

CSE 564
VISUALIZATION & VISUAL ANALYTICS
INTERACTION & INFORMATION
NAVIGATION

KLAUS MUELLER

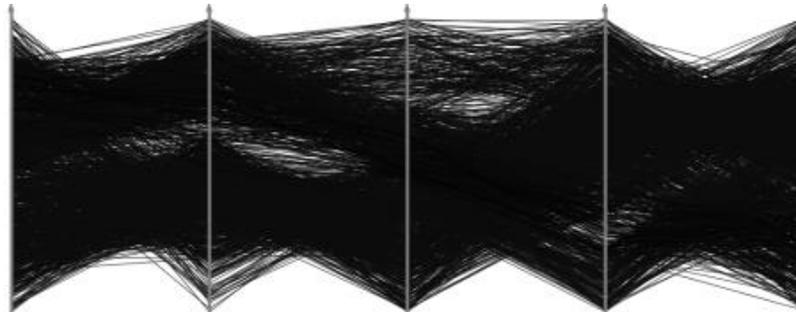
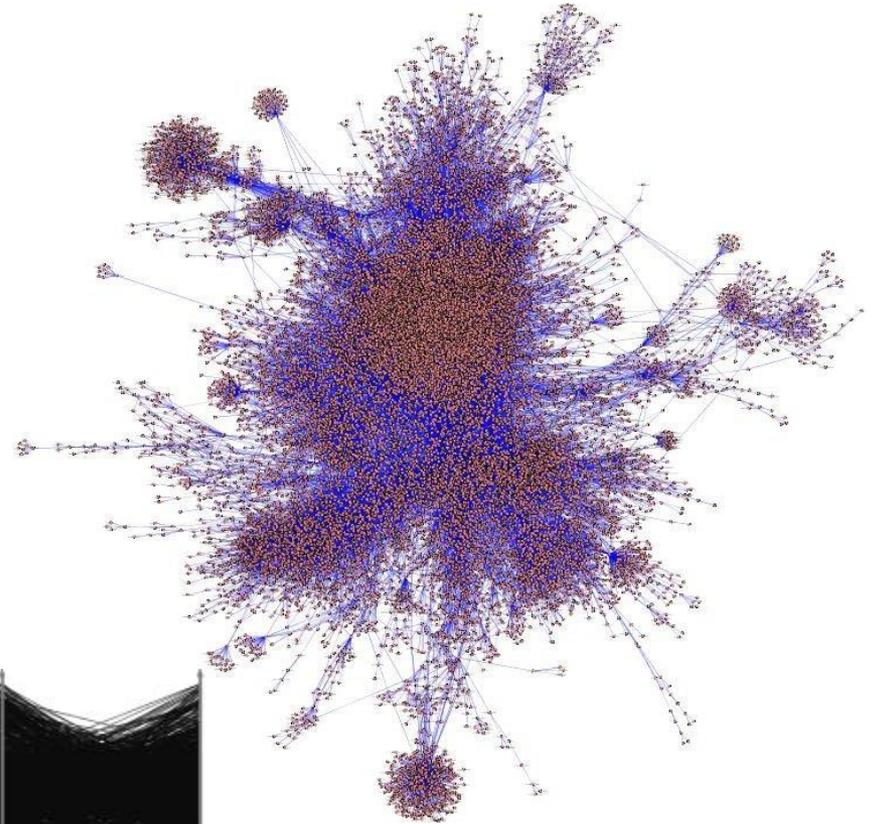
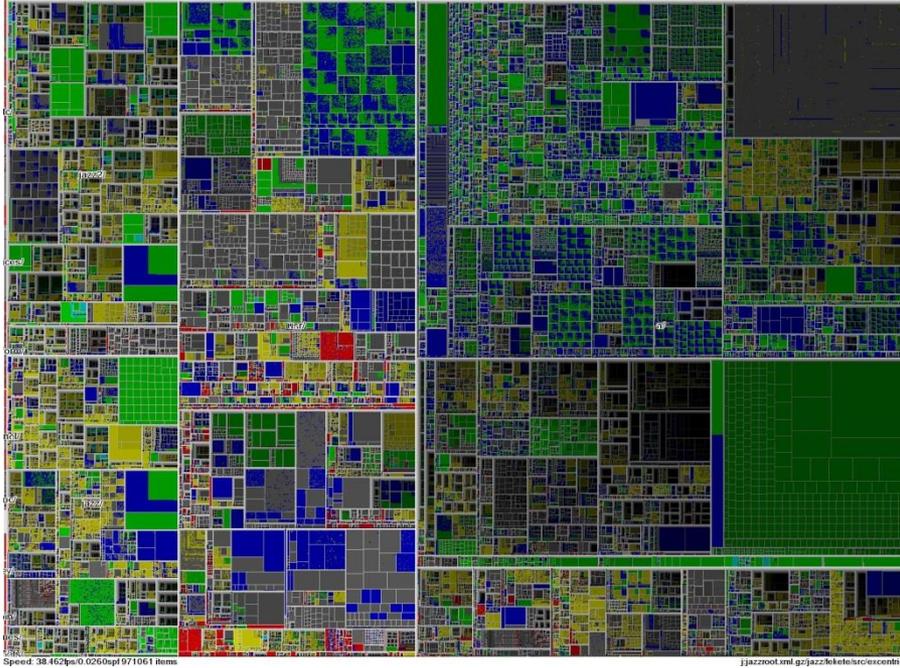
COMPUTER SCIENCE DEPARTMENT
STONY BROOK UNIVERSITY

Lecture	Topic	Projects
1	Intro and logistics	
2	Basic visualizations and tasks, data types, examples, ethical considerations	
3	Data preparation (cleaning, imputation, data set integration)	
4	AI-assisted coding for VIS applications (design, debugging, refactoring)	Project #1 out
5	Big data and data reduction (distance/sim metrics, intro to clustering)	
6	High-D data: concept, subspaces, dimension reduction, PCA	
7	Cluster analysis: hierarchical, density, model, embedding, temporal	
8	Perception and cognition (human visual system, color, contrast)	Project #2(a) out
9	Visual design and aesthetics	
10	Visualization of multivariate and high-D data: linear methods, projections	
11	Vis. of multivariate and high-D data: non-linear methods, embeddings	
12	Visualization and AI: mutual support and capabilities (VIS4AI, AI4VIS)	Project #2(b) out
13	Principles of interaction: drive what is visualized, analyzed & how (HCI4VIS)	
14	Visual analytics (VA), human-centered AI, mixed-initiative system	
15	Midterm #1 (tentative date)	
16	VA system design and evaluation, collaborative VA, uncertainty, provenance	
17	Midterm #1 discussion (tentative date)	Final proj. proposal call out
18	Visualization of hierarchical data	
19	Visualization of maps and data with geo-reference	
20	Visualization of graphs, networks (incl. derivation of causal networks)	Final project proposal due
21	Vis. of time-varying, time-series, streaming data, progressive visualization	
22	Visualization of text, LLMs, and semantic data	
23	Ed Tufte revisited: principles, critiques and limits, responsible visualization	
24	Design of effective infographics	Final proj. prelim report due
25	Foundations scientific and medical visualization, intro to volume rendering	
26	Scientific visualization	Bonus project out (Vol Ren)
27	Story telling with data, data journalism	
28	Midterm #2 (tentative date)	
Final	Final project demo on zoom (public)	All final proj. materials due

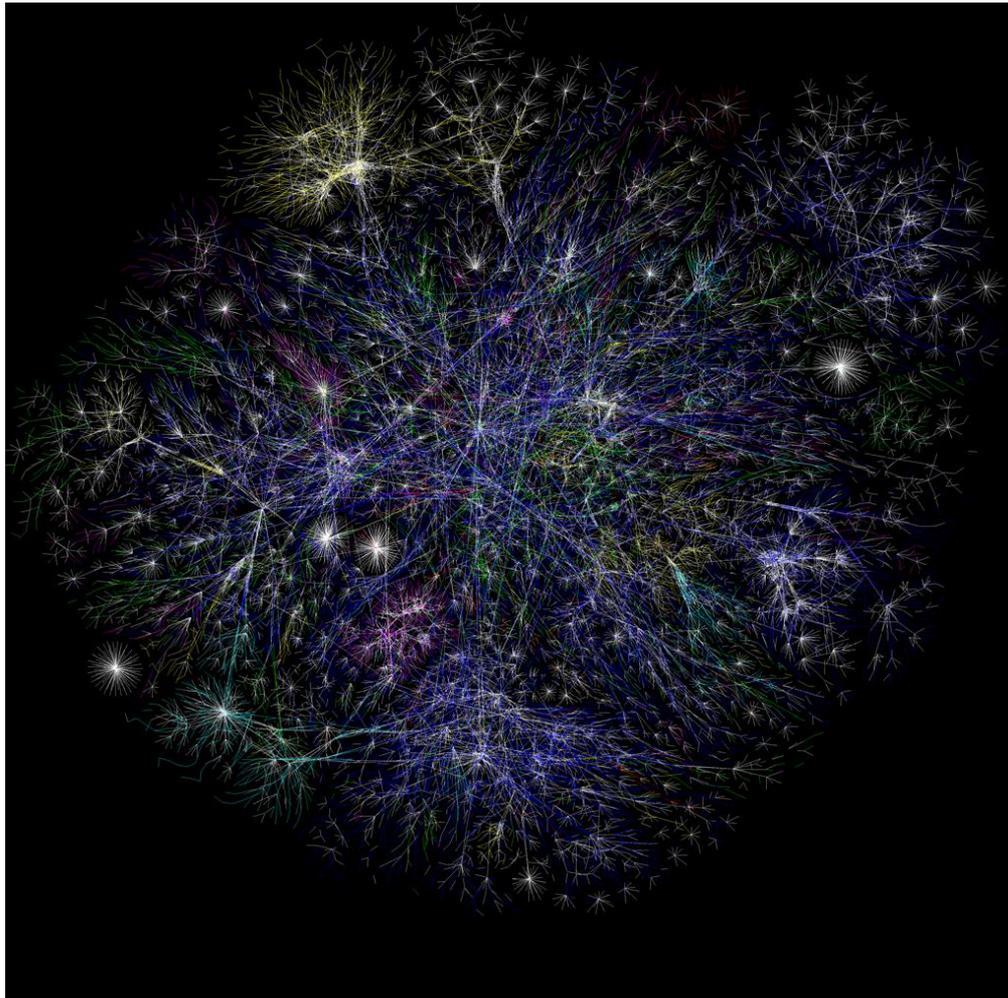
Too Much Data?

How can we deal with data overload

- see the forest for the trees (or the other way around)



Too Much Data?



Internet routes (1/15/2005)

(NY Museum of Modern Art)

The Key to Overcome the Data Deluge: Interaction

Allow users to control what is currently shown:

- level of detail
- extent of the data (spatial, values)
- aspects of the data (attributes)

But do not leave the user lost in the forest

- provide navigation hints

Two powerful paradigms:

- overview, and detail on demand (forest and trees)
- focus and context (trees and forest)

Interaction needs to be interactive (as in responsive)

- user needs get quick visual feedback on actions

Interaction: Key to Visual Analytics

Puts the human in the loop

- appeals to human's expertise and intuition

Requires a suitable human-computer interface

- recall the lectures on color and perception

Interaction can help with:

- making sense of it all
- putting things in proper context
- data overload (scalability)
- telling stories with data (explain findings to others)

Evaluate effectiveness

- do human users actually benefit?
- user studies!

A Taxonomy of Fundamental Interaction Types

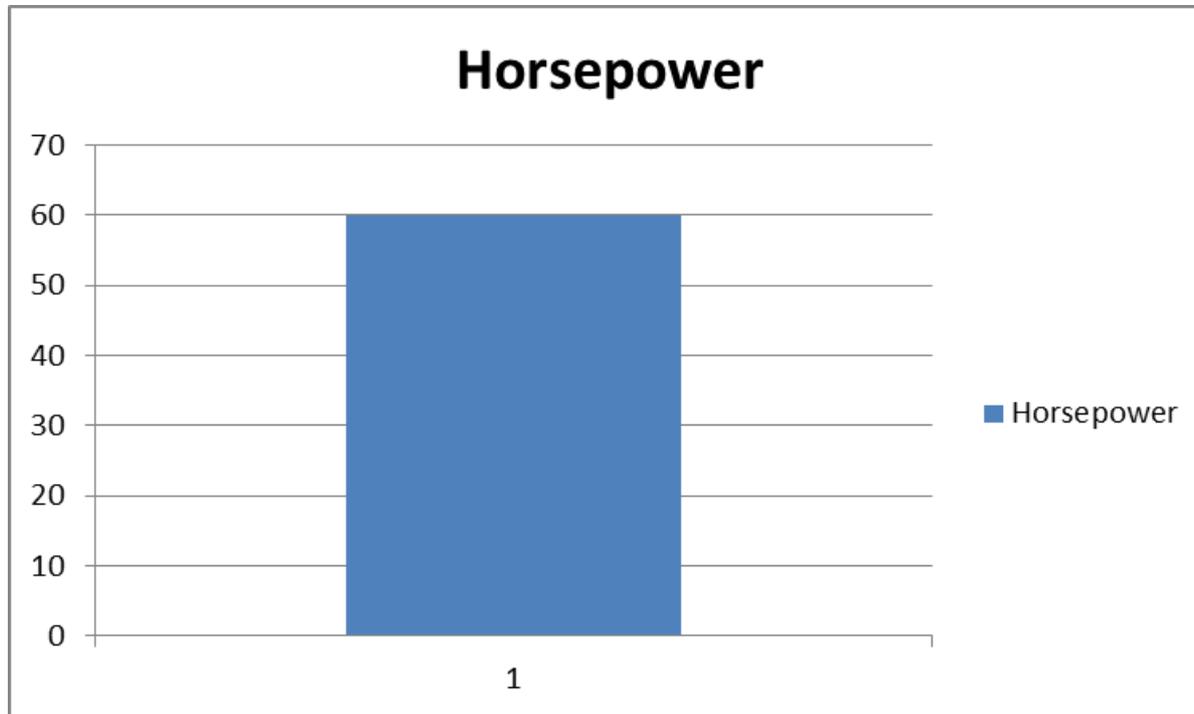
Stephen Few (chapter 4):

- compare
- sort
- add variables
- re-scale
- re-express
- filter
- highlight
- annotate
- bookmark
- aggregate
- re-visualize
- zoom and pan
- details on demand

