

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

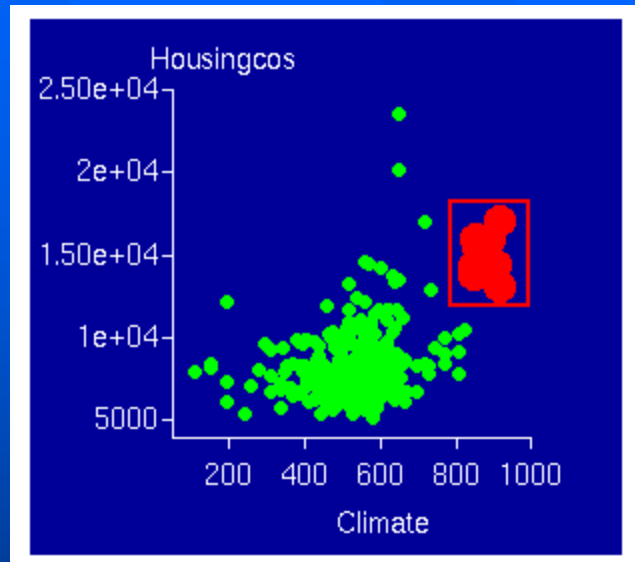
Härdle, Klinke, et al.

- 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

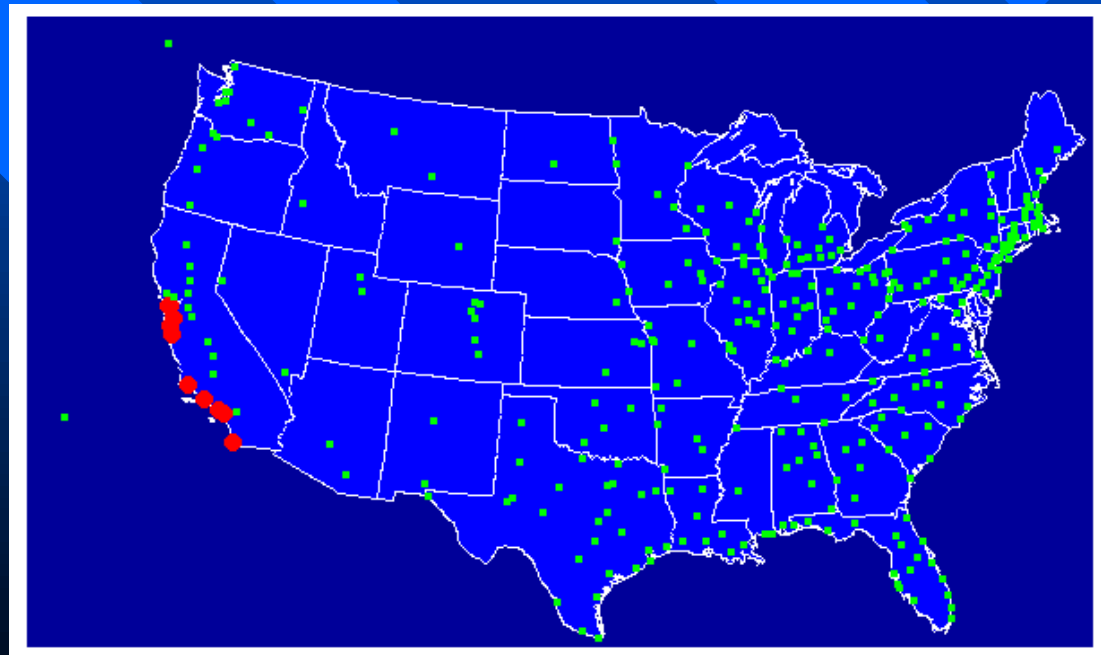
- Links these 3 packages using RPCs
- Remote Procedure Calls (RPCs):
 - Process on the local system (client) invokes a procedure on a remote system (server)
 - Request = client's desire to execute a particular remote procedure
 - Response = result produced by the remote procedure
- ArcView, XGobi, XploRe: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

