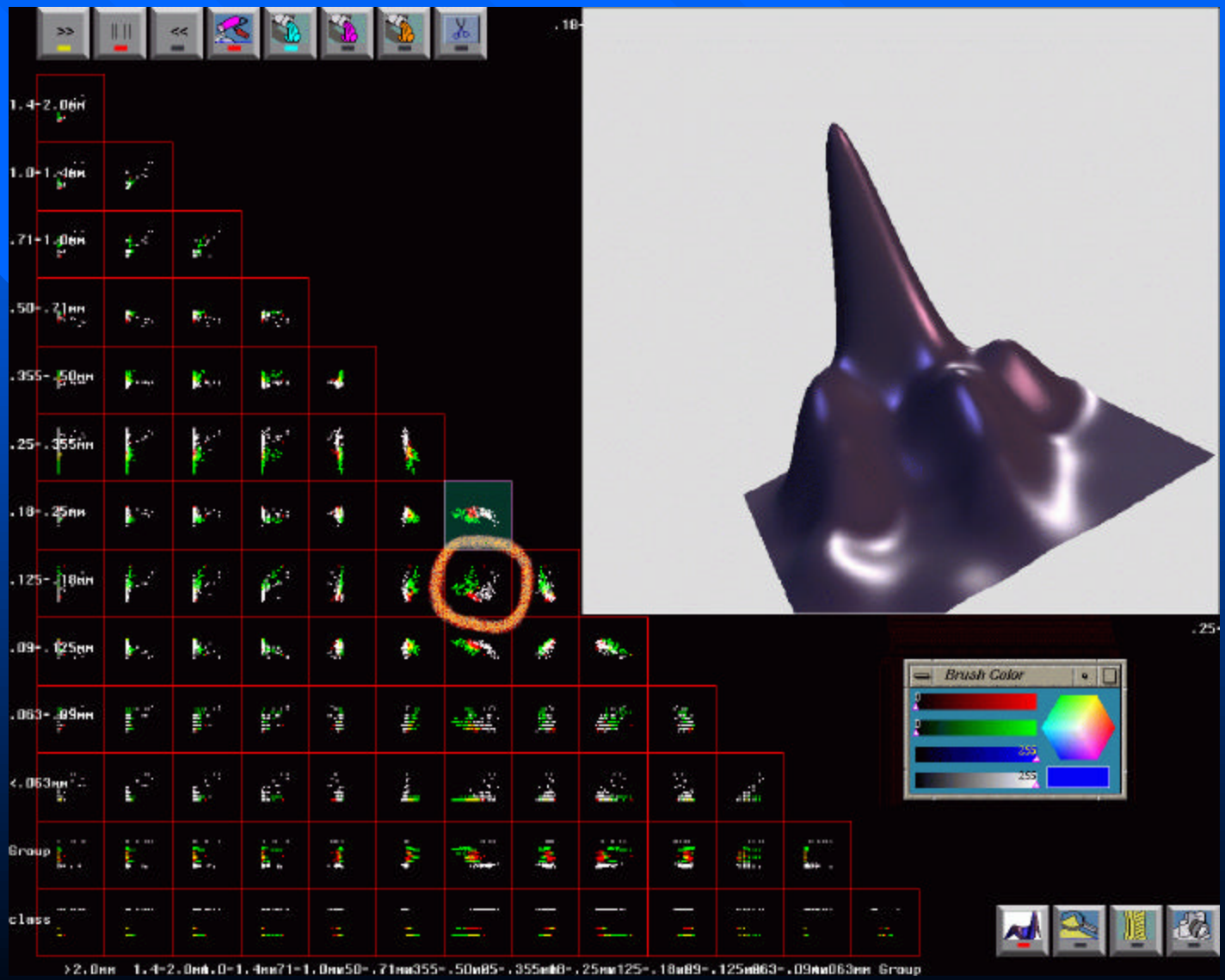
Tools: Parallel Coordinate Plots

ExplorN



Tools: Scatterplot Matrix

ExplorN

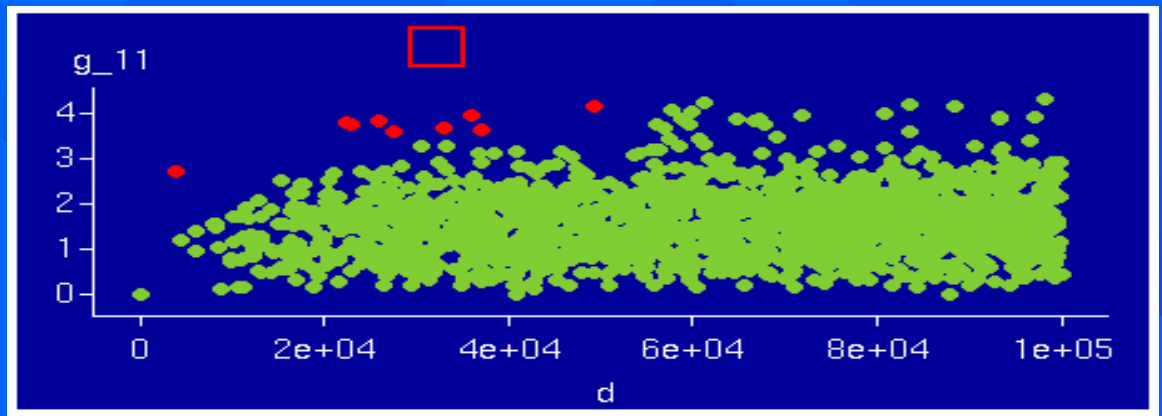


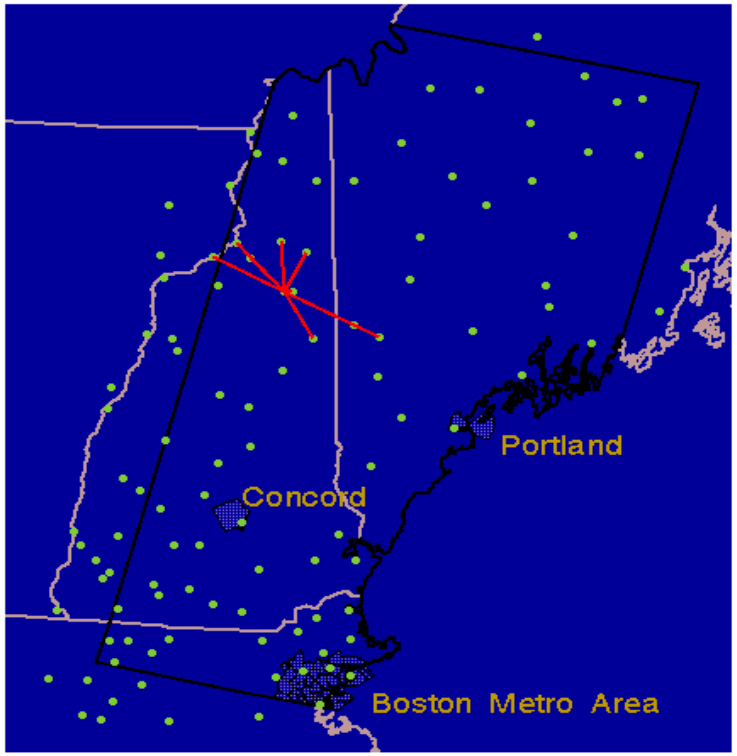
Applications

- Environmental Data
- Remote Sensing
