visualization

Santiago V Lombeyda | Center for Advanced Computing Research | Caltech
data \rightarrow \text{questions} \rightarrow \text{analytical answers} \rightarrow \text{new questions} \rightarrow \text{visual inspection answers}
what includes visualization
data
more understandable
visualization =

science

+ 

computer graphics/hci

+ 

graphic design/art
finding solution(s) via purpose

for what purpose do we visualize?

quick view/demonstration
we want to look at/show something particular

analysis
we know what we are looking for

explore
we do not know what we are looking for

debugging
we want to assure there is nothing odd

...
process dictated by the data

A CLOSER LOOK AT THE “DATA”
data: GEOMETRIC STRUCTURE

abstract

multi-dimensional data RECORDS
mapping + paradigms \rightarrow interaction

2d/3d data

scalar/vector/tensor + time
paradigms \rightarrow interaction
the more main stream viz
data: GEOMETRIC STRUCTURE

abstract

MPG Cylinders Horsepower Weight Acceleration Year Origin
8.5 0.4 2.8 8.2 4 40.4 1500.4 5500.4 65.30.4 69.5 82.5 4.8 3.2 3
18.000000 8.000000 30.000000 3504.000000 12.000000 70.000000 1.000000
15.000000 8.000000 165.000000 3693.000000 11.500000 70.000000 1.000000
18.000000 8.000000 150.000000 3436.000000 11.000000 70.000000 1.000000
16.000000 8.000000 150.000000 3433.000000 12.000000 70.000000 1.000000
17.000000 8.000000 140.000000 3449.000000 10.500000 70.000000 1.000000
...
understanding

THE VISUALIZATION PROCESS
[ Ben Fry, *Visualizing Data*, O'Reilley Media, 2008 ]
**Acquire**
- Obtain the data, whether from a file on a disk or a source over a network.

**Parse**
- Provide some structure for the data's meaning, and order it into categories.

**Filter**
- Remove all but the data of interest.

**Mine**
- Apply methods from statistics or data mining as a way to discern patterns or place data in mathematical context.

**Represent**
- Choose a basic visual model, such as a bar graph, list, or tree.

**Refine**
- Improve the basic representation to make it clearer and more visually engaging.

**Interact**
- Add methods for manipulating the data or controlling what features are visible.

[Ben Fry, *Visualizing Data*, O'Reilly Media, 2008]
usual visualization engine

data processing algorithms

geometry generation

rendering

ui

interaction
"the" visualization toolkit

- rendering
- geometry generation
- processing algorithms
- data
- interaction
- ui

Languages:
- VTK
- c/c++
- tcl/tk
- python
- java
- R
interactive renderers

- processing algorithms
- geometry generation
- rendering
- data
- interaction
- ui

Higher level:
- OpenGL
- OpenInventor
- COIN
- Java3D
- 03D (by Google)
- WebGL
- mesaGL
- DirectX
- C++/GL
- Java/GL
ray tracers

processing algorithms

geometry generation

rendering

interaction

data

ui

POVRay

RenderMan

Blender

Maya

MODELLERS:
tsunami (artic region sc)
gui toolkits

- QT
- GTK+
- TK (TCL/TK)
- Java Swing, Cocoa, Motiff…

data

processing algorithms

geometry generation

rendering

interaction

ui
web based ui

data

processing algorithms

geometry generation

rendering

interaction

ui

dHTML (HTML5 or SVG)/ JavaScript
Processing
Java
JProcessing
Flash
visualization system

- Paraview\textsuperscript{VTK}
- LLNL VisIt\textsuperscript{VTK}
- EnSight\$ 
- Protovis\textsuperscript{WWW}
- Many Eyes\textsuperscript{WWW}
- Wiki based
- Modrian\textsuperscript{R},
  TopCat,
  Mollegro, …
an overview: tools & techniques
INFOVIZ
data: GEOMETRIC STRUCTURE

abstract

MPG Cylinders Horsepower Weight Acceleration Year Origin
8.5 0.4 2.8 8.2 4 40.2 1500.4 5500.4 5.30.4 69.5 82.5 4.8 3.2 3
18.000000 8.000000 130.000000 3504.000000 12.000000 70.000000 1.000000
16.000000 8.000000 165.000000 3693.000000 11.500000 70.000000 1.000000
18.000000 8.000000 150.000000 3436.000000 11.000000 70.000000 1.000000
16.000000 8.000000 150.000000 3433.000000 12.000000 70.000000 1.000000
17.000000 8.000000 140.000000 3449.000000 10.500000 70.000000 1.000000
... ... ... ... ...

2d/3d data
basic infovis techniques:
N-DIMENSIONAL RECORD DATA
visual analytics goal:

(mapping data)

detect, classify, and measure

trends, outliers, patterns, clusters, correlations
principal component analysis:

\[ \text{pca} = \text{data transformation} \]

\[ n \ "\text{correlated}" \ variables \]

\[ \downarrow \]

\[ n \ "\text{uncorrelated}" \ variables \]

(akin decomposition into orthogonal element basis)
choose a layout strategy..
PARALLEL COORDINATES + CLUSTERING
3D PARALLEL COORDINATES
basic infovis techniques:

HEI RARCHI CAL DATA
DPOSS & Fisher's Iris:
MONDRIAN+R
quick demo
infovis packages

mondrian (R based)
rosuda.org/Mondrian

xmdv
davis.wpi.edu/~xmdv

molegro data modeller (bio)

topcat (astro)

many eyes (online wiki)

protovis/polaris (+DATACUBES)
COMERCIAL: tableau.com

iVu (local)

www.infovis-wiki.net
data: GEOMETRIC STRUCTURE

abstract

MPG Cylinders Horsepower Weight Acceleration Year Origin
8. 50.4 2.8 8.2 4 40. 4 1500. 5500. 4 5. 30. 4 69.5 82.5 4 .8 3.2  3
18.000000 8.000000 130.000000 3504.000000 12.000000 70.000000 1.000000
16.000000 8.000000 165.000000 3693.000000 11.500000 70.000000 1.000000
18.000000 8.000000 150.000000 3436.000000 11.000000 70.000000 1.000000
16.000000 8.000000 150.000000 3433.000000 12.000000 70.000000 1.000000
17.000000 8.000000 140.000000 3449.000000 10.500000 70.000000 1.000000
... ... ... ... ... ... ... ... ...
multivariable visualization...

how many variables can you visualize? as many as you want!

how many variables can you understand? 3? 4?
Collin Ware study:

**encoding 4 variables**

i.e. encoding 2 variables on map

\[(x, y, u, v)\]

**Qton:**

putative units of pre-attentive human texture perception, analogous to a phoneme in speech recognition
error: integral separable
reading order/focal flow

what do you see first? second?
why is it important?
what does unknown order create?
science (data analysis, visual analytics) +

+ graphics/ hci

+ graphic design/ art (human perception, aesthetics)

visualization
visualization =

science

+ 

computer graphics/hci

+ 

graphic design/art
quick looks at...
visualization

ay/bi199: methods of computational science

santiago v lombeyda | center for advanced computing research | caltech