Tools: Linked Brushing 1

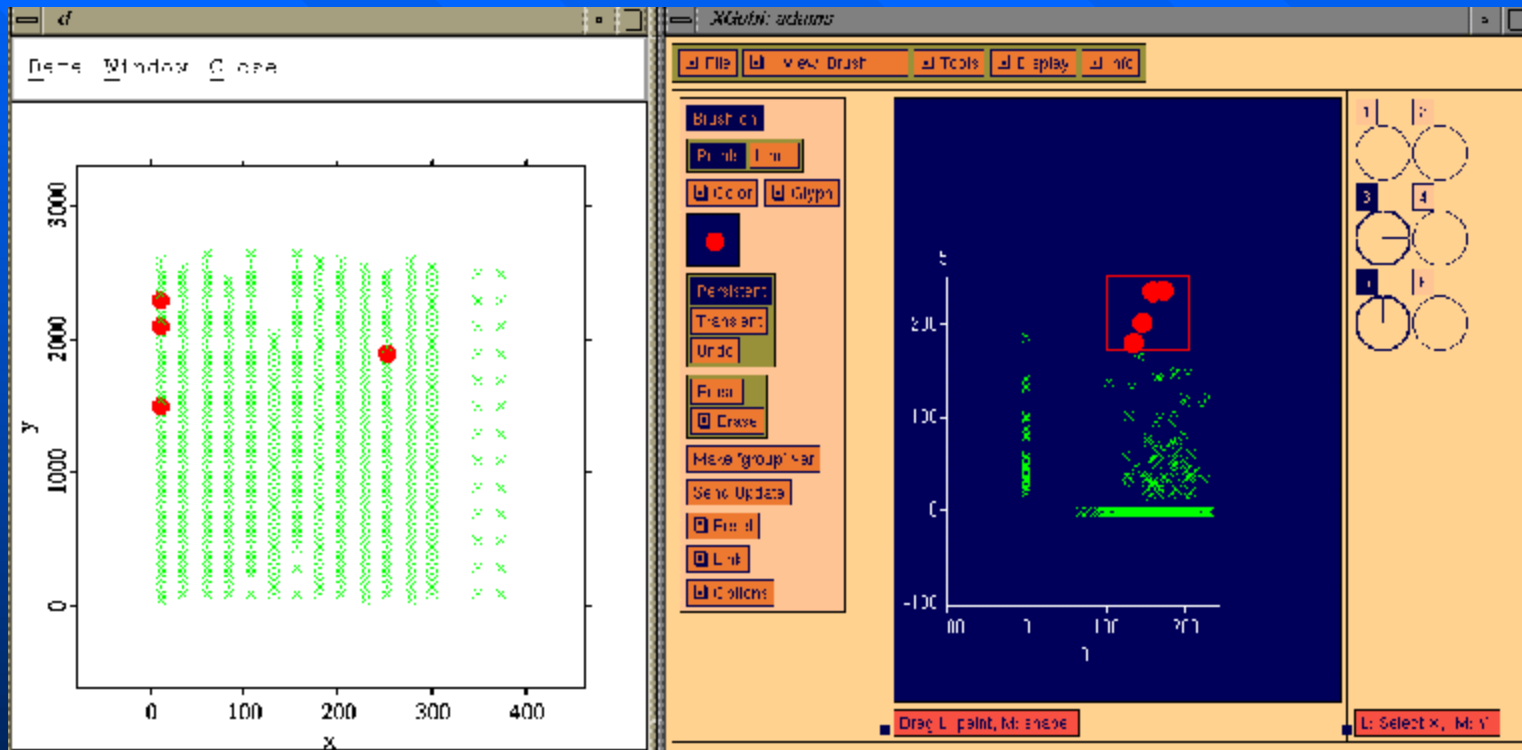


XGobi

ArcView



Tools: Linked Brushing 2

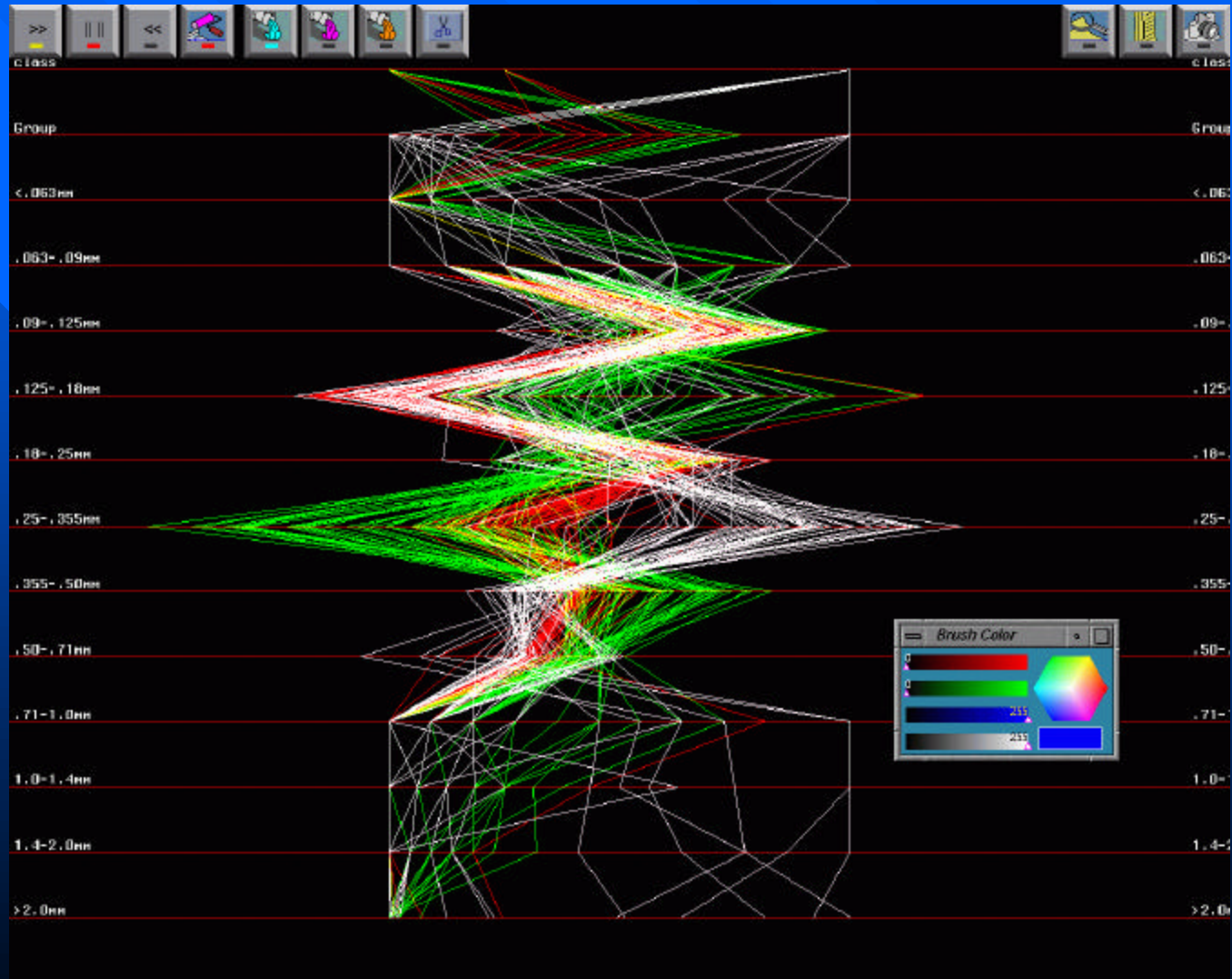


XploRe

XGobi