- Archaeological Data
- Human Motion Data

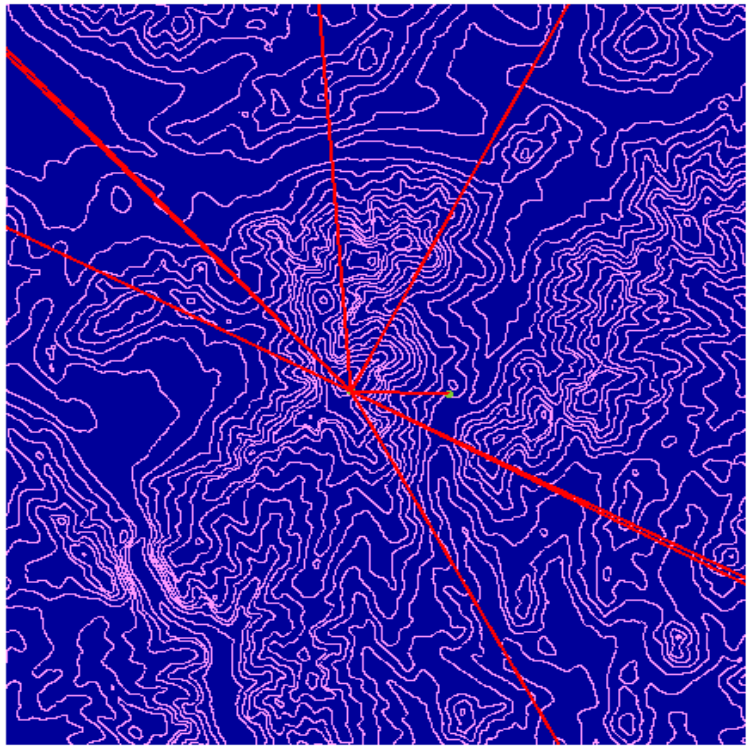
Environmental Data

XGobi

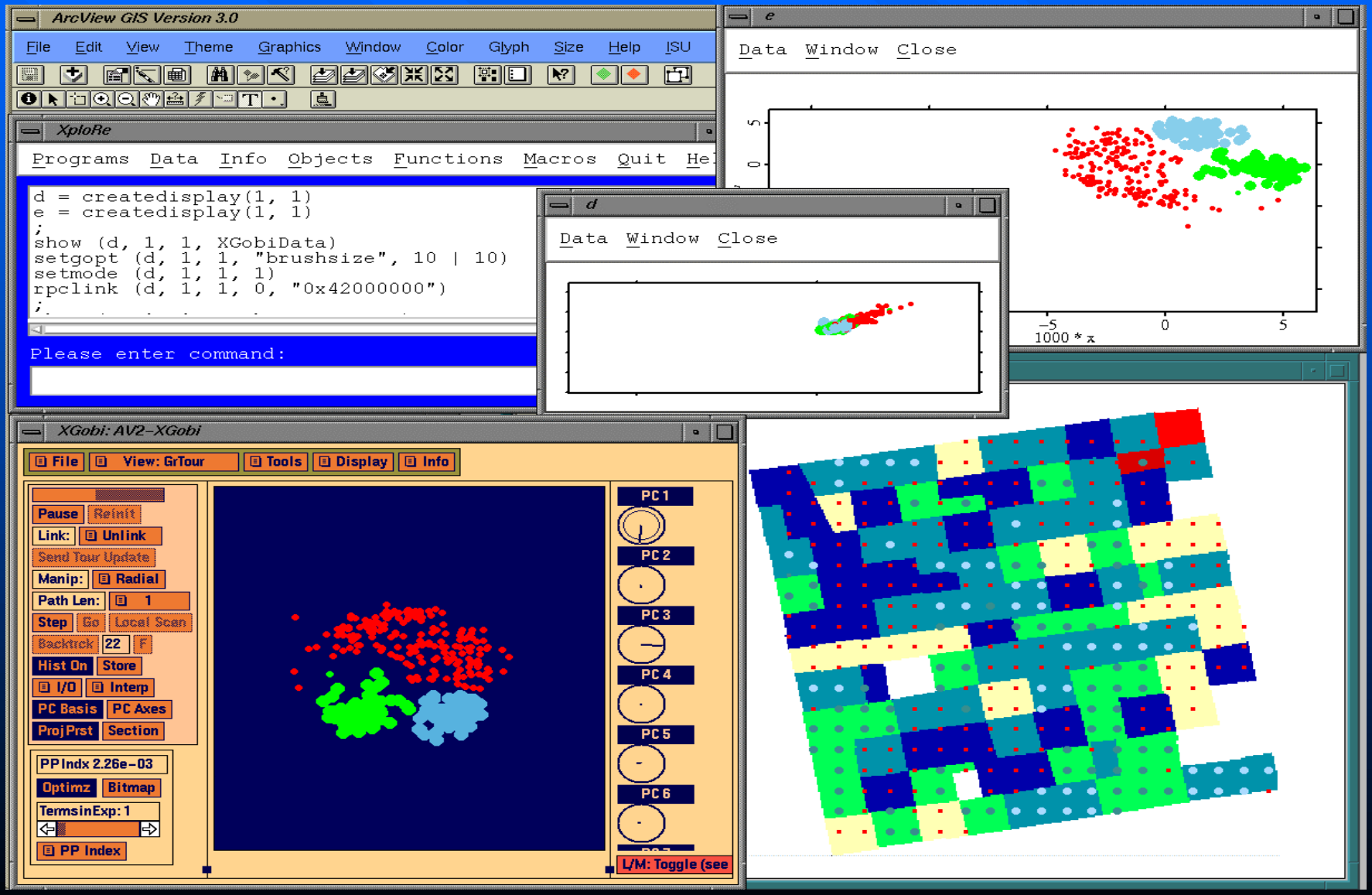




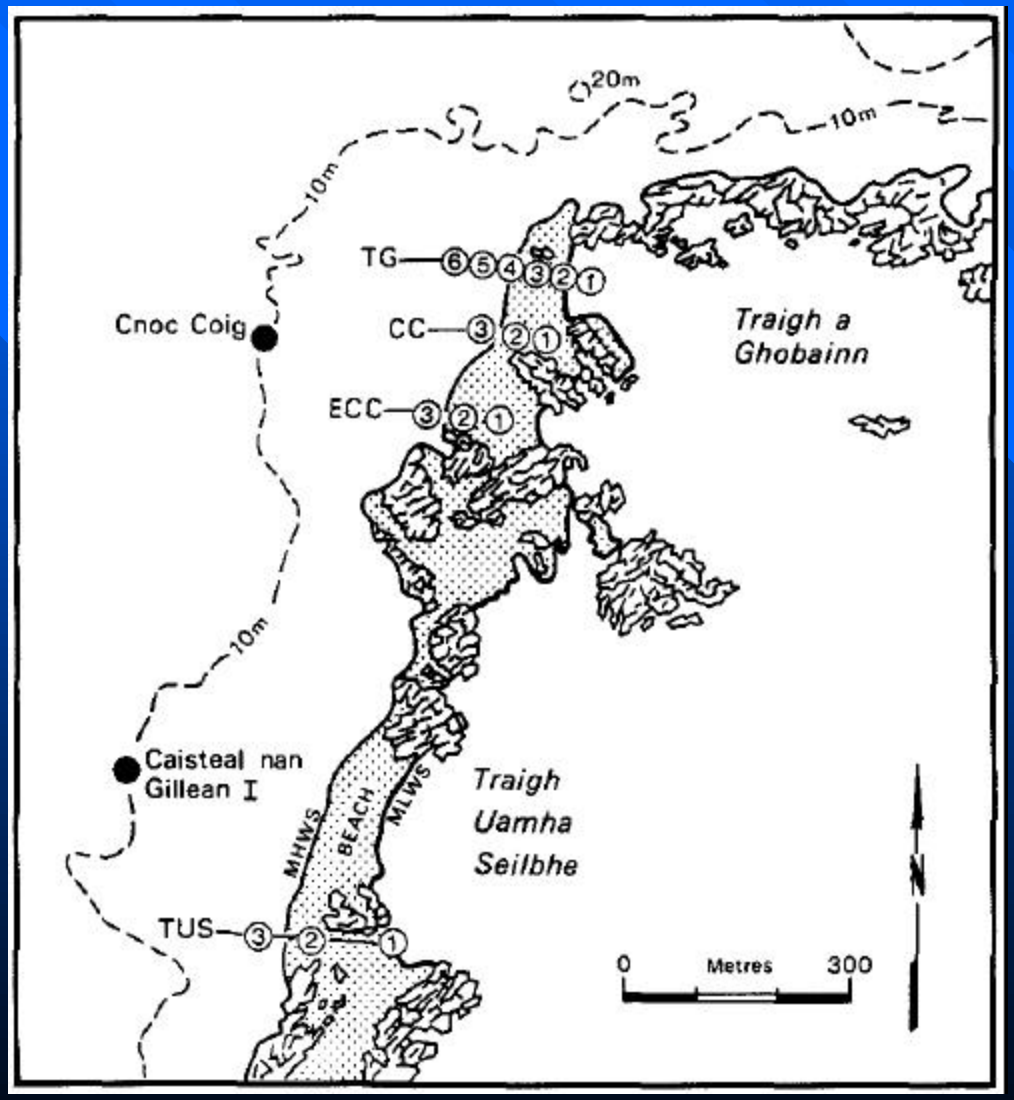
ArcView



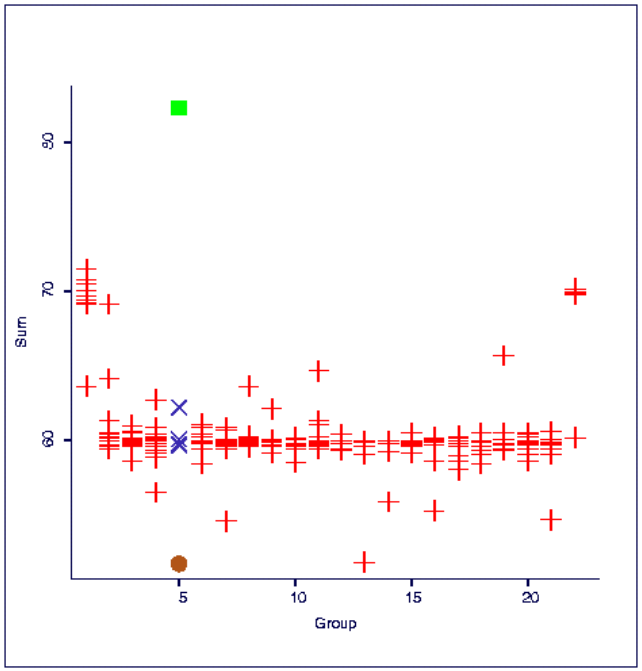
Remote Sensing