Example

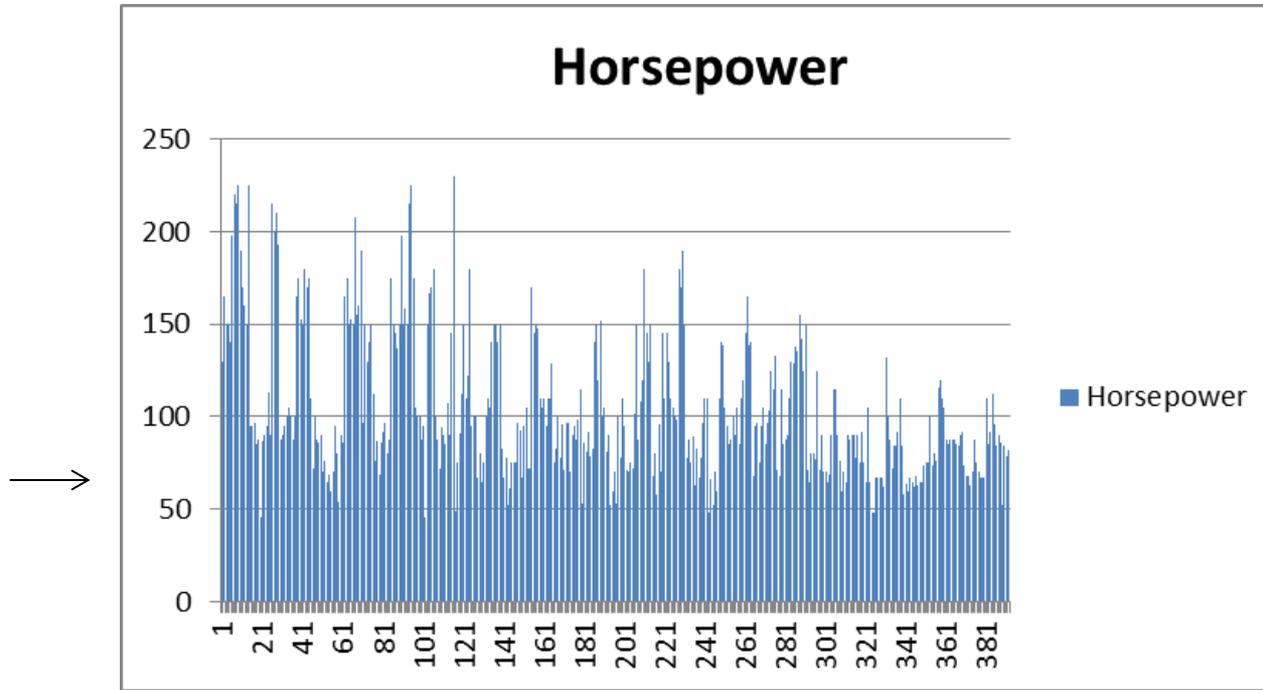
Assume you have been offered a car to buy

- assume you are mostly interested in horsepower, weight, acceleration
- the car you have been offered has 60 hp, 1834 kg, 8 s



Compare

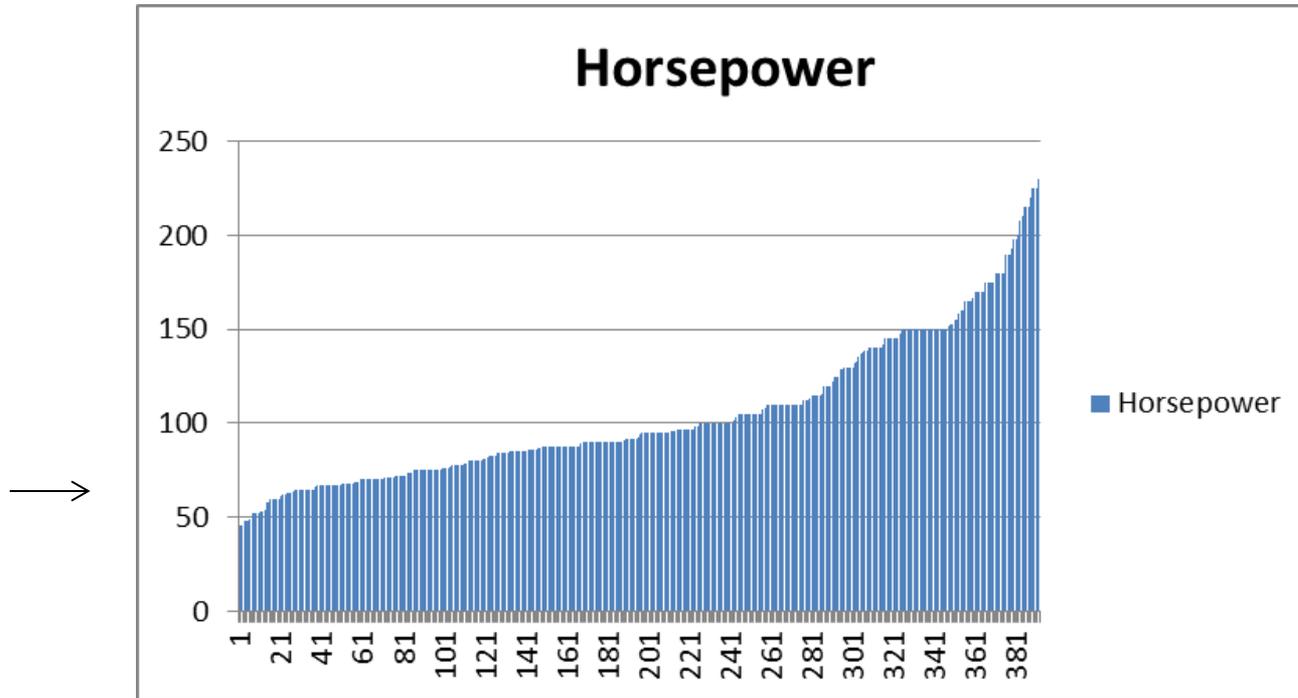
See the car with other available cars



hard to see how it really ranks

Sort

See the car in the context of other available cars

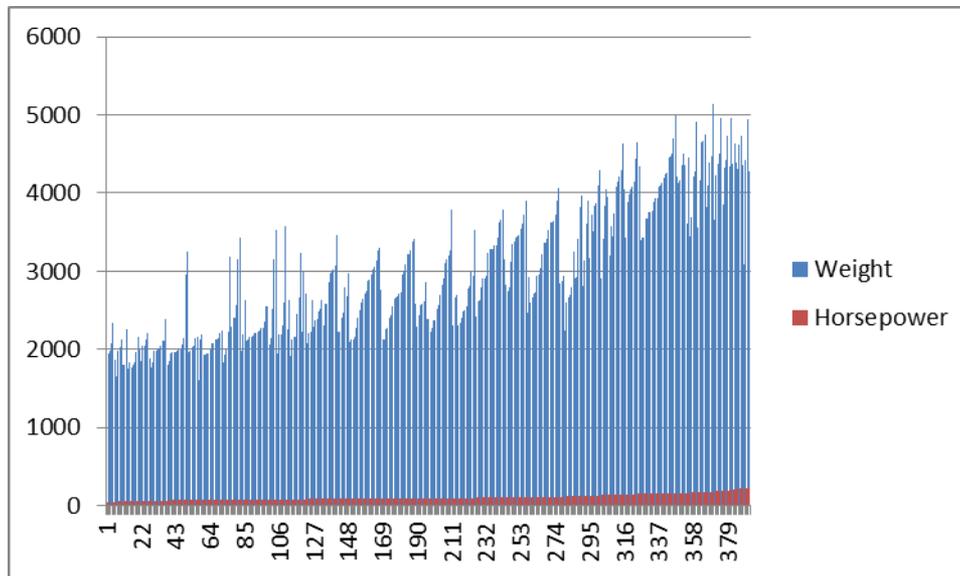


it is a low-horsepower car

Additional Variables

Is horsepower correlated to weight?

- are there trade-offs?

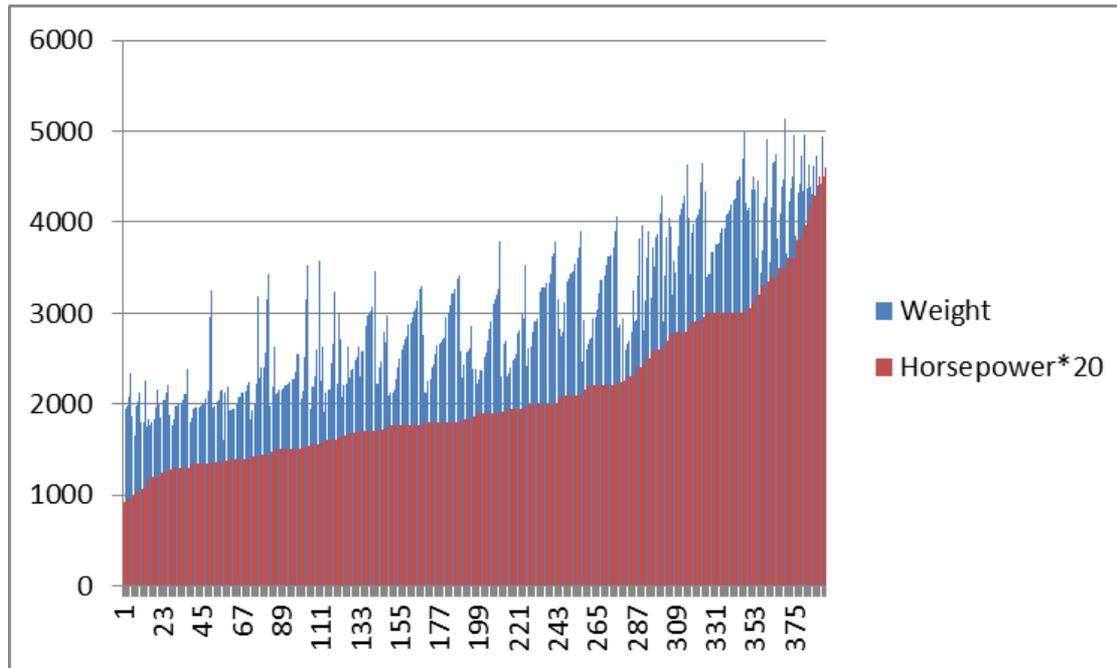


hard to see what is going on

Re-Scale

Scale horsepower into the same range than weight

- could also normalize each to (0.0, 1.0)

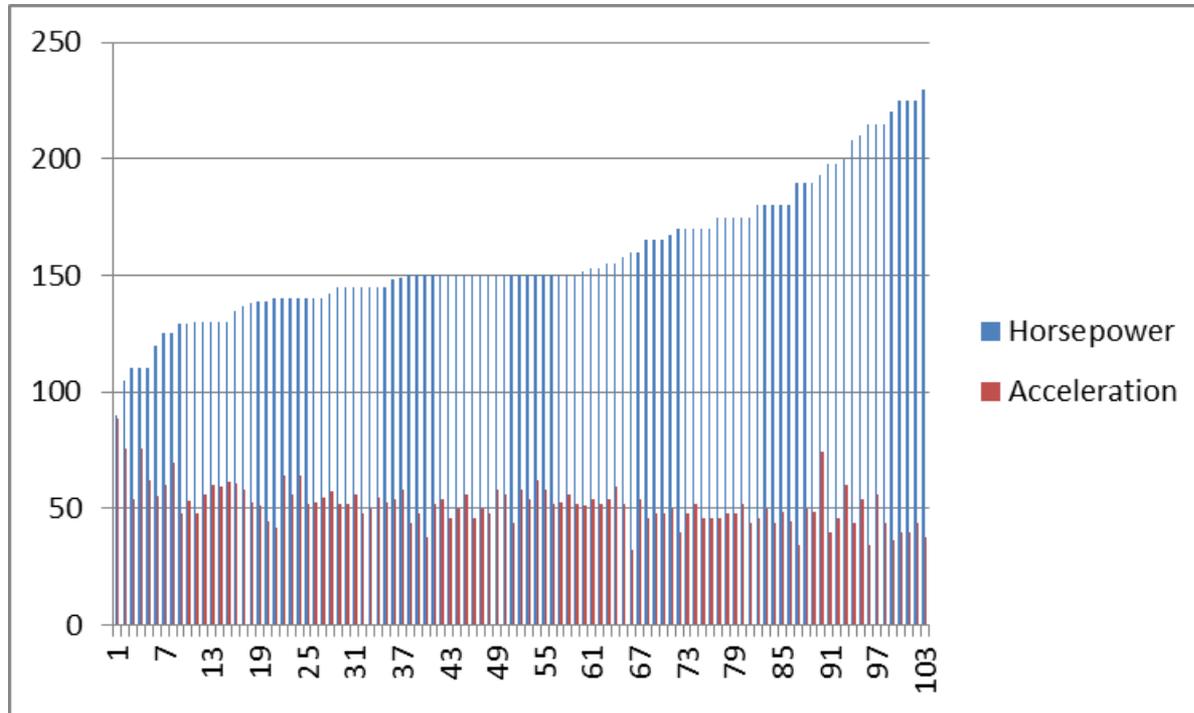


There seems to be a positive correlation

- cars with higher horsepower are also heavier

Another Variable

How does it relate to acceleration?

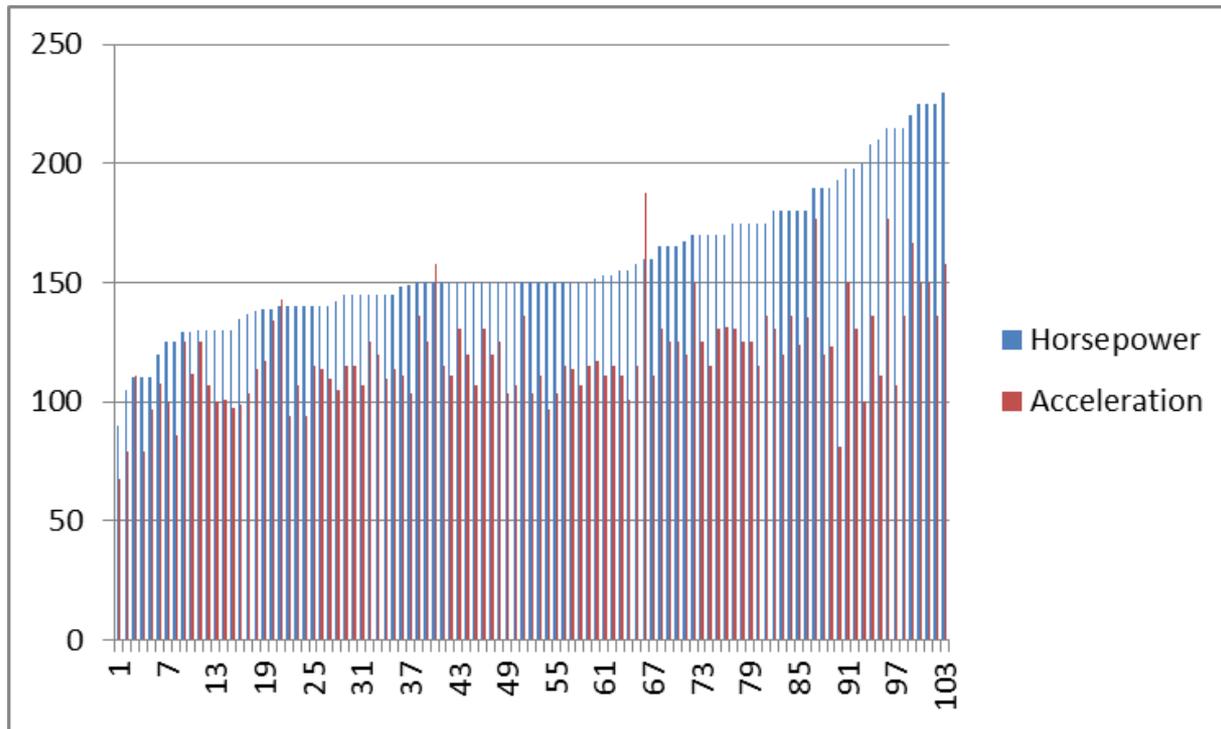


non-intuitive that acceleration is less for high horsepower cars

Re-Express

Acceleration should really be $1/\text{acceleration}$

- should be measured in $1/\text{sec}$ (and not sec)