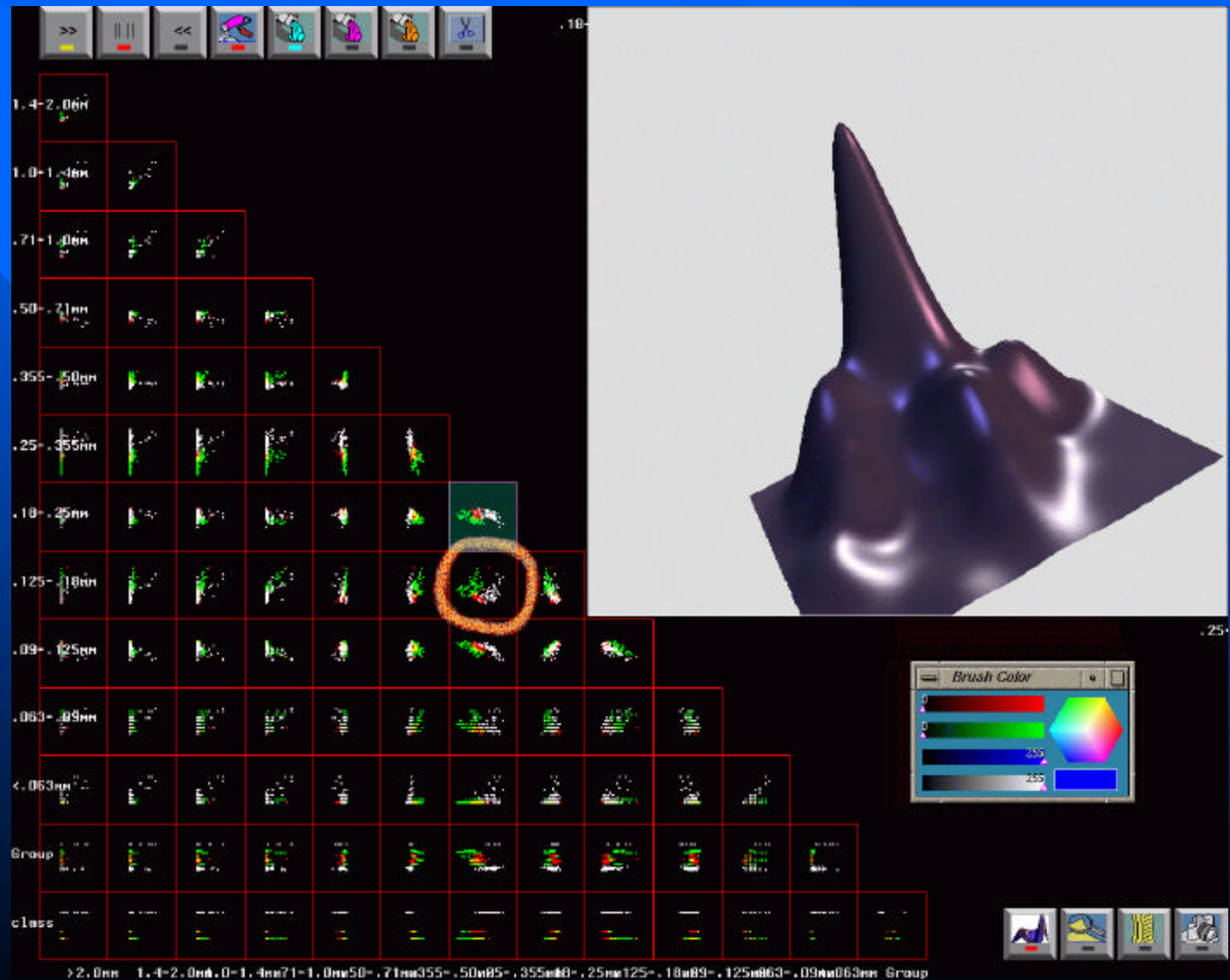
Tools: Parallel Coordinate Plots

ExplorN



Tools: Scatterplot Matrix

ExplorN

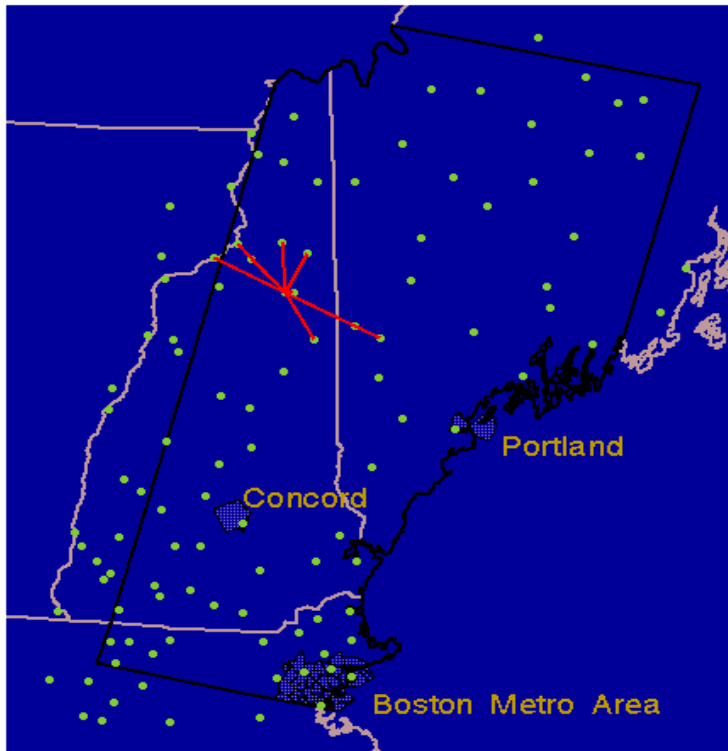
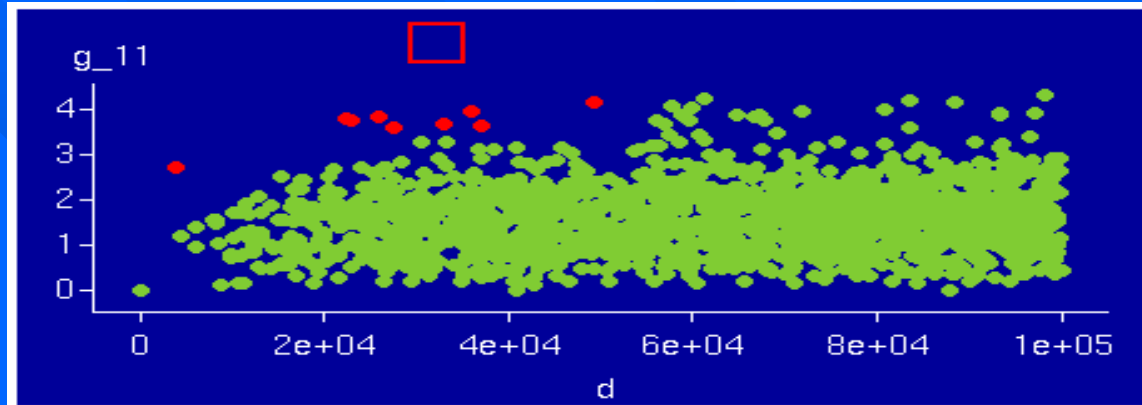