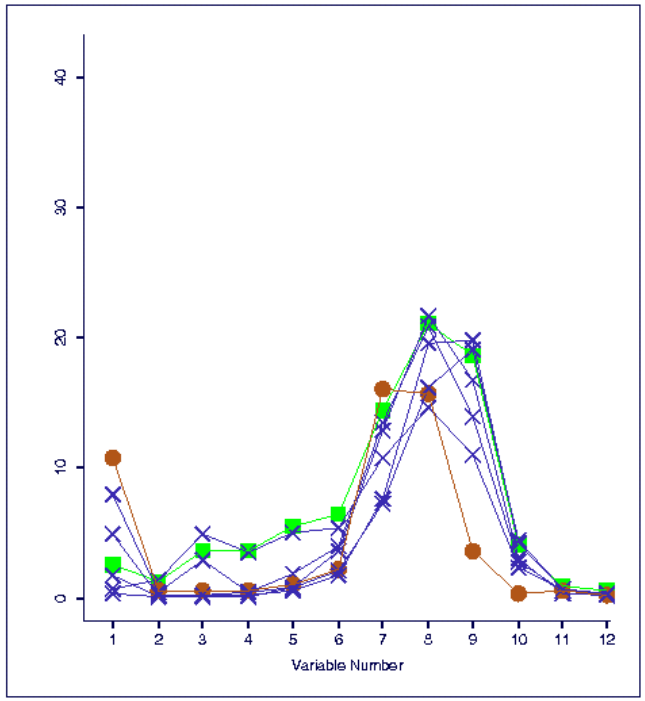


Archaeological Data 1

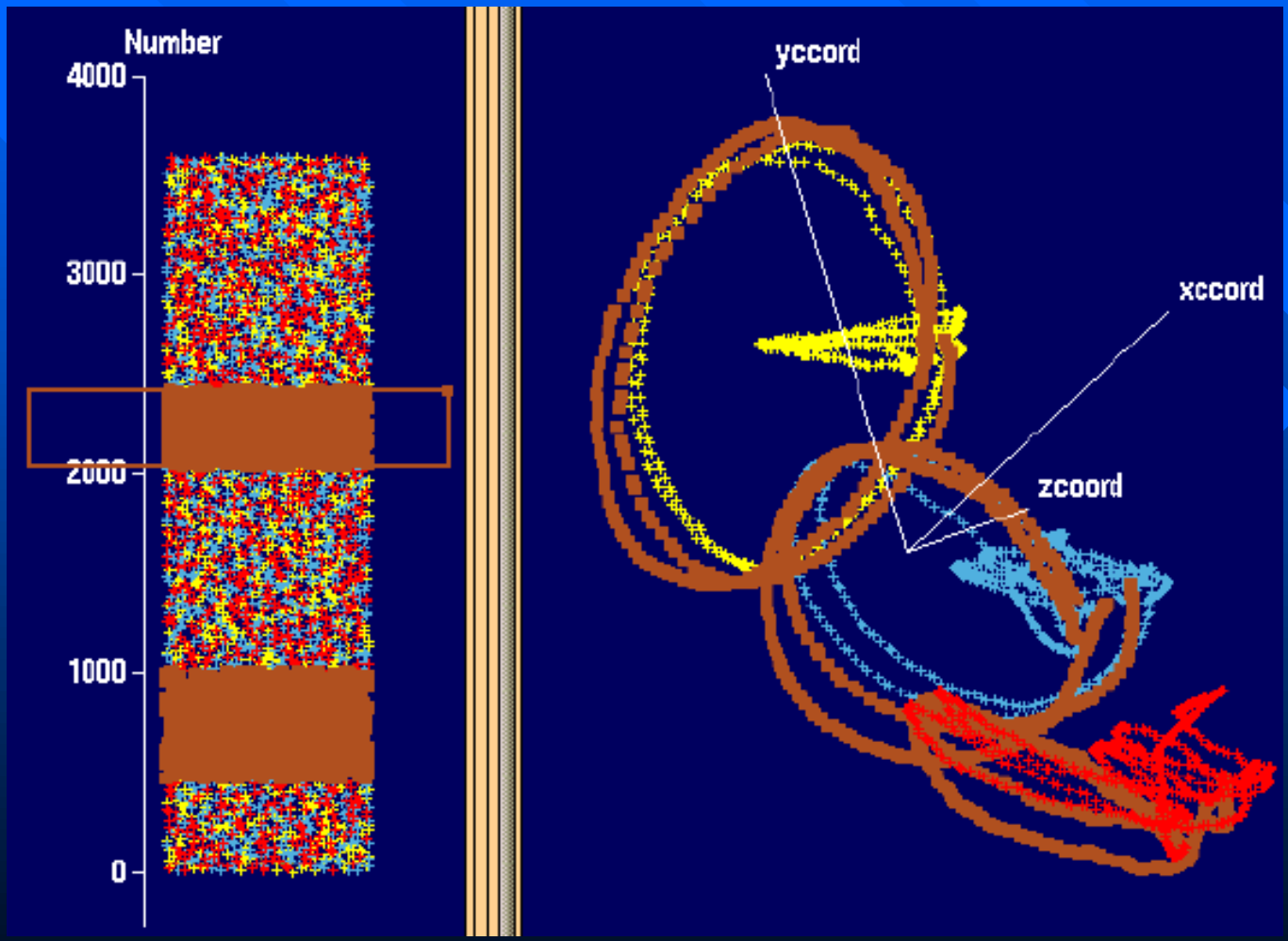


Archaeological Data 2

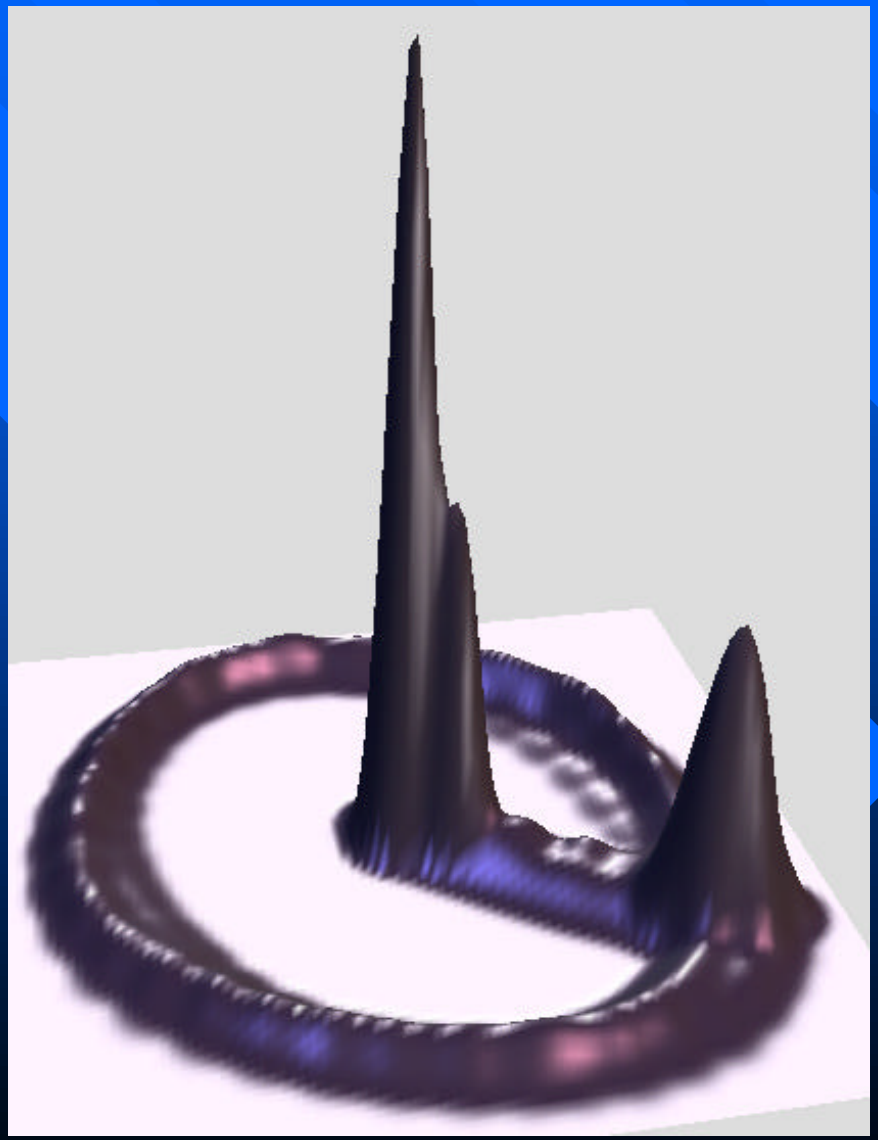




Human Motion Data 1



Human Motion Data 2

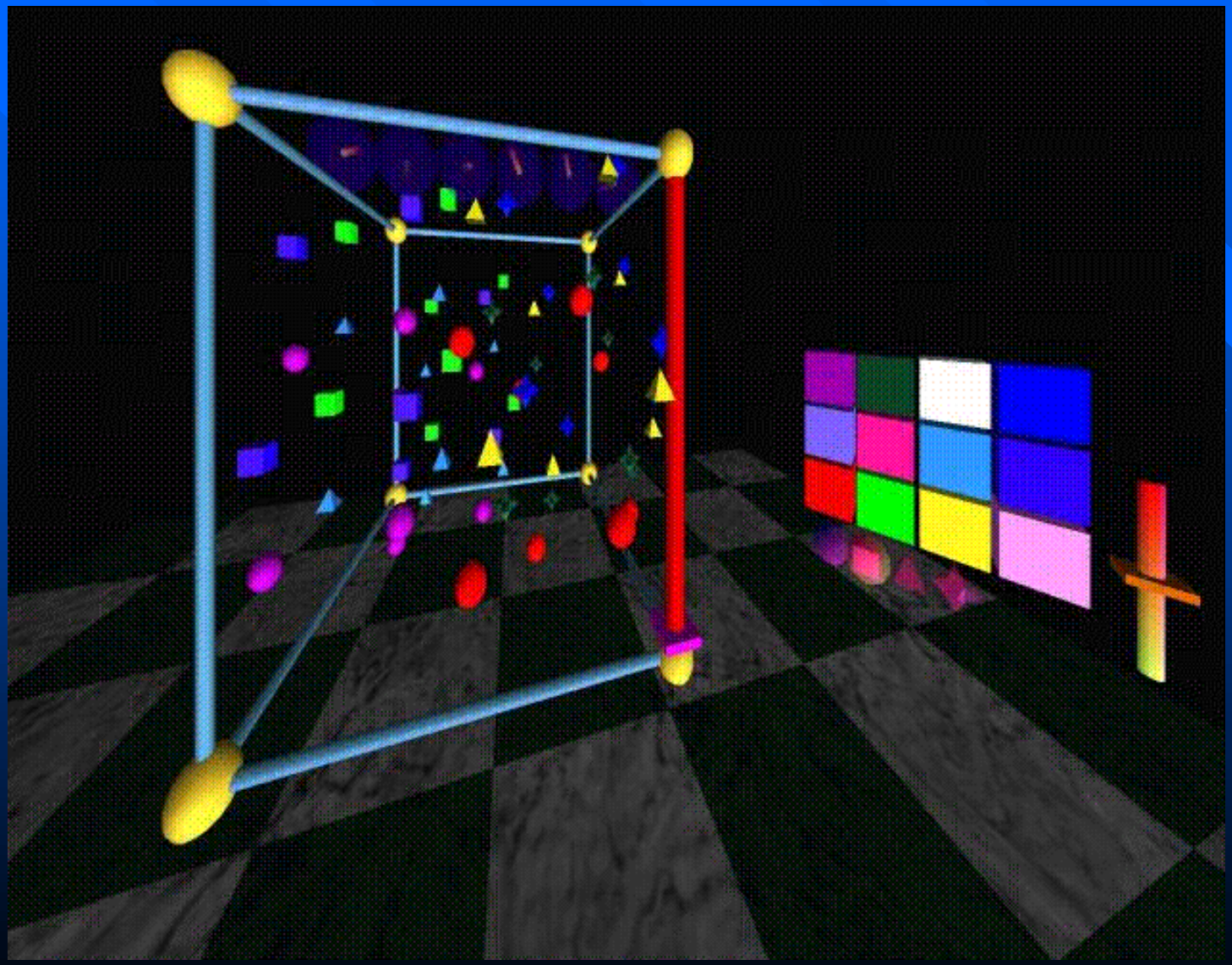