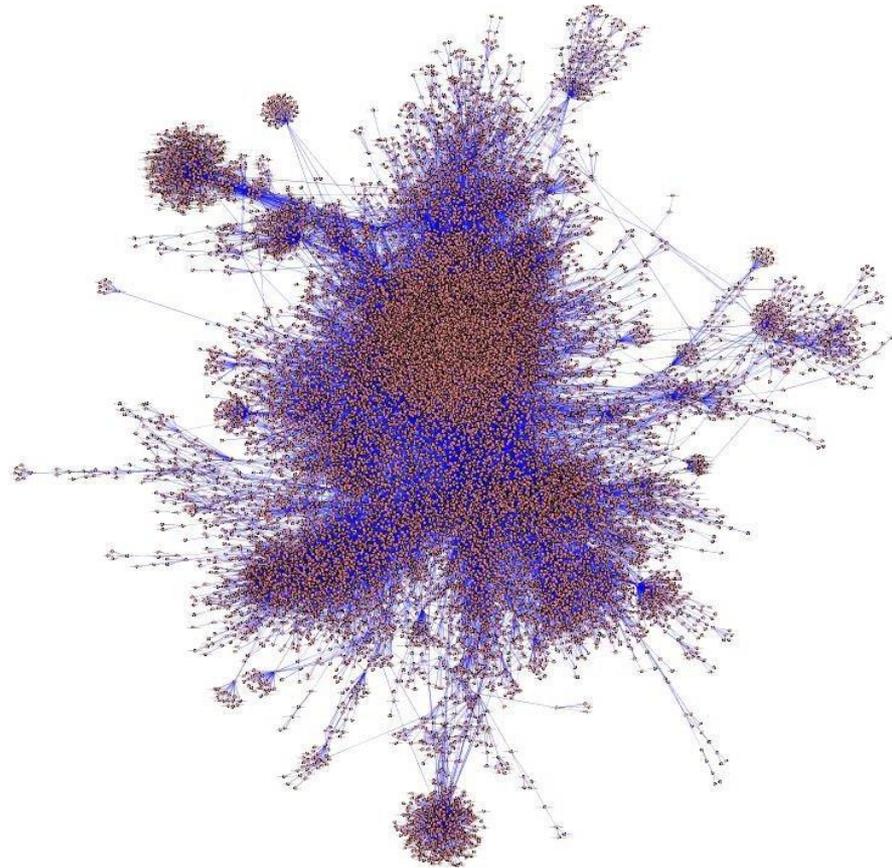


- now higher horsepower cars also seem to have higher acceleration (but the influence is quite minor) → is there a higher-D relationship?

Filtering

Example: Graph of concepts

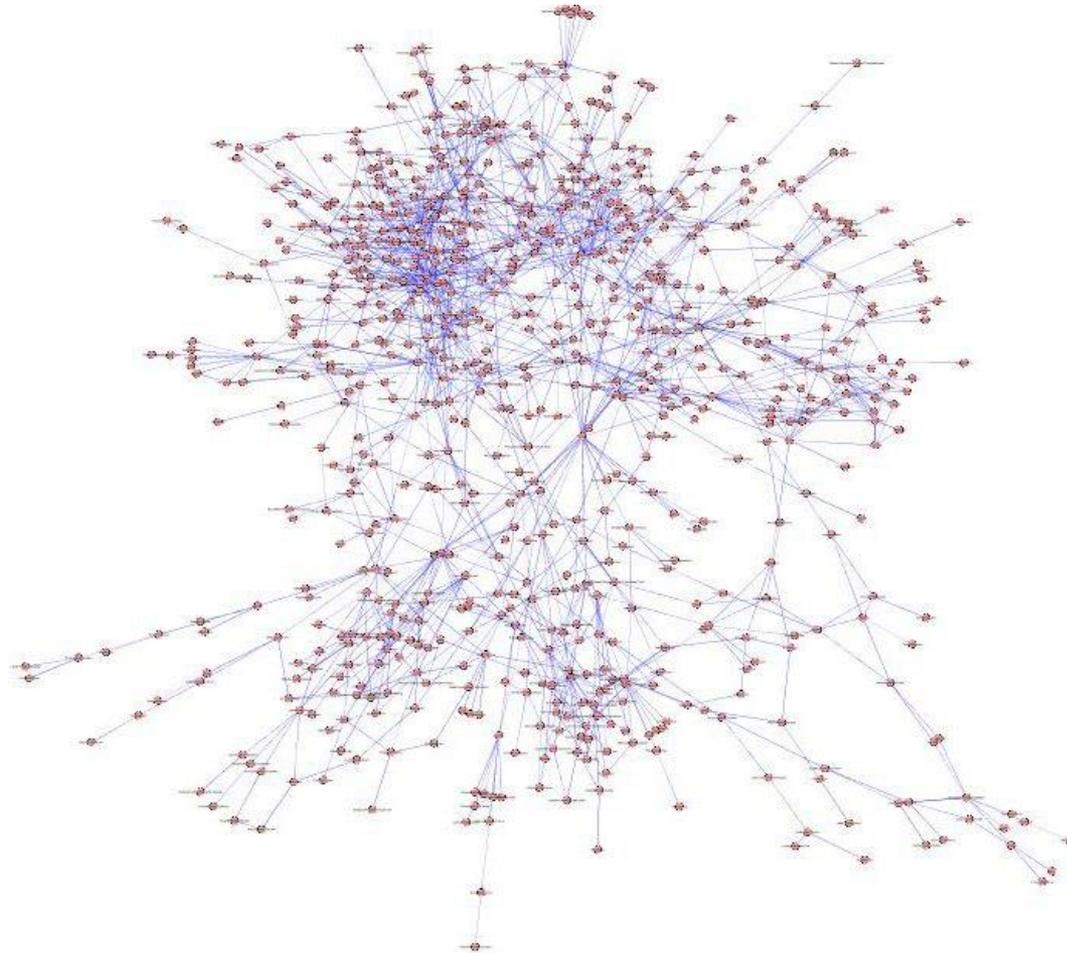
- related concepts group closer



Filtering

Example: Graph of concepts

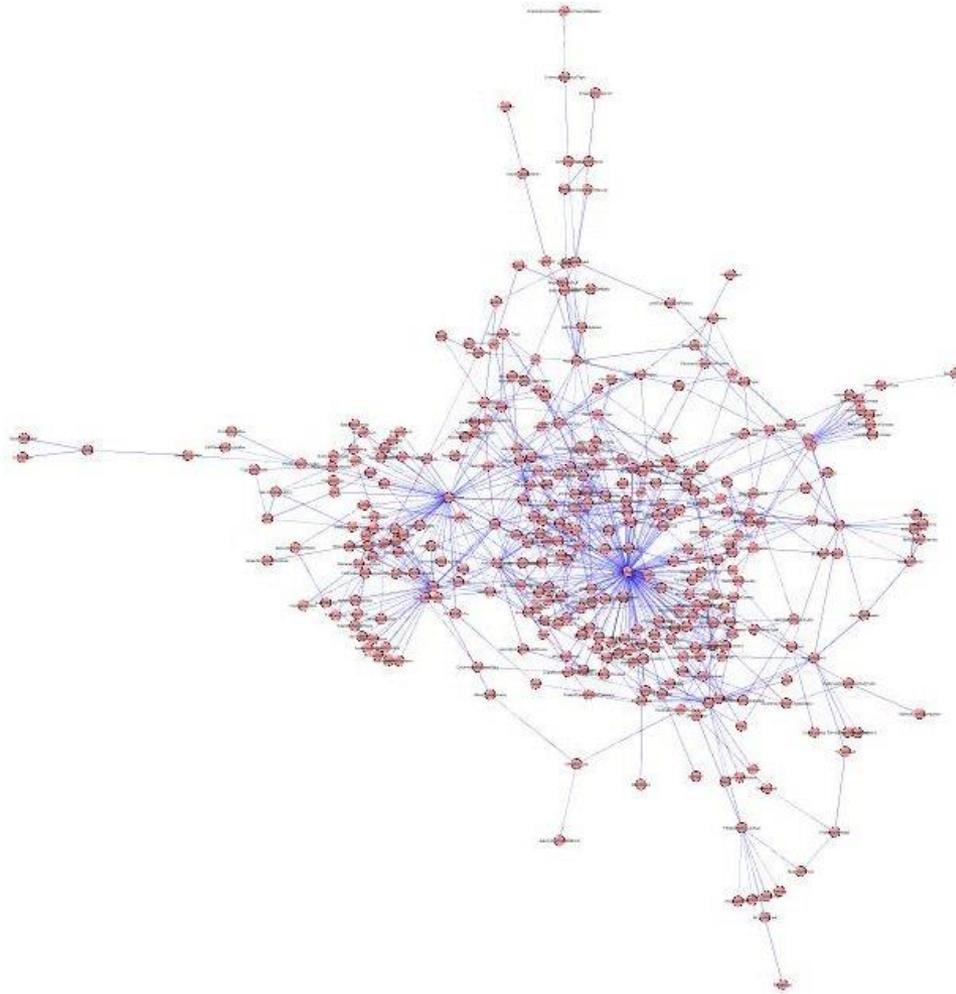
- only keep top 750 connected nodes



Filtering

Example: Graph of concepts

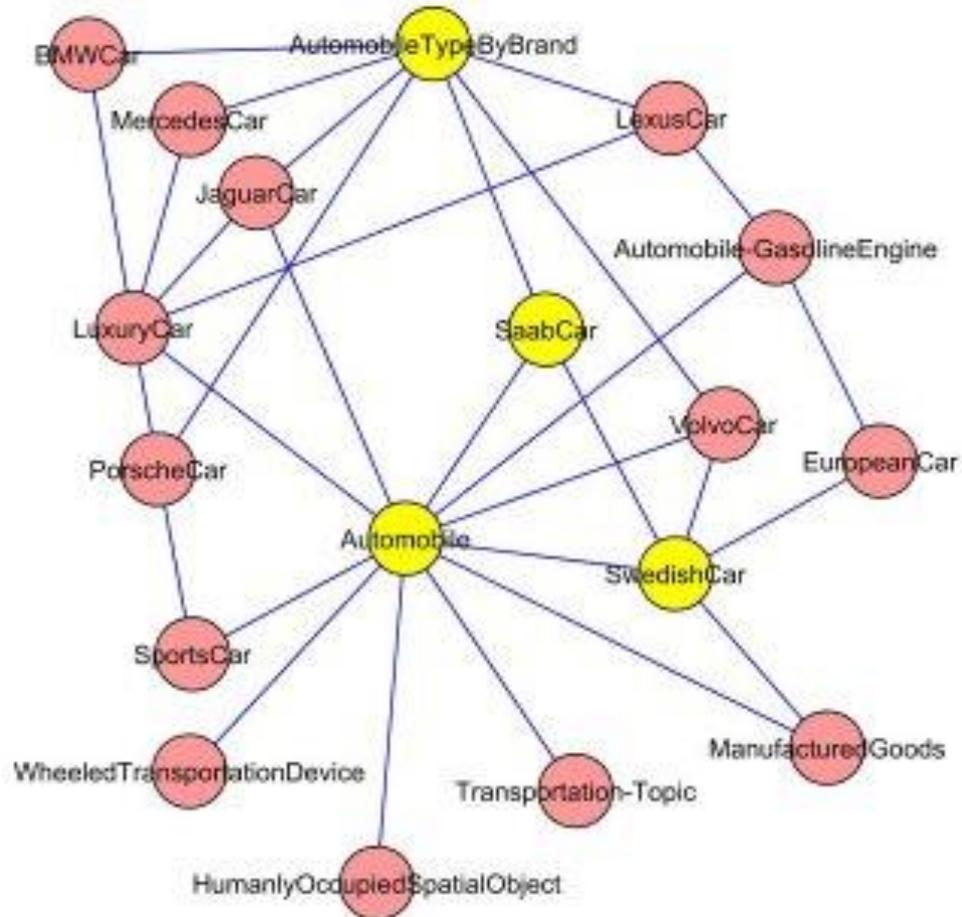
- only keep top 350 connected nodes



Zooming

Example: Graph of concepts

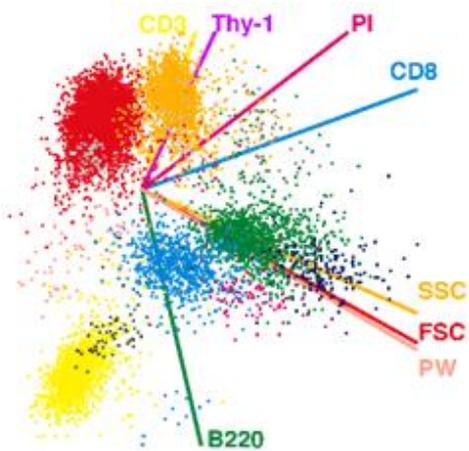
- only keep Saab neighborhood, zoom in more



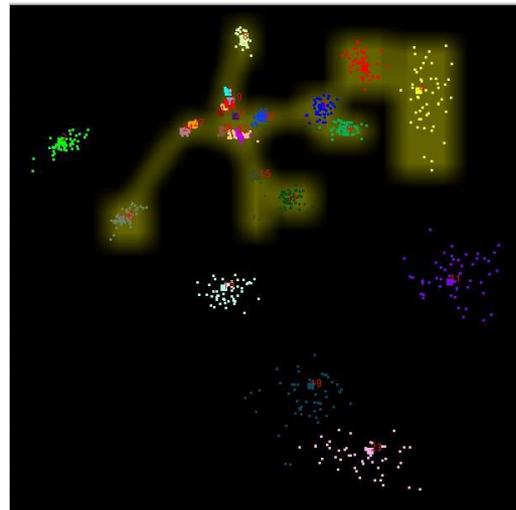
Aggregate

As discussed, good ways to aggregate all data into a single display are:

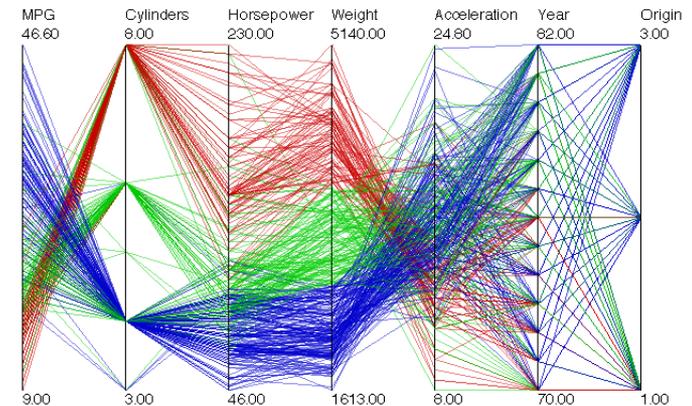
- biplots (project all data into a PCA vector basis)
- multidimensional Scaling (MDS)
- parallel coordinates



biplot



MDS



parallel coordinates

Overview and Detail

The Visual Information-Seeking Mantra

- devised 1996 by Ben Shneiderman (U Maryland, College Park)
- summarizes many visual design guidelines
- in some ways inspired by human vision/behavior
- provides an excellent framework for designing Information visualization applications.