Applications

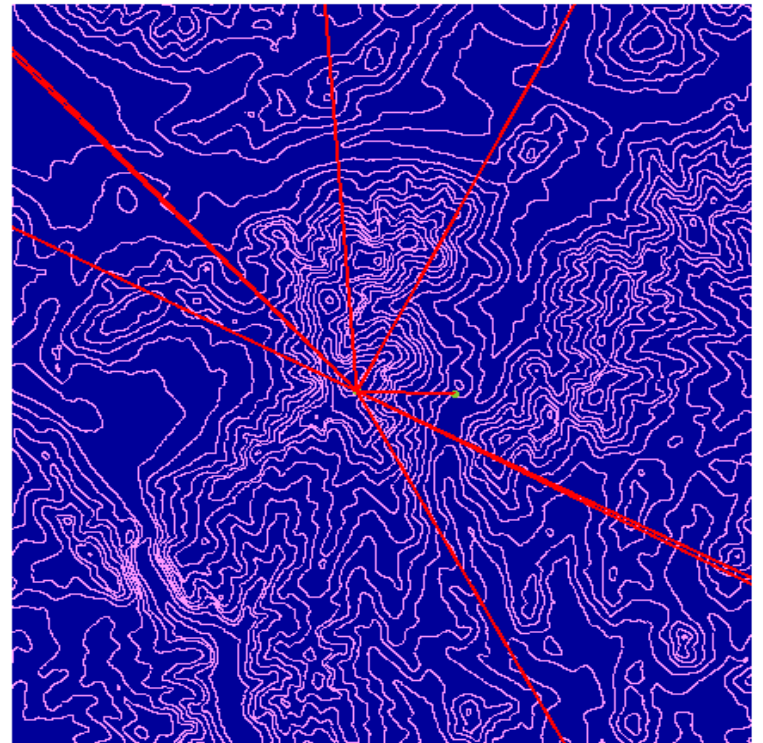
- Environmental Data
- Remote Sensing
- Archaeological Data
- Human Motion Data

Environmental Data

XGobi



ArcView



Remote Sensing

The image displays a composite of software interfaces for remote sensing data analysis. At the top, the title "Remote Sensing" is written in a large, yellow, underlined font against a blue background.

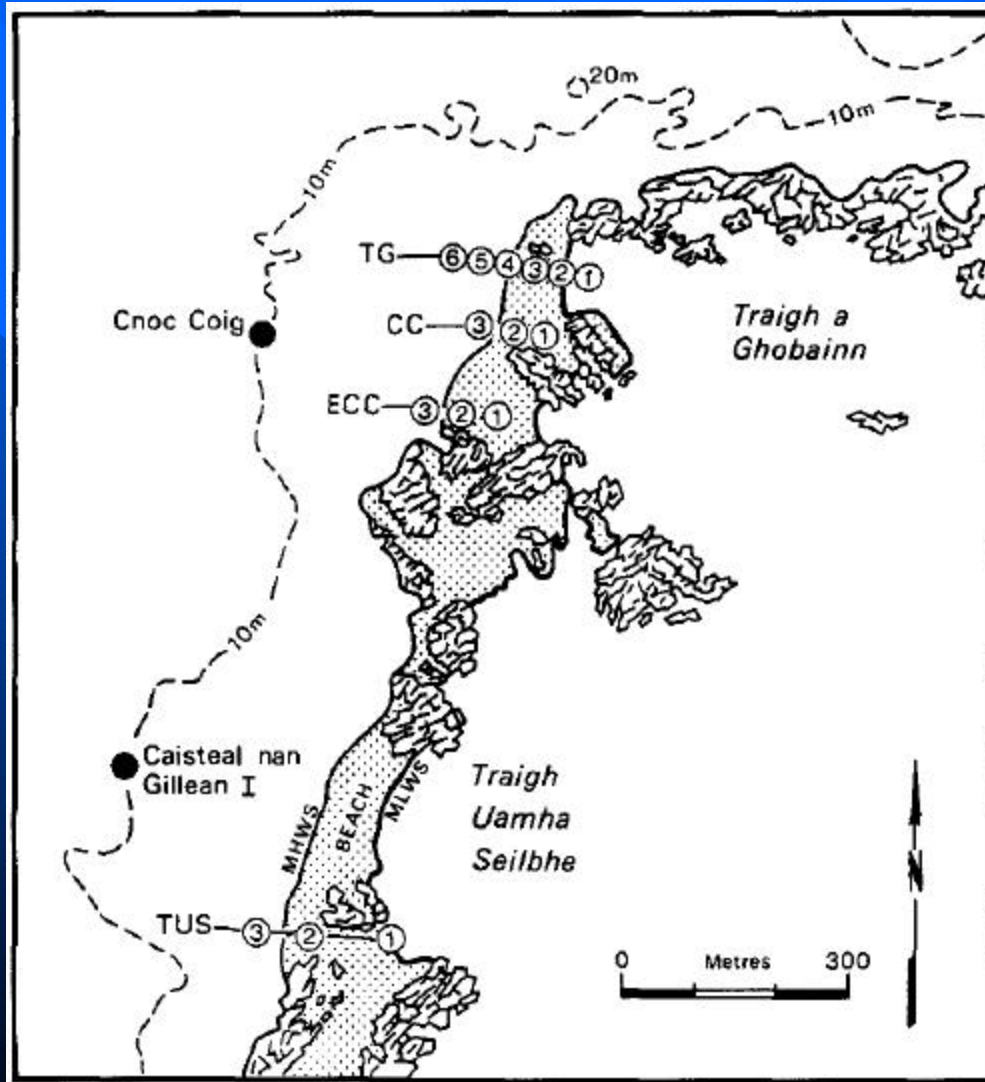
The main interface is ArcView GIS Version 3.0. It features a menu bar (File, Edit, View, Theme, Graphics, Window, Color, Glyph, Size, Help, ISU) and a toolbar with various icons. A script window titled "XploRe" contains the following commands:

```
d = createdisplay(1, 1)
e = createdisplay(1, 1)
;
show (d, 1, 1, XGobiData)
setgopt (d, 1, 1, "brushsize", 10 | 10)
setmode (d, 1, 1, 1)
rpclink (d, 1, 1, 0, "0x42000000")
;
```