IPT / CAVE / MiniCAVE

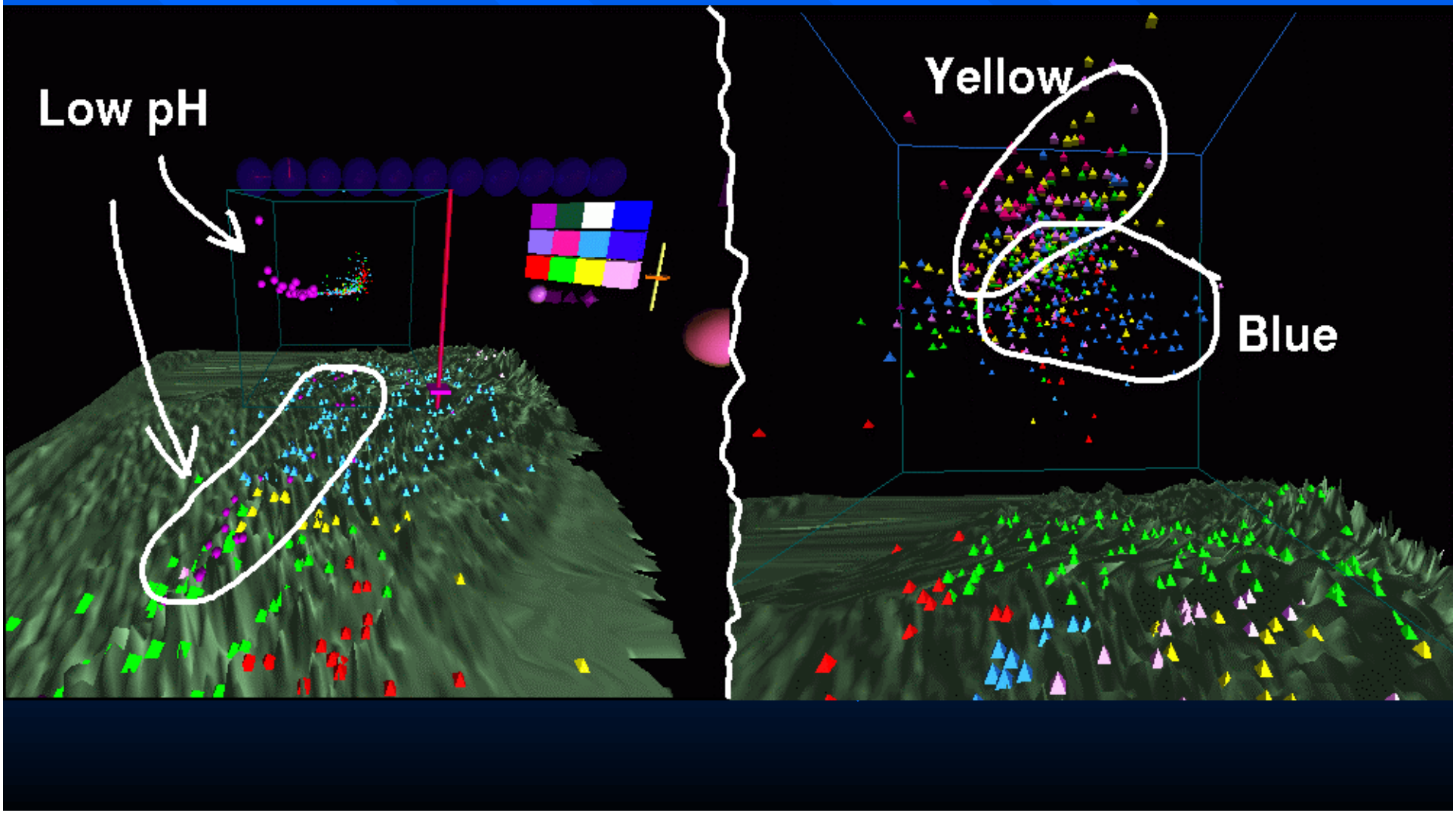
CAVE Concept

- A Projection-based Immersive VR System
  - Silicon Graphics-based with 8 to 12 processors
  - RE^2 or RE Infinity graphics engines
  - CRT-based projection system