Overview, zoom and filter, then details-on-demand

....

Overview, zoom and filter, then details-on-demand

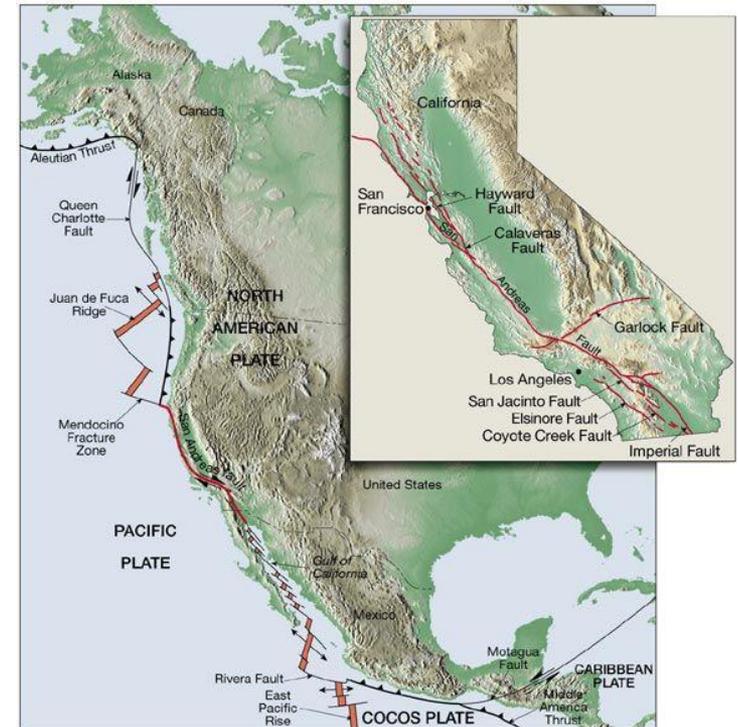
Overview and Detail

Information space overview plus some detail

- maintains (some) context with the detail currently focused on



Leica Microsystems



WikiViz

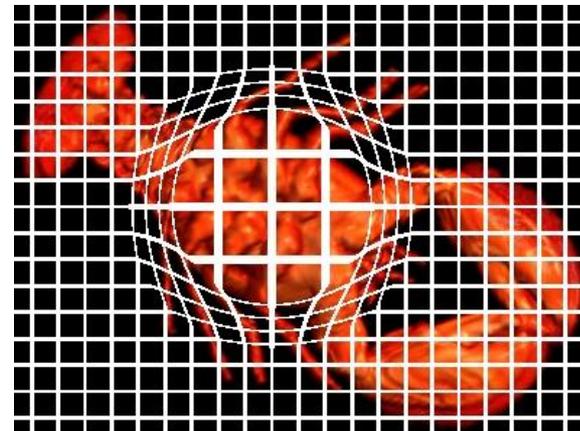
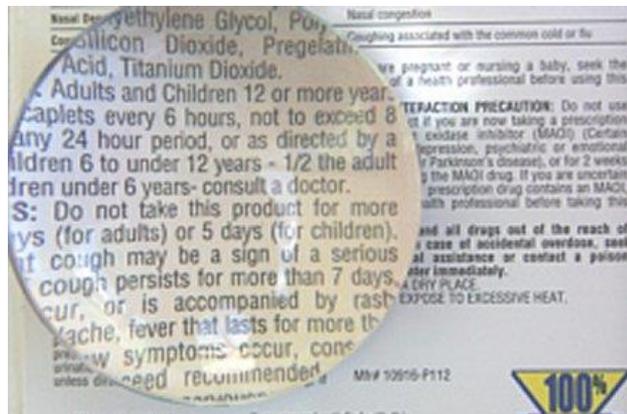
Focus + Context

Overview and detail

- disjoint views, maybe connected by a fan
- but: they simultaneously shows both overview and details
- require the viewer to consciously shift his/her focus of attention

Focus + context

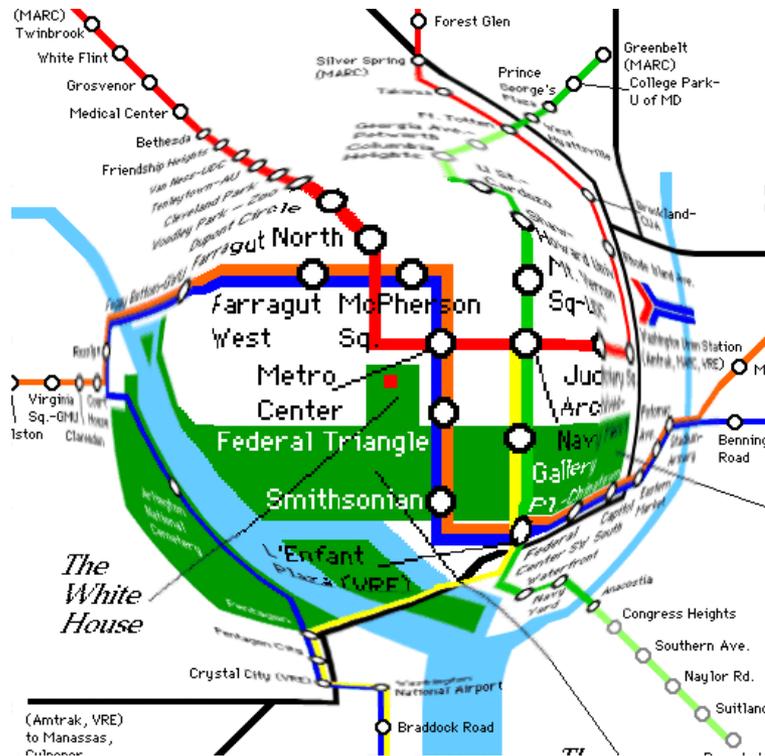
- one single view which shows information in direct context
- maintains continuity across the display
- do not require viewer to shift back and forth
- a good F+C paradigm is the *lens*
- but: there will be distortion



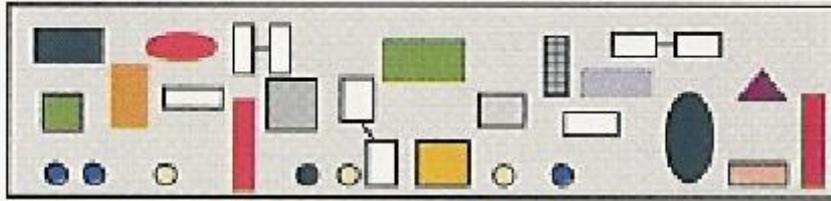
Fisheye Lenses

Fisheye lenses

- physically correct and therefore familiar
- keep target item in focus
- less relevant peripheral items are dropped or reduces in size
- distortion

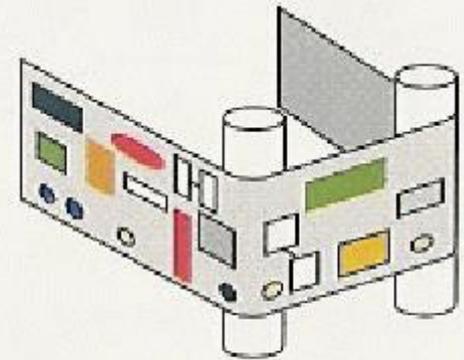


Bifocal Lens

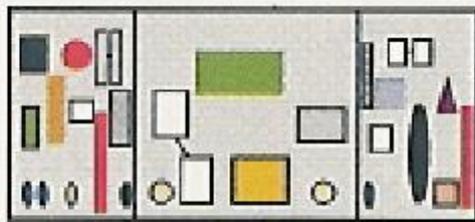


Complete Mapped Information Space

fold



Principle of the Bifocal Display



Bifocal Display Seen by the User

project

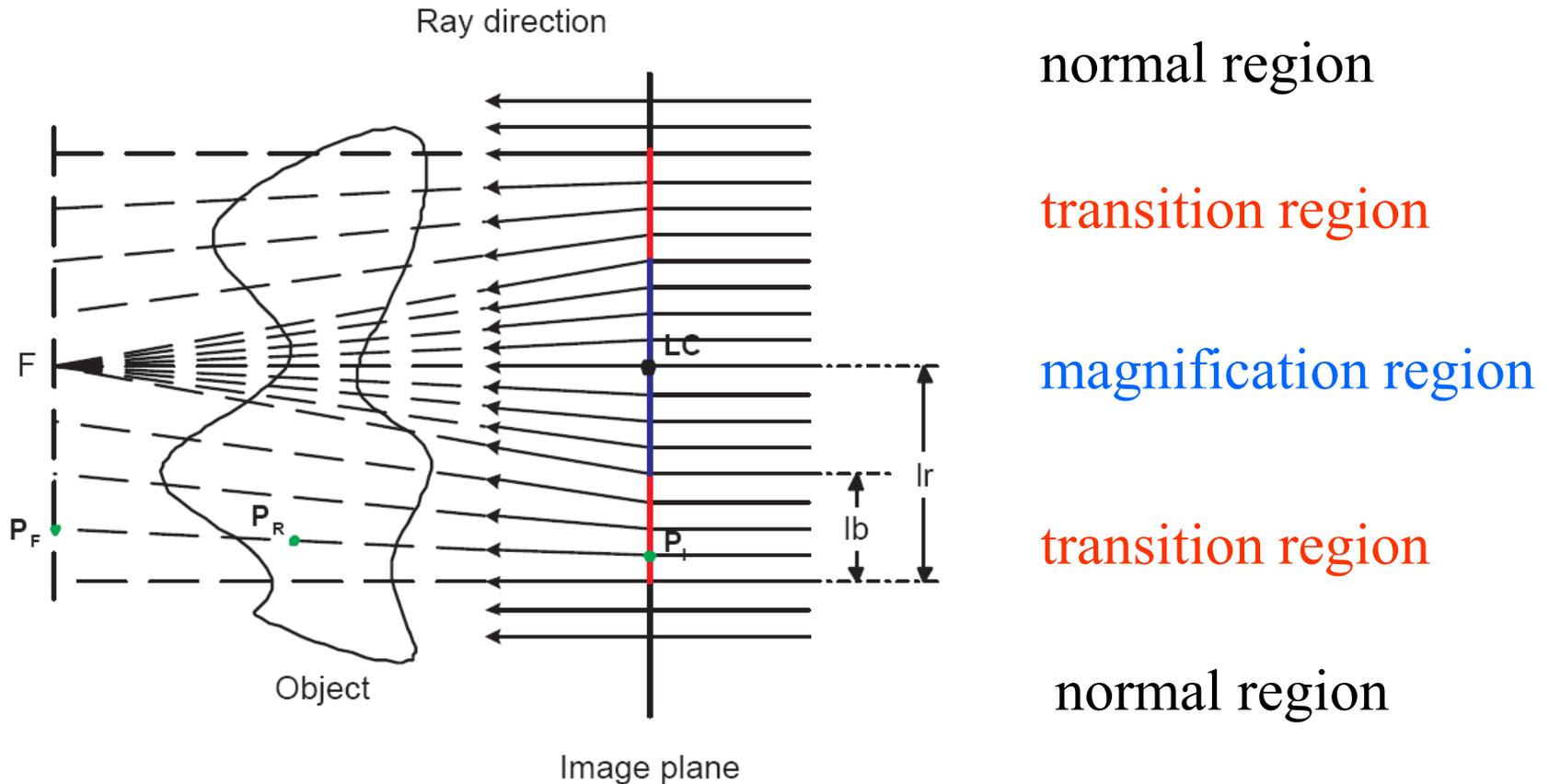


Bifocal Lens

London subway map



(Volumetric) Magnification Lenses

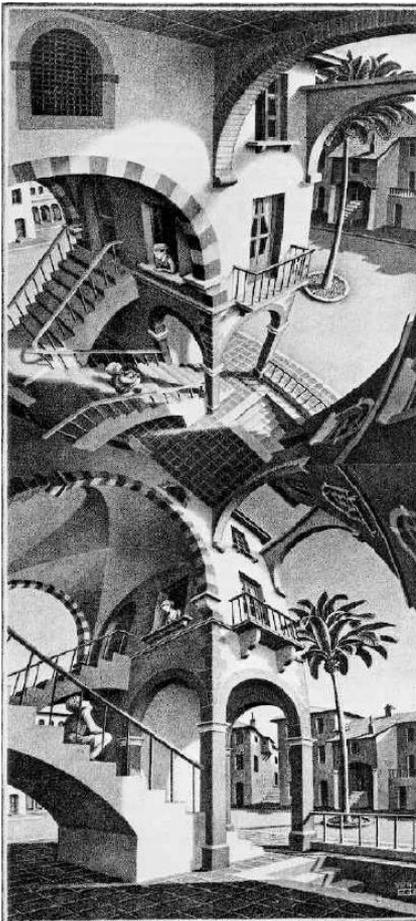


Avoid aliasing in transition regions by low-pass filtering

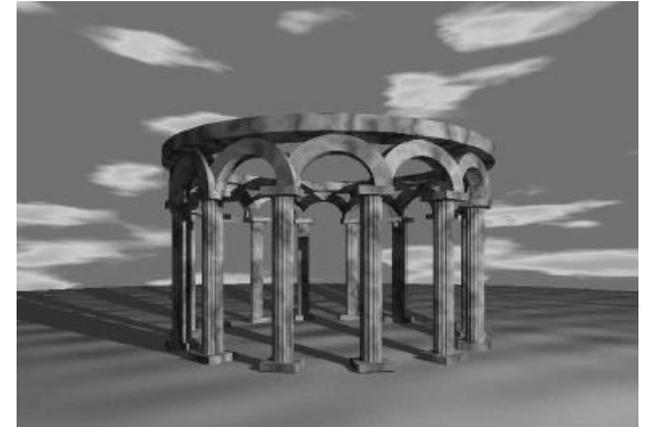
Generalized Lenses

Computers can go beyond (stretch) the laws of physics

- example: multi-perspective lens rendering, gaze-directed, ...