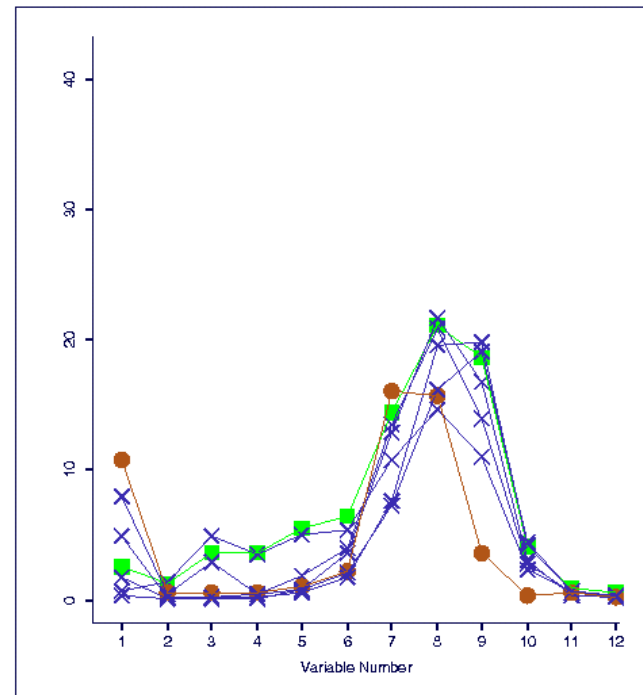
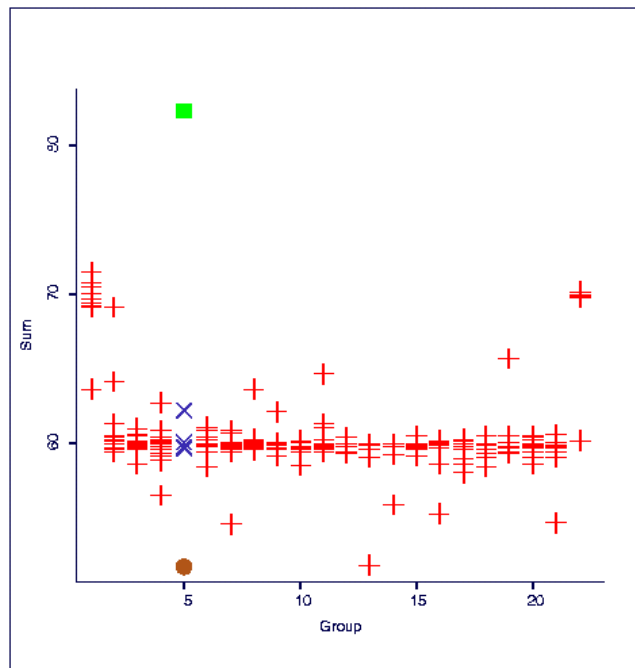
Below the script is a prompt: "Please enter command:". To the right, two data windows are visible. The top window, titled "e", shows a scatter plot of data points colored red, blue, and green. The bottom window, titled "d", shows a smaller version of the same data. The x-axis of the "e" window is labeled $1000 * x$ with tick marks at -5, 0, and 5.

At the bottom, the "XGobi: AV2-XGobi" interface is shown. It has a menu bar (File, View: GrTour, Tools, Display, Info) and a control panel with buttons for "Pause", "Reinit", "Link", "Unlink", "Send Tour Update", "Manip", "Radial", "Path Len", "1", "Step", "Go", "Local Scan", "Backtrck", "22", "F", "Hist On", "Store", "I/O", "Interp", "PC Basis", "PC Axes", "ProjPrst", "Section", "PP Indx", "2.26e-03", "Optimz", "Bitmap", "TermsinExp", "1", and "PP Index". The main display area shows a 2D scatter plot of data points colored red, blue, and green. To the right of the plot are six circular buttons labeled "PC 1" through "PC 6". At the bottom right, there is a button labeled "L/M: Toggle (see)".