  - Stereographics Crystal Eyes shutter glasses
  - Head tracking
  - Usually 3 to 5 wall cube
- Developed originally at U. Illinois, Chicago
  - Carolina Cruz-Neira

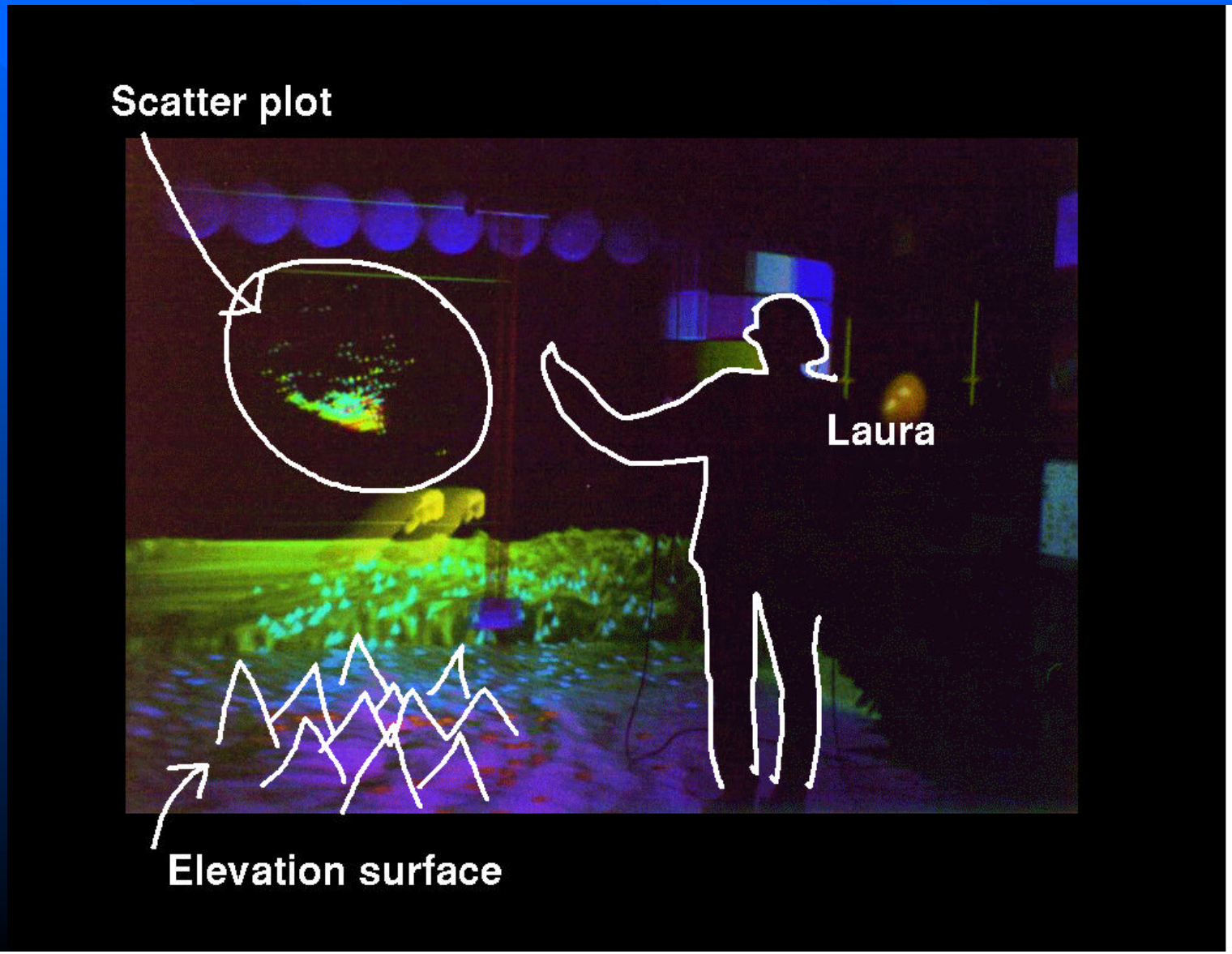
Statistics in the C2 - 1



Statistics in the C2 - 2



Statistics in the C2 - 3



CAVE Strengths

- Effective immersive environment
  - Lightweight non-intrusive glasses
  - Can see own hands and other participants
- Effective for group VR
  - Good tool for group collaboration

CAVE Weaknesses 1

- CRT Projectors
  - Projectors not very bright
  - Shock, vibration & heat, hard to keep focus
  - Geometric distortion at wall interfaces
- Tracking
  - One user tracked, badly distorted stereo for users not at viewpoint
- User Interface
  - Usually 3-D extension of desktop metaphor