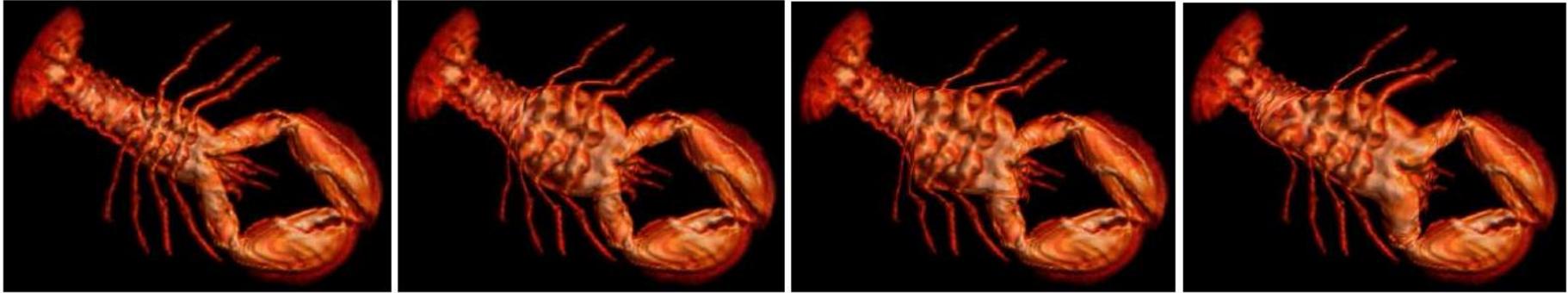
Rademacher/Bishop



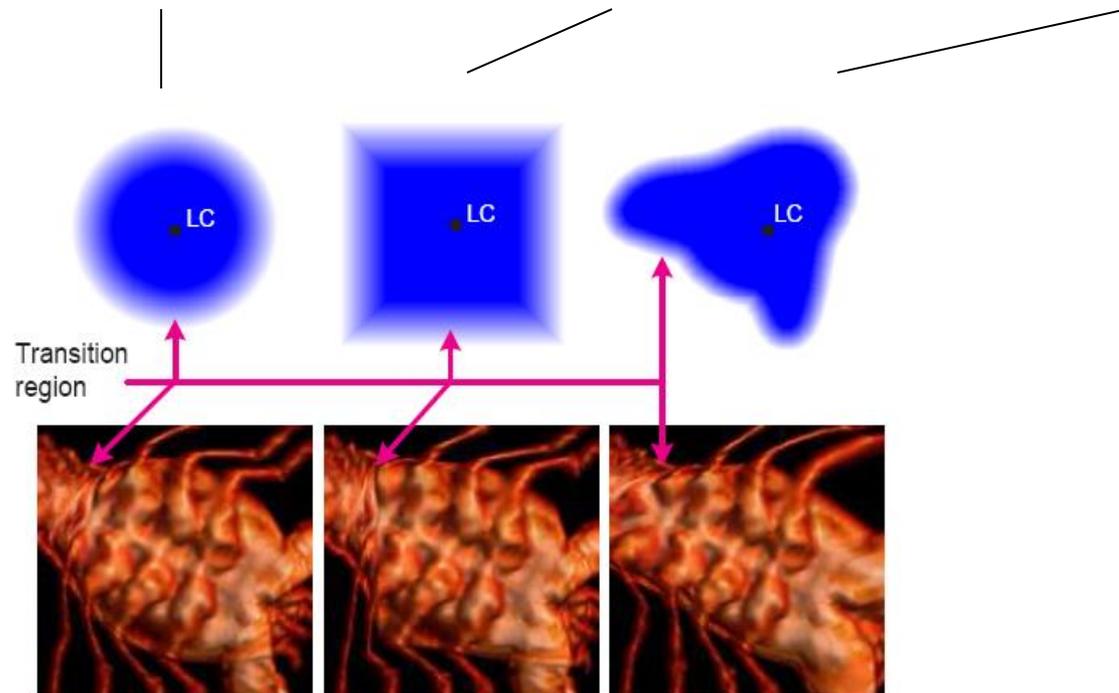
MC Escher

Loeffelmann/Groeller

Generalized Lenses



no lens

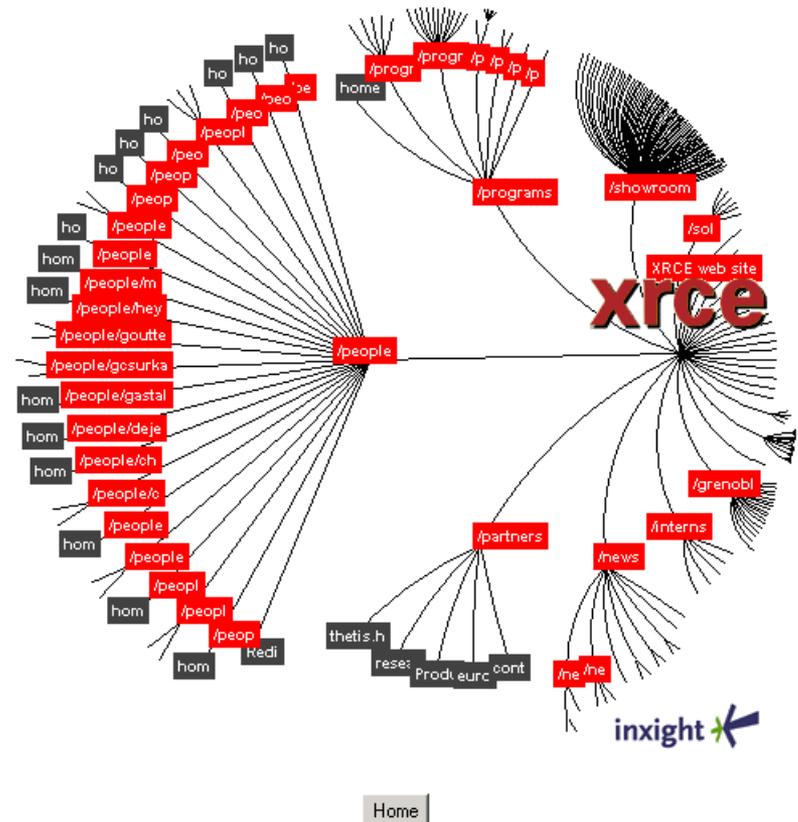
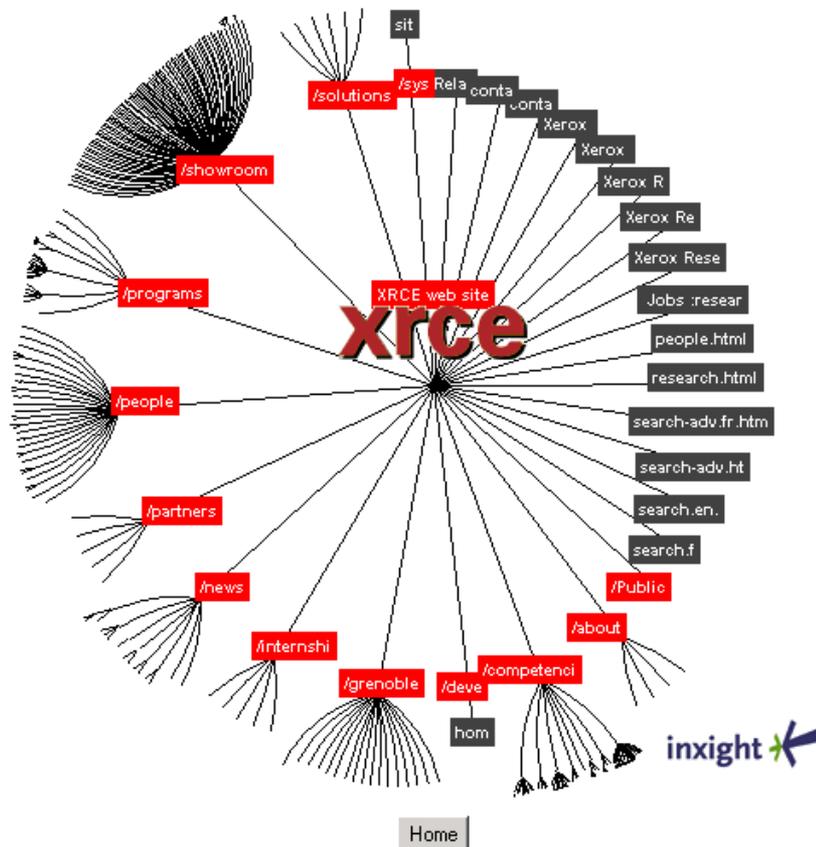


Wang et al., 2005

Lenses in Information Visualization

Hyperbolic Tree fisheye lens

- Xerox PARC/Inxight
- selectively and smoothly reduce complexity as user focus changes



Lenses in Information Visualization

Table Lens (Rao and Card, 1994)

- uses a DOI (degree of interest) lens
- fuses symbolic and graphical detail driven by the DOI lens

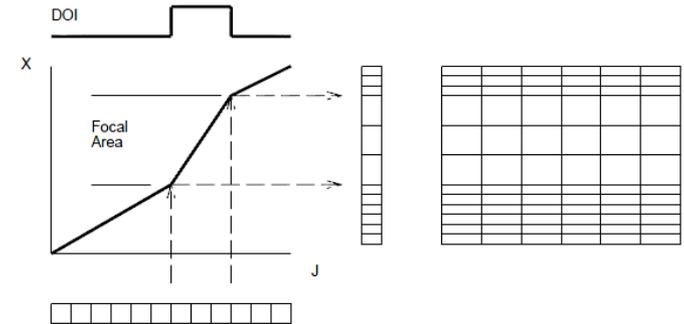


Table Lens: Baseball Player Statistics

Calculate: "Hits" / "At Bats" = "Avg"

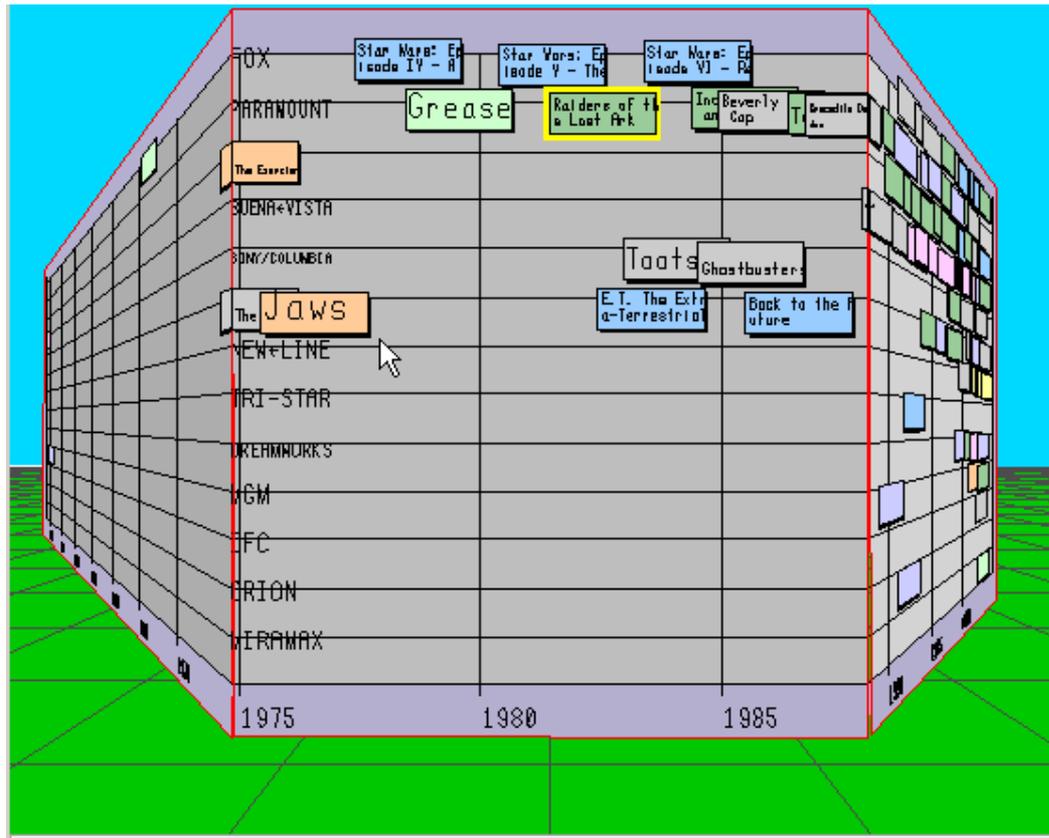
	Avg	Career Avg	Team	Salary 87
Larry Herndon	0.24734983	0.27282876	Det.	225
Jesse Barfield	0.2886248	0.27268818	Tor.	1237.5
Jeffrey Leonar	0.27859238	0.27260458	S. F.	900
Donnie Hill	0.28318584	0.2725564	Oak.	275
Billy Sample	0.285	0.2718601	Atl.	NA
Howard Johnson	0.24545455	0.25232068	N. Y.	297.5
Andres Thomas	0.250774	0.2521994	Atl.	75
Billy Hatcher	0.25775656	0.25211507	Hou.	110
Omar Moreno	0.2339833	0.2518029	Atl.	NA
Darnell Coles	0.2725528	0.25153375	Det.	105

Row 304: Mike Lavalliere; Column 20: Put Outs Value: 468 810 -- 2163

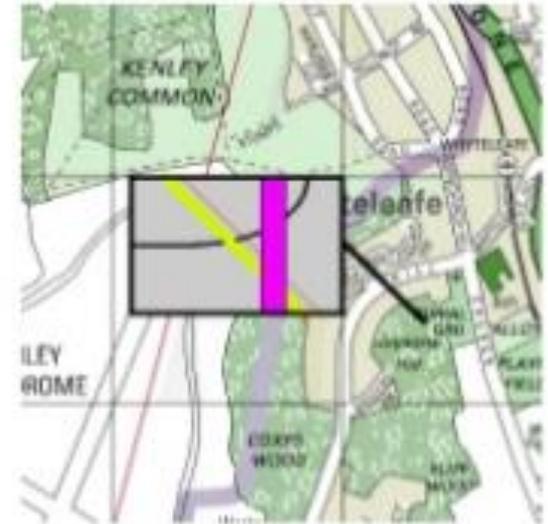
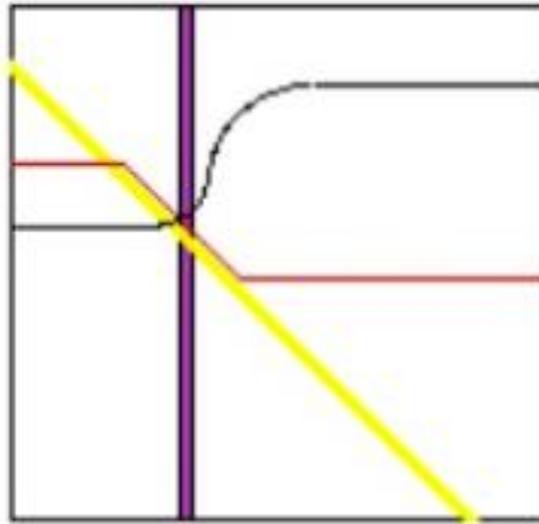
Lenses in Information Visualization

Perspective Wall

- Xerox PARC/Inxight
- details on the center panel are at least three times larger than the details on a flat wall that fits the field of view



Magic Lens



Illustrating the concept of a magic lens. (a) shows a conventional map of an area, (b) shows the location of services (gas, water and electricity pipes) in the same area, and (c) a (movable) magic lens shows services in an area of interest, in context

[Video](#)