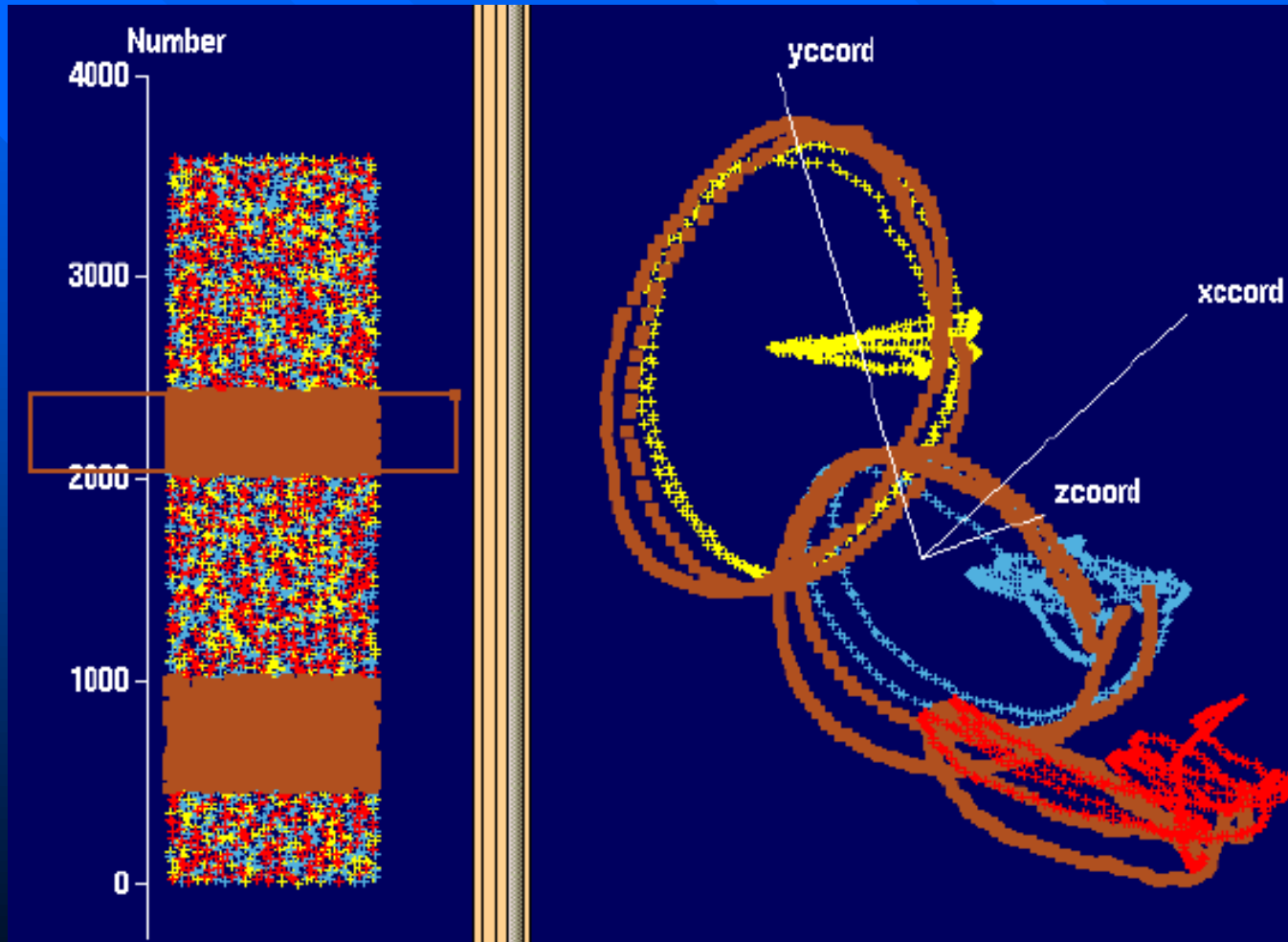
Archaeological Data 1



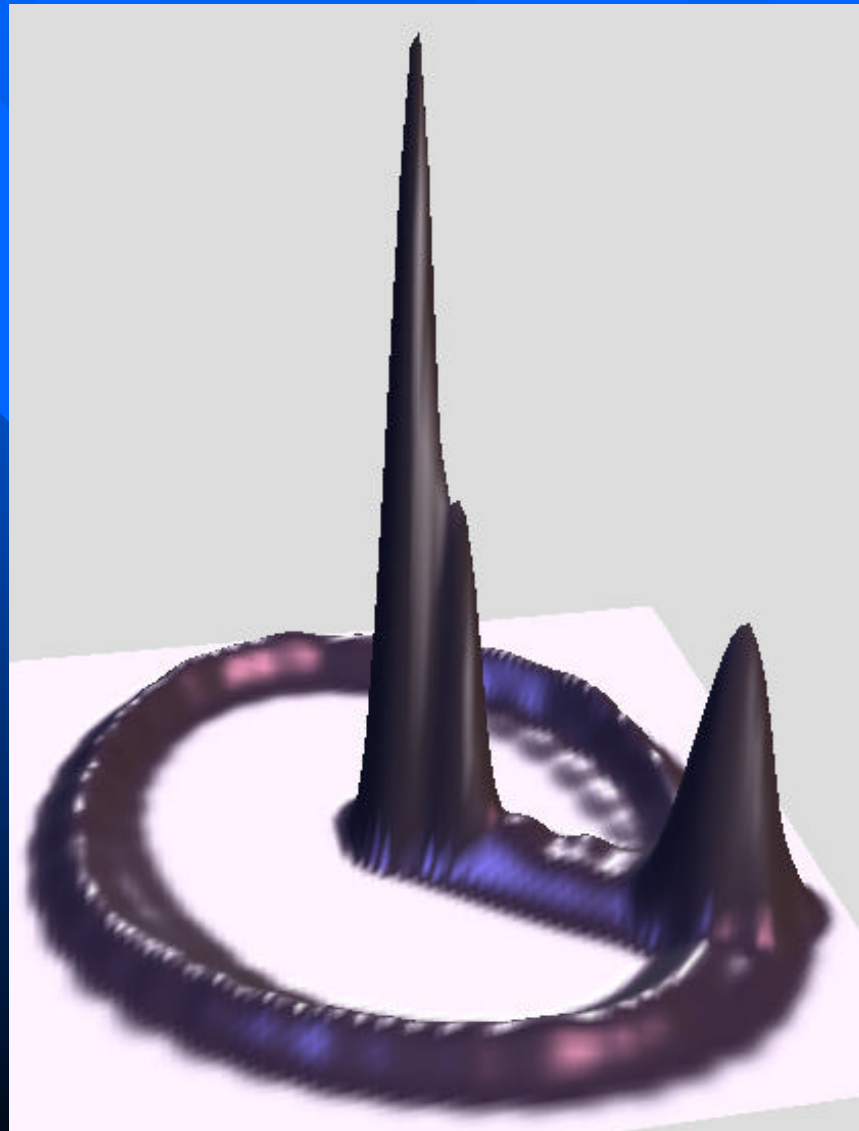
Archaeological Data 2



Human Motion Data 1



Human Motion Data 2

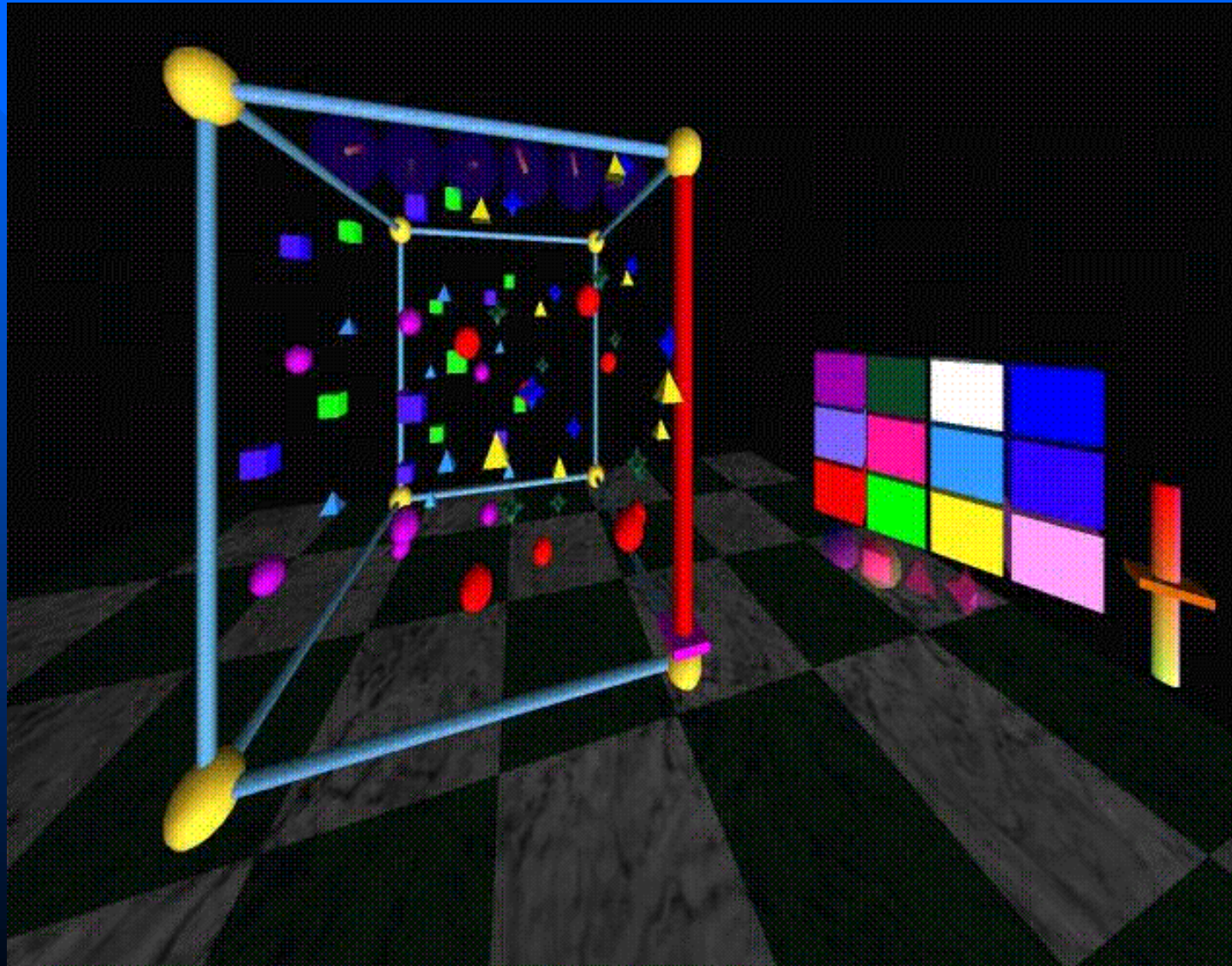