CAVE Weaknesses 2

- Expensive
  - $1,000,000 fully outfitted
  - $600,000+ SGI computers
  - $30,000 per projector

Motivation for MiniCAVE

- Installed MATLAB 5 on SGI Onyx and Pentium
  - Benchmarks on 200 megahertz Pentium Pro ($3000) and 200 megahertz SGI Onyx ($120,000) similar
- Liquid Crystal Projectors sharp, bright, and stable under shock, vibration and temperature variation
- Stereographics Crystal Eyes technology available for Windows NT

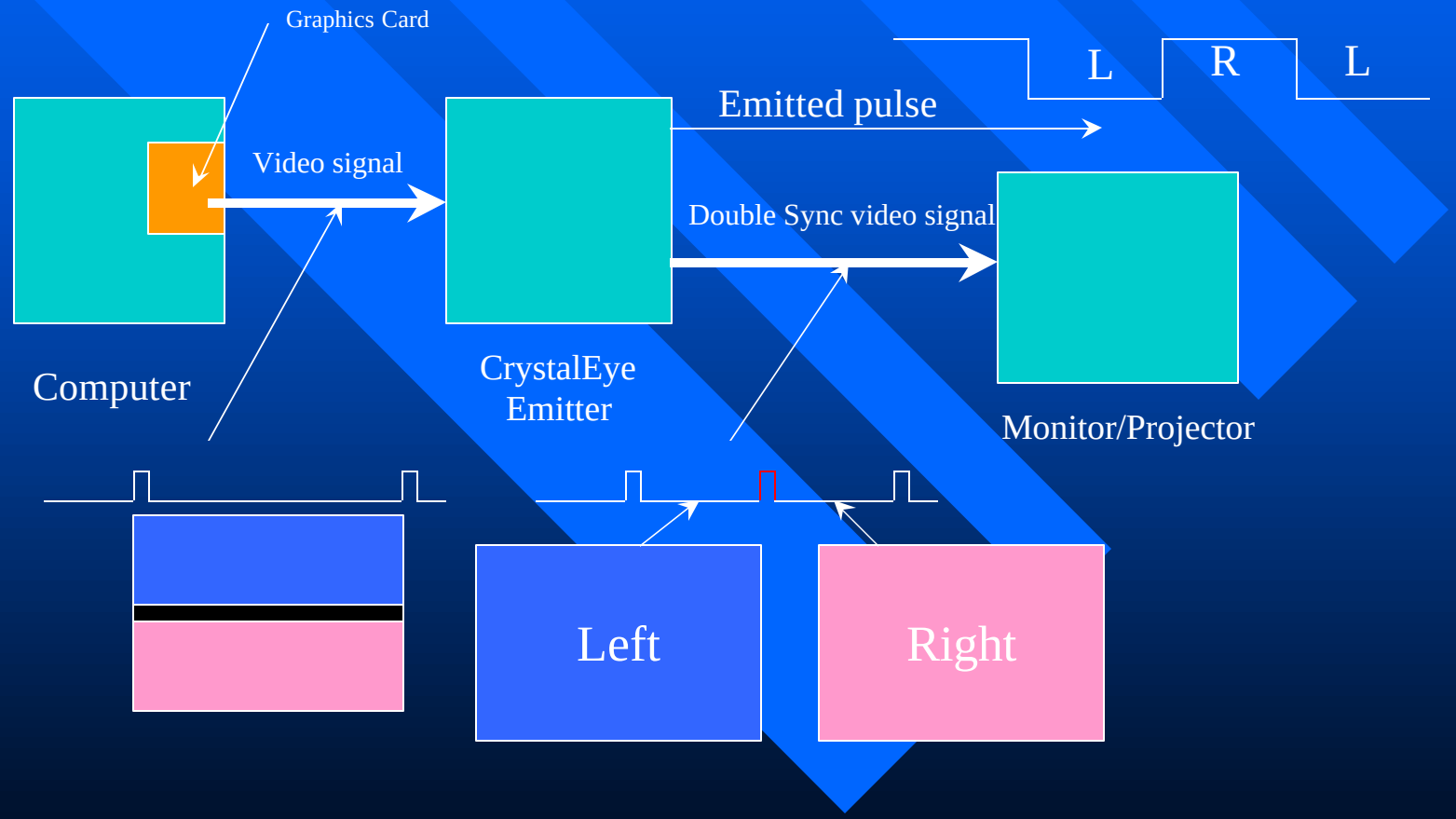
MiniCAVE Concept

- Windows NT/Intel Pentium II 400 mhz
- LCD-based projection systems
- 12 ft cubes scaled to 6 ft cubes
- Tracking optional, reduced latency
- Voice command metaphor
- $100,000 entry level

Implementation Using Monitor

- Stereo using CrystalEyes
  - Above-below stereo
  - Image resolution 1024x384 each eye
  - Vertical refresh rate 120-150 (60-75 each eye)
    - SGI monitor can handle both 120 and 150
    - CRT projector can only handle 120 refresh rate

Principles of Above-Below CrystalEyes Stereo



Speech Motivation

- User Interfaces (Van Dam)
- Shortcuts in XGobi
- User Controls in the C2 Stats Application

MiniCAVE - Successes

- Port of SkyFly Stereoscopic Demo to NT successful with adequate frame rates on 333 megahertz machine
- CrystalEyes interface on NT successful
- Speech recognition using Dragon Dictate successful
  - but requires training of speech recognizer

MiniCAVE - Future Steps

- NSF proposal submitted - waiting for funding of required hardware
- Use MiniCAVE for Visual Data Mining

Contact

- Jürgen Symanzik
  - symanzik@sunfs.math.usu.edu
- Website
  - www.math.usu.edu/~symanzik