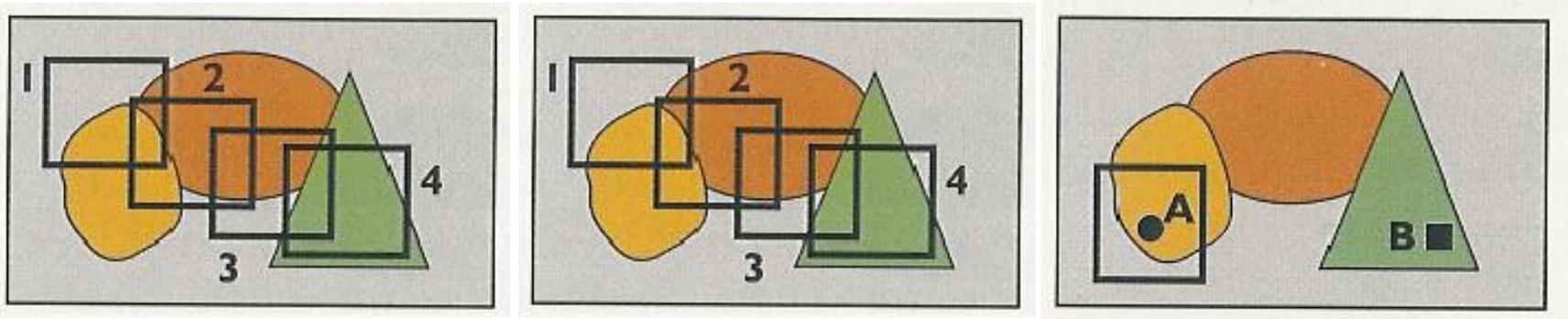
Zoom and Pan

Panning

- smooth movement of a viewing frame over a 2D image of greater size
- Zooming
 - increasing magnification of a decreasing fraction (or vice-versa) of a 2D image under the constraint of a viewing frame of constant size

Transfer of the focus of attention:

- zoom out → pan → zoom in

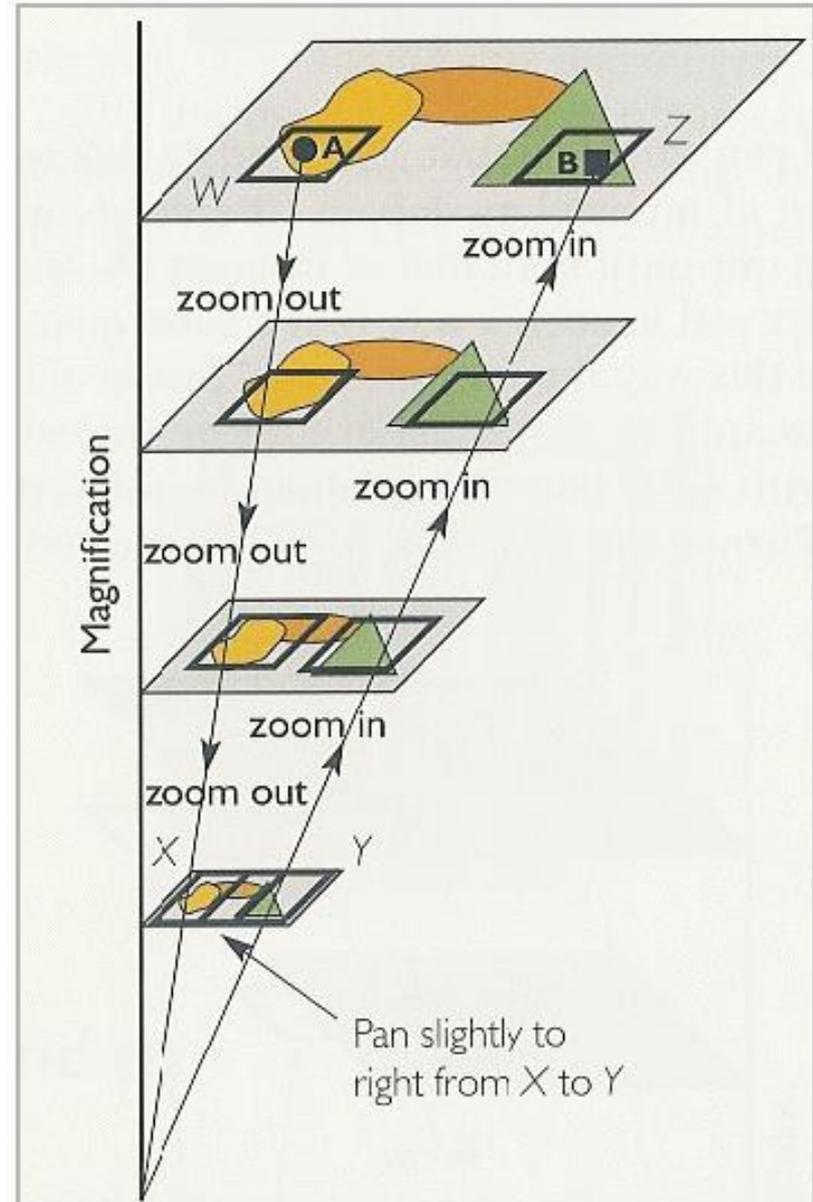


Scale-Space Diagrams

Efficient transfer of the focus of attention:

- zoom out → pan → zoom in

Furnas, Bederson, 1995



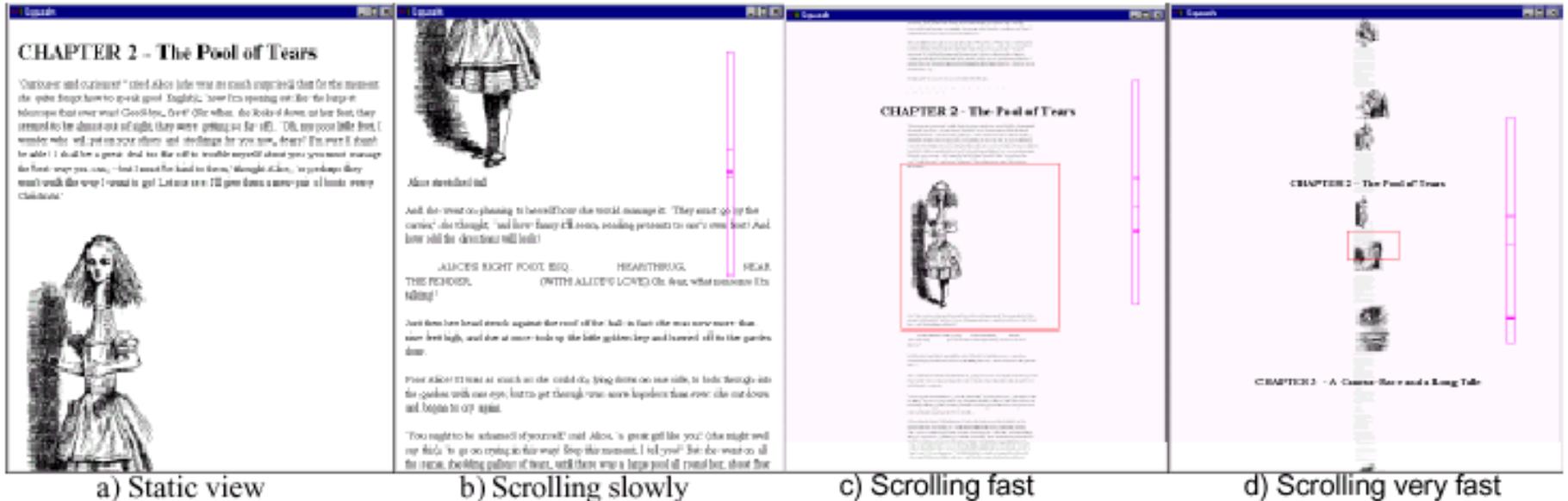
Scale-Space Diagrams Application

[video on youtube](#)

Intelligent Zooming

Depending on scrolling speed, zoom more or less

- allows efficient navigation of large documents
- employs semantic zooming



Igarashi, Hinckley, 2000

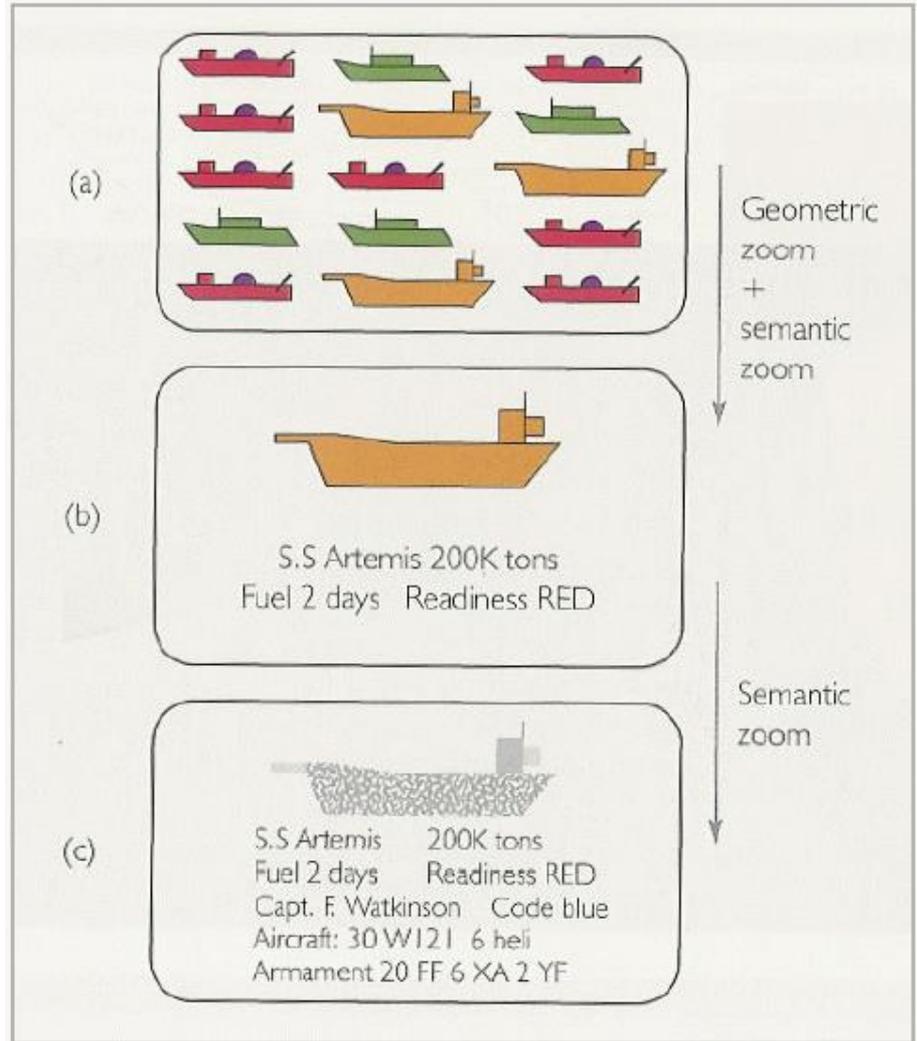
Semantic Zoom

Standard zoom:

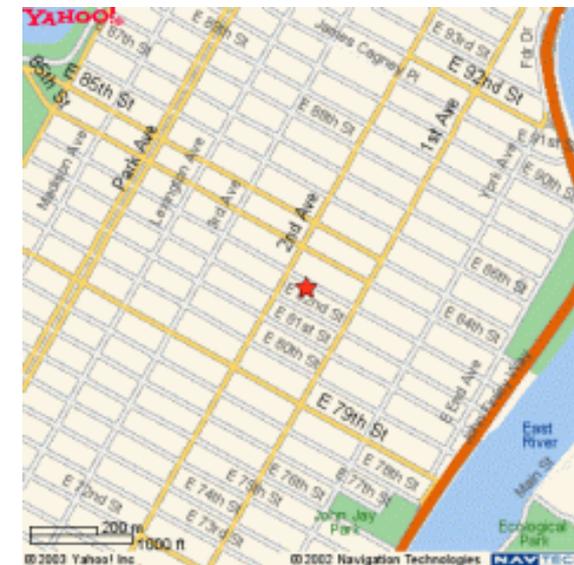
shows a down/up scaled version of the object/image

Semantic zoom:

- shows a different representation determined by the space available



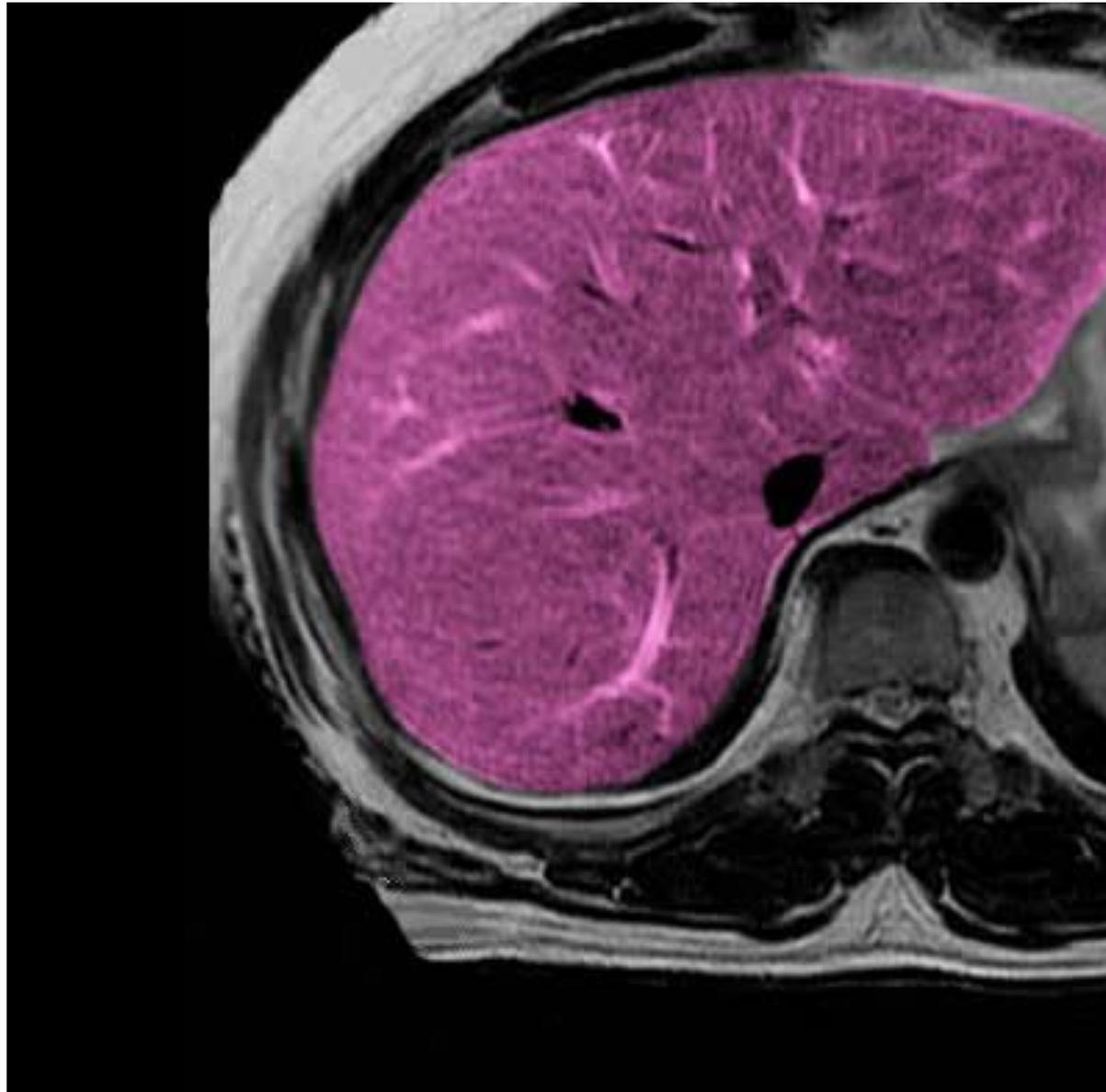
Semantic Zooms: Maps



Semantic Zoom: The Infinite Microscope

L. Wang, K. Mueller, “Enhancing volumetric datasets with sub-resolution detail using texture synthesis,” IEEE Visualization Conference, 2004