IPT / CAVE / MiniCAVE

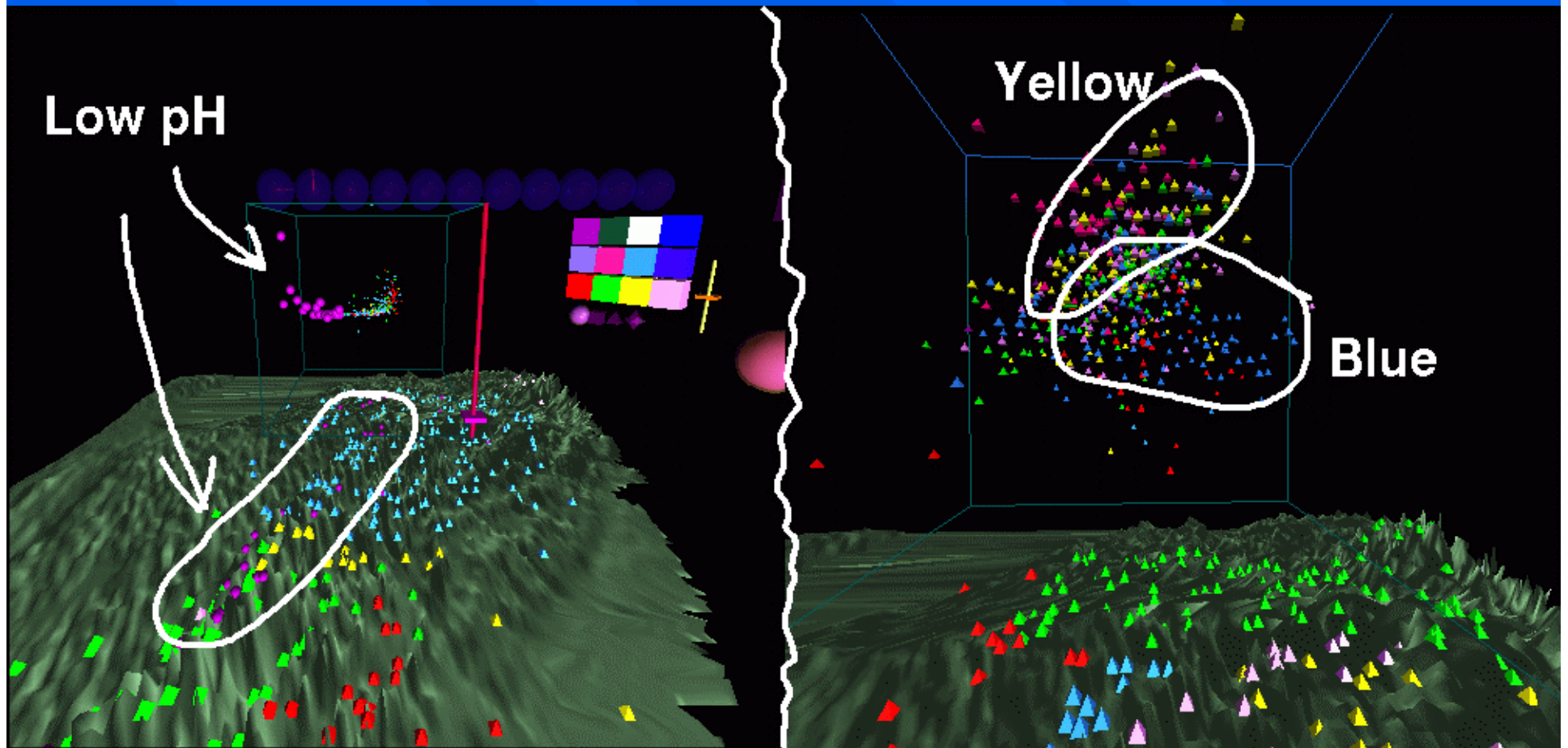
CAVE Concept

- A Projection-based Immersive VR System
 - Silicon Graphics-based with 8 to 12 processors
 - RE² 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