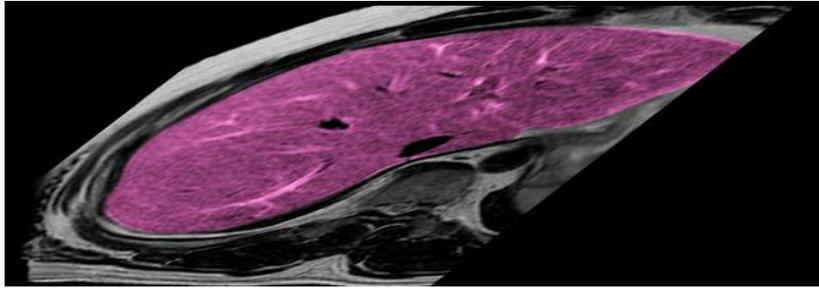
Semantic Zoom: The Finite Microscope



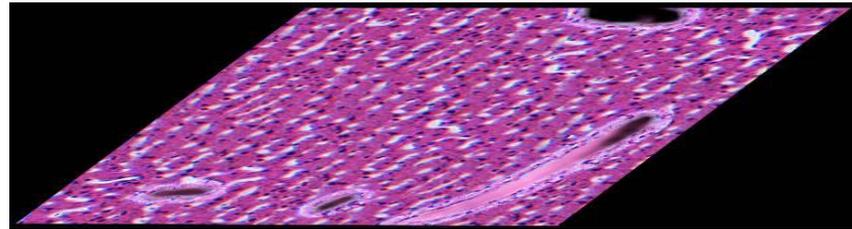
Semantic Zoom: The Multi-Sample Microscope

What to do normally, in this case:

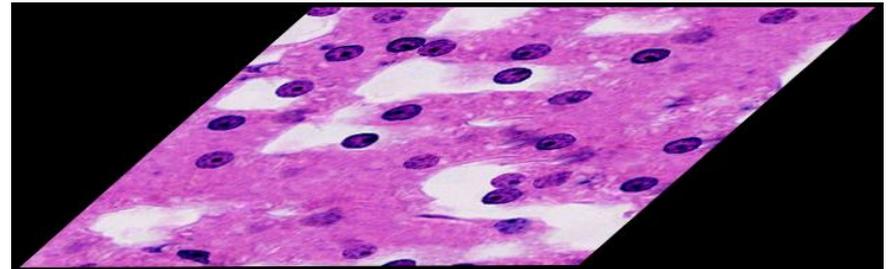
- switch to the dataset acquired with suitable modality (higher resolution)



level 1



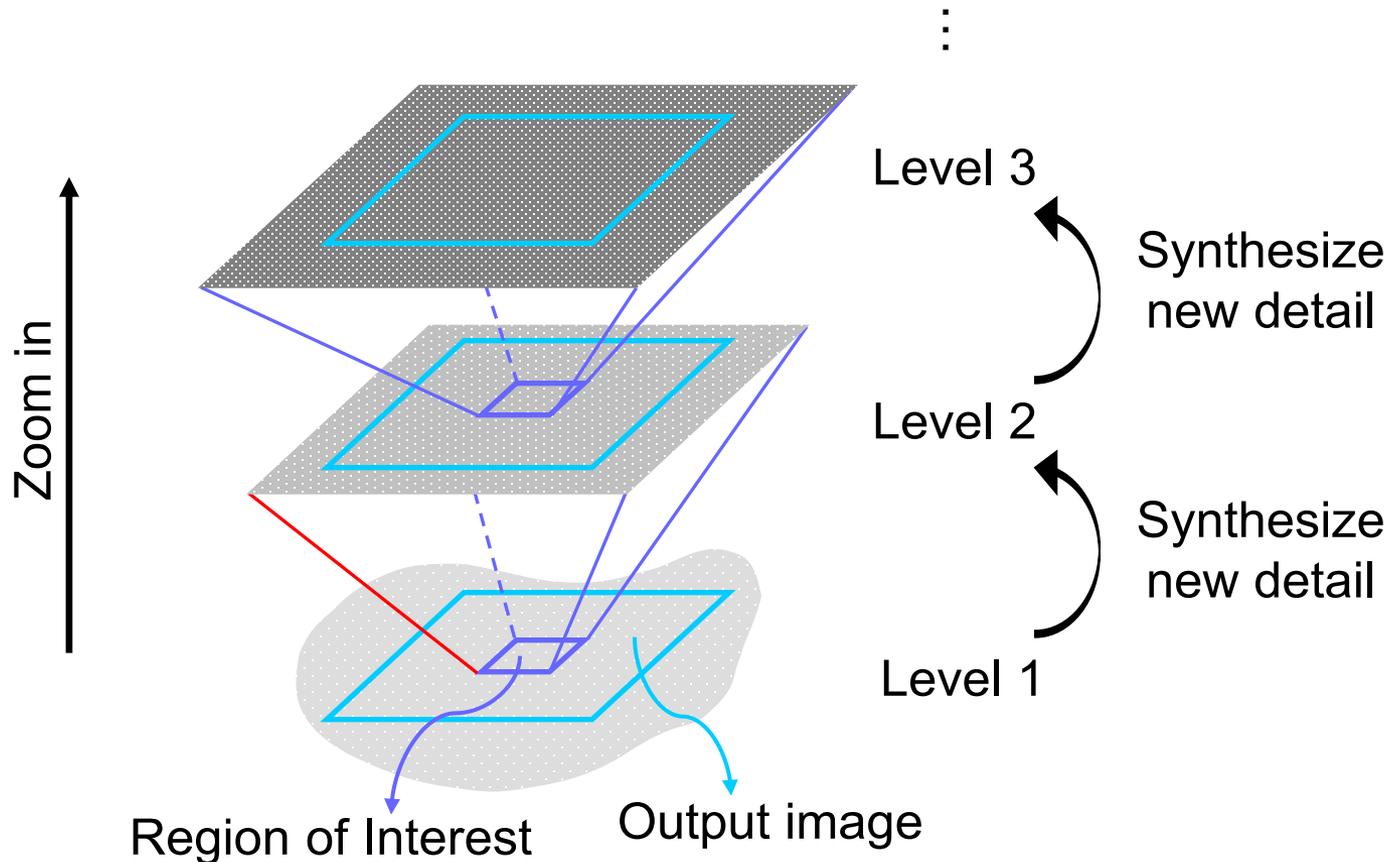
level 2



level 3

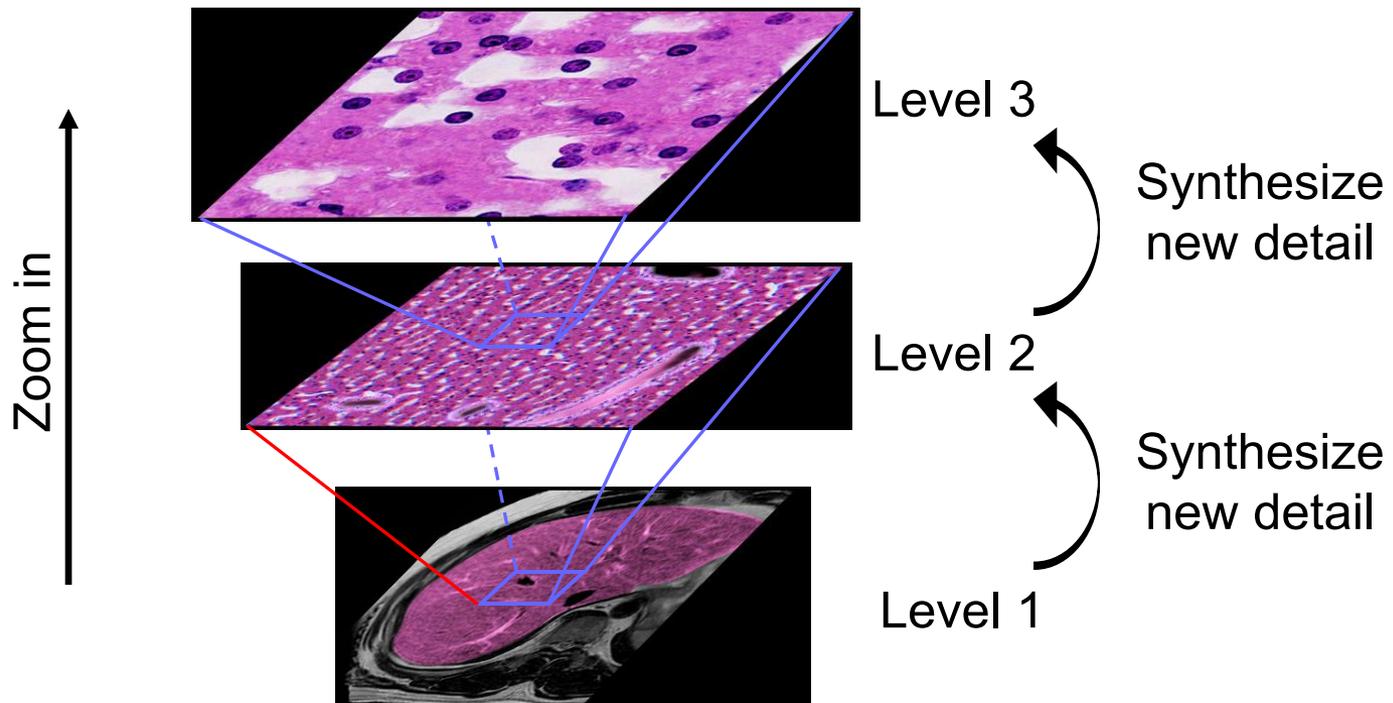
Semantic Zoom: The Infinite Microscope

Apply constrained multi-scale texture synthesis to facilitate semantic zooms



Semantic Zoom: The Infinite Microscope

Apply constrained multi-scale texture synthesis to facilitate semantic zooms

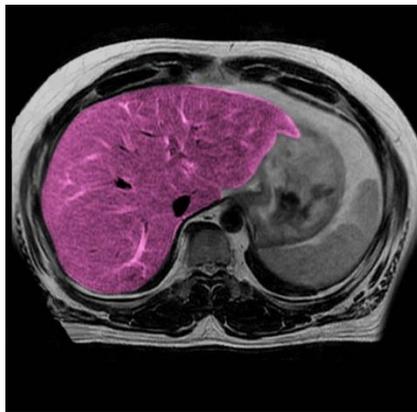


Semantic Zoom: The Infinite Microscope

Data collection (sample images)

Colorization

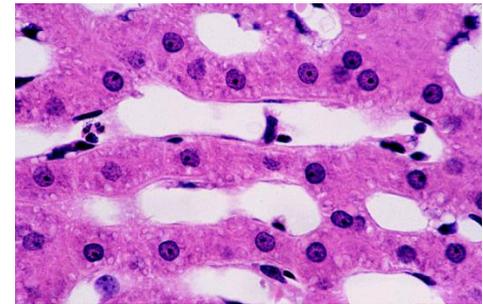
- reduce the distinct discontinuities during zoom



MRI level



histology level



cell level

Sample Images

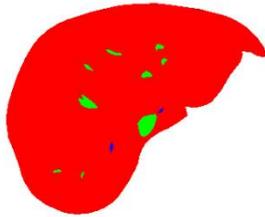
Semantic Zoom: The Infinite Microscope

Data collection (sample images)

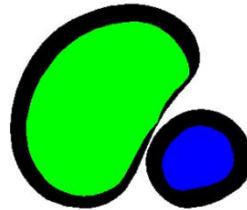
Colorization

- reduce the distinct discontinuities during zoom

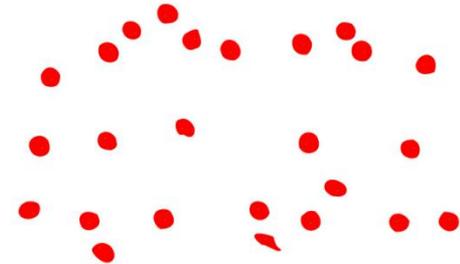
Segmentation (to enable constrained synthesis)



MRI level



histology level



cell level

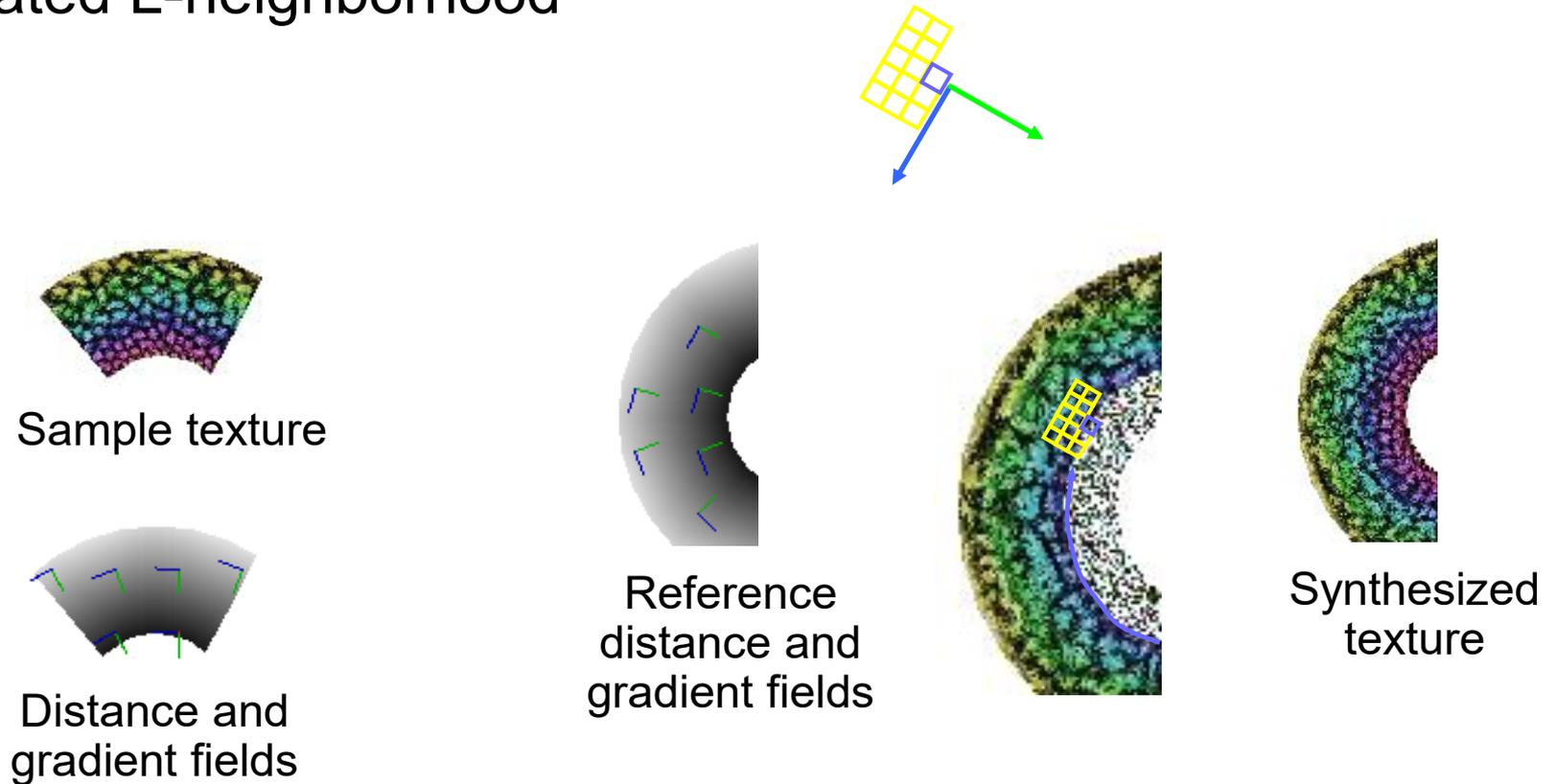
Tag Images

Constrained Synthesis

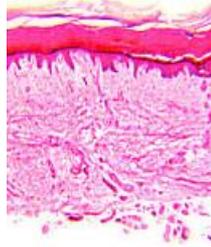
Standard L-neighborhood will not work

- distance and shape dependencies exist

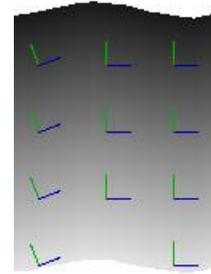
Rotated L-neighborhood



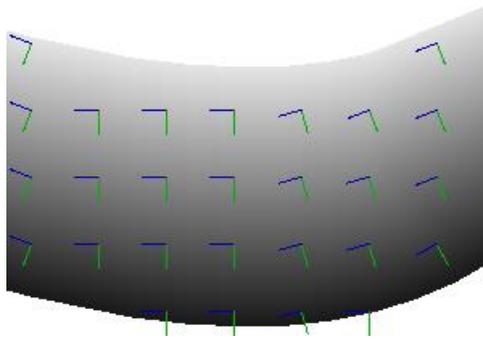
Example: Skin Histology



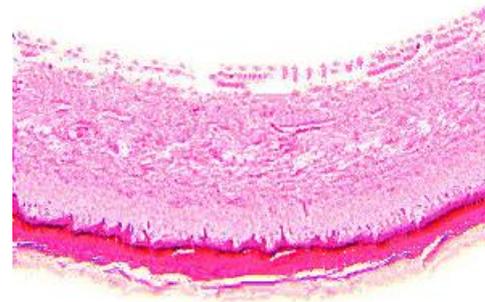
Sample image



Distance and
gradient fields



Reference
distance and
gradient fields



Synthesized skin
histology image

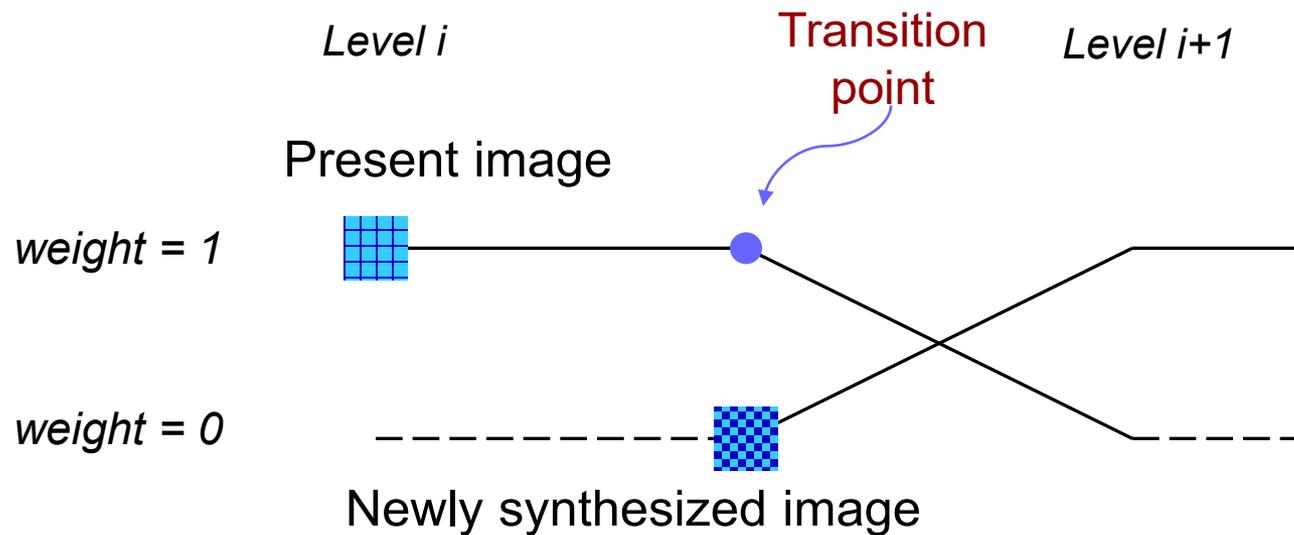
Smooth Semantic Zooms

Magnify low resolution level image

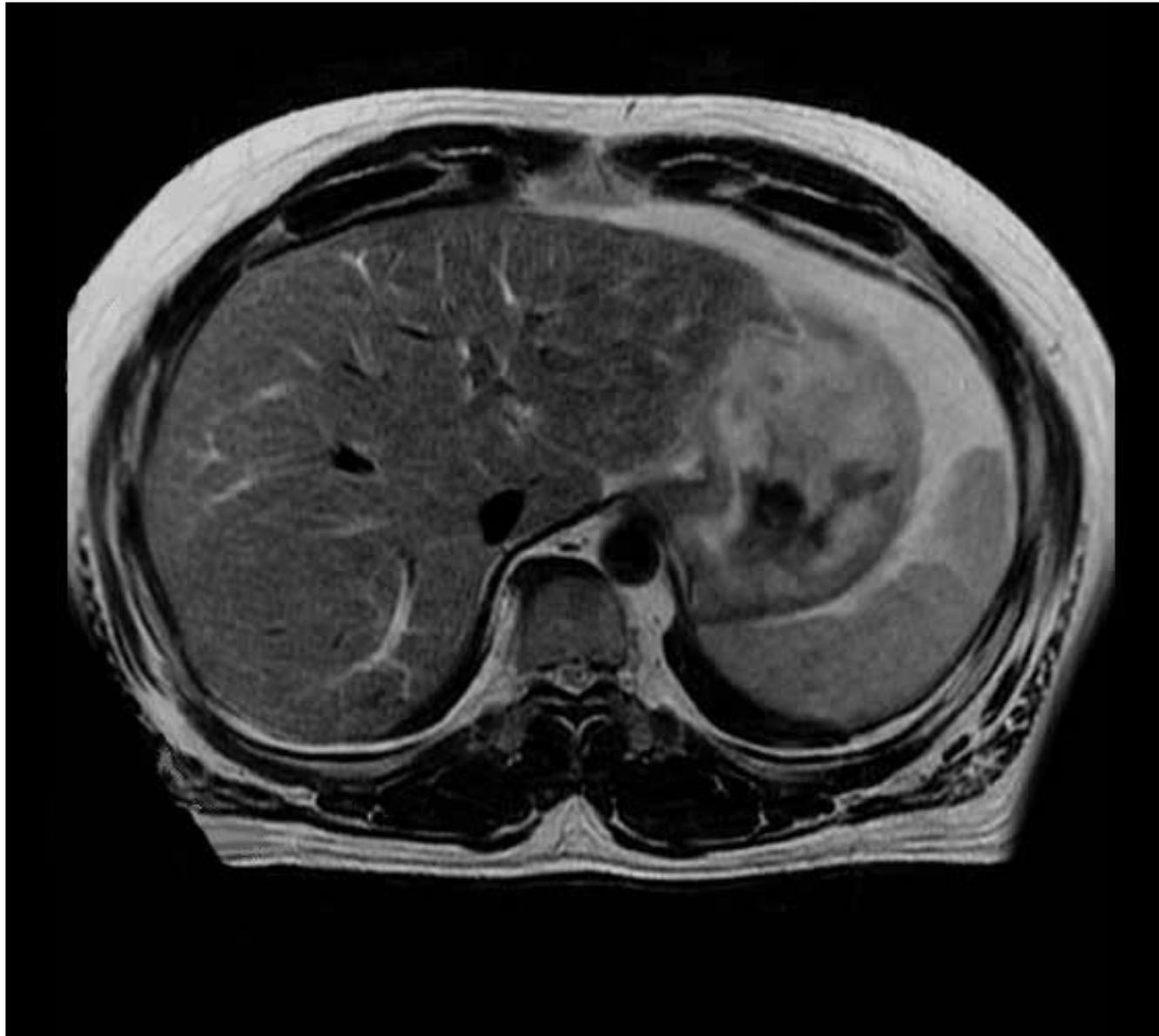
Synthesize new detail at transition point

Minify synthesized next high level image

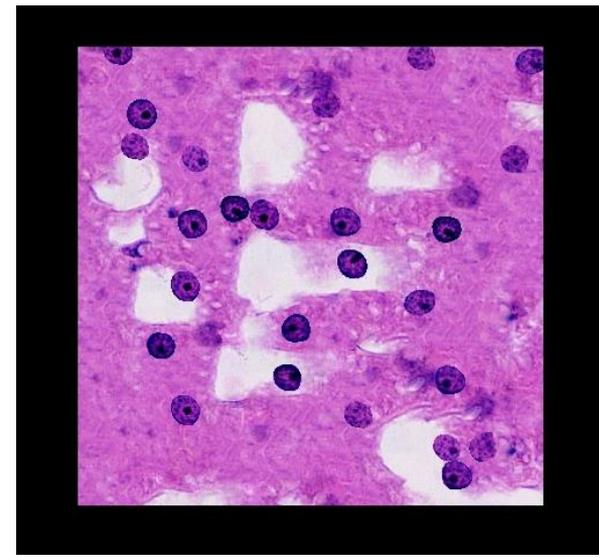
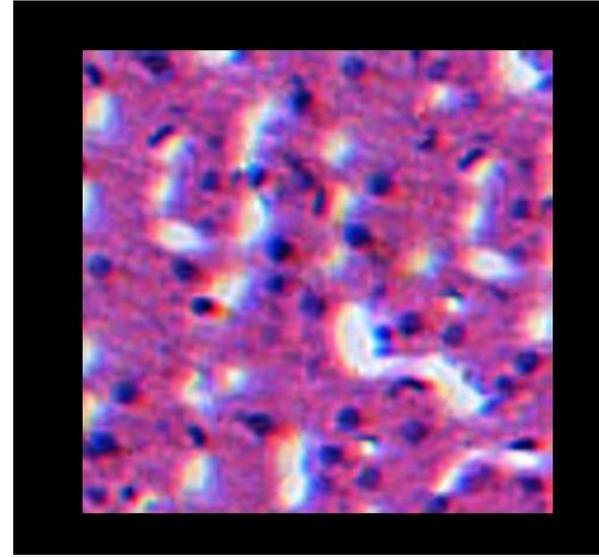
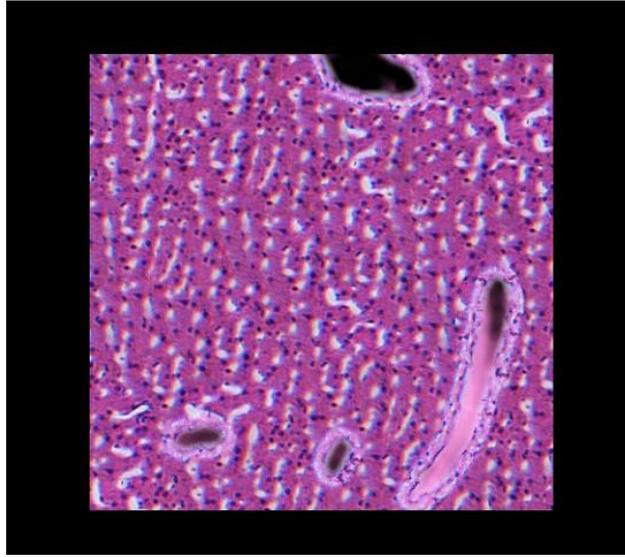
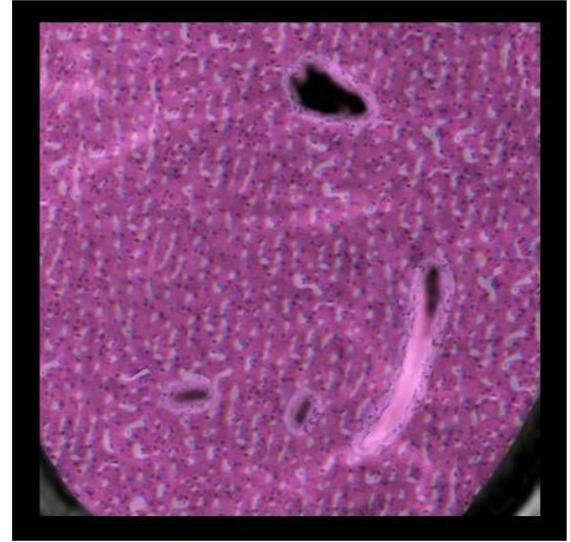
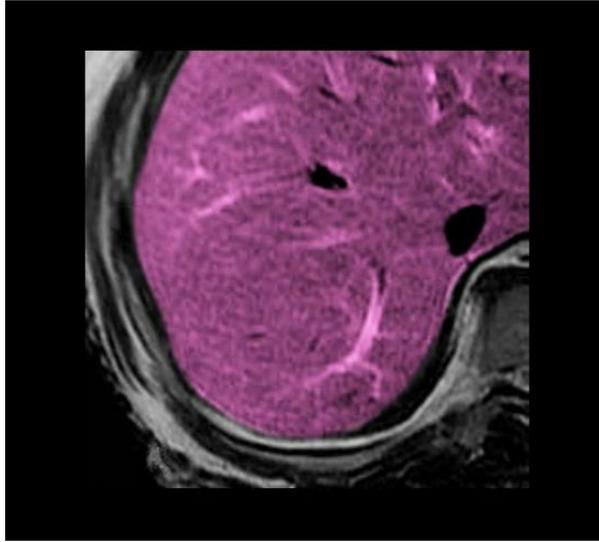