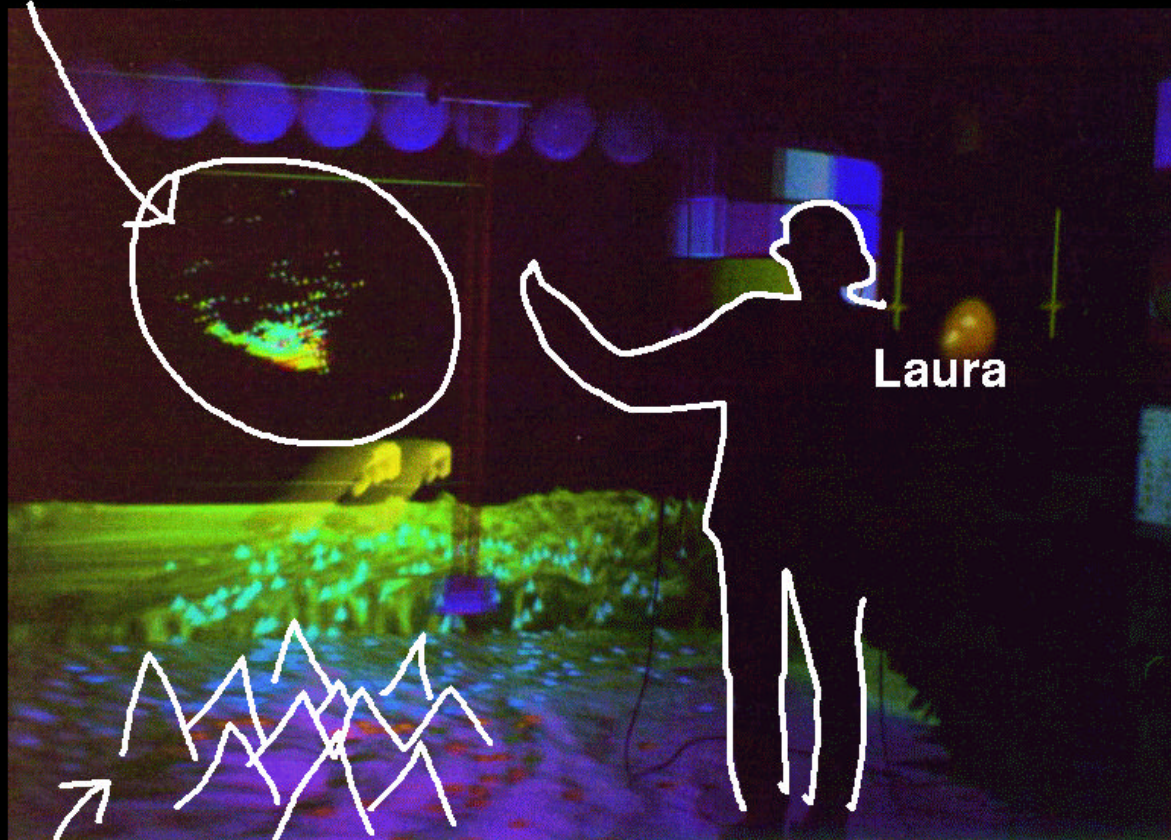


Statistics in the C2 - 2



Statistics in the C2 - 3

Scatter plot



Laura

Elevation surface

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

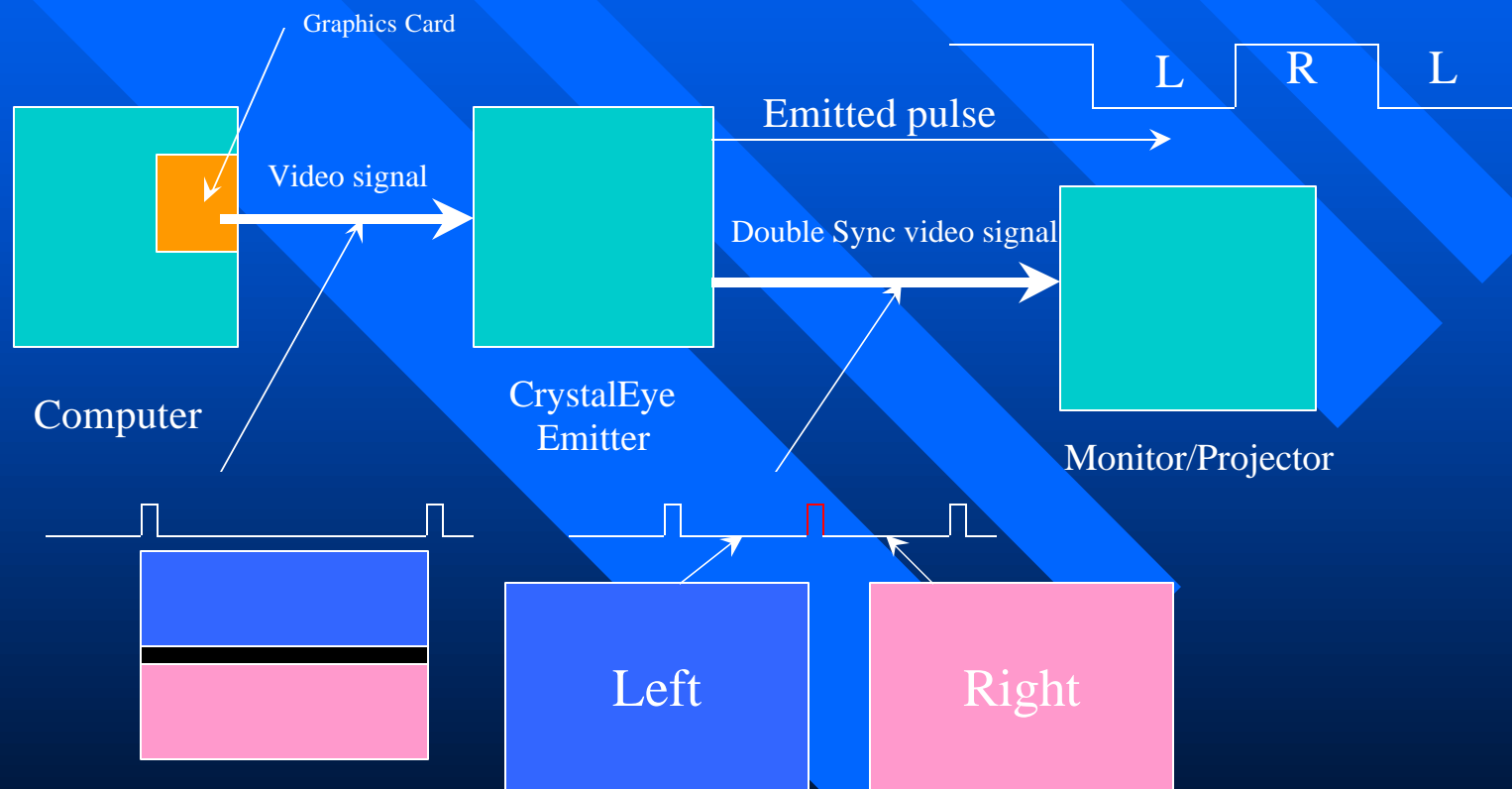
- Windows NT/Intel Pentium II 400 mhz
- LCD-based projection systems
- 12 ft cubes scaled to 6 ft cubes
- Tracking optional, reduced latency
- Voice command metaphor
- \$100,000 entry level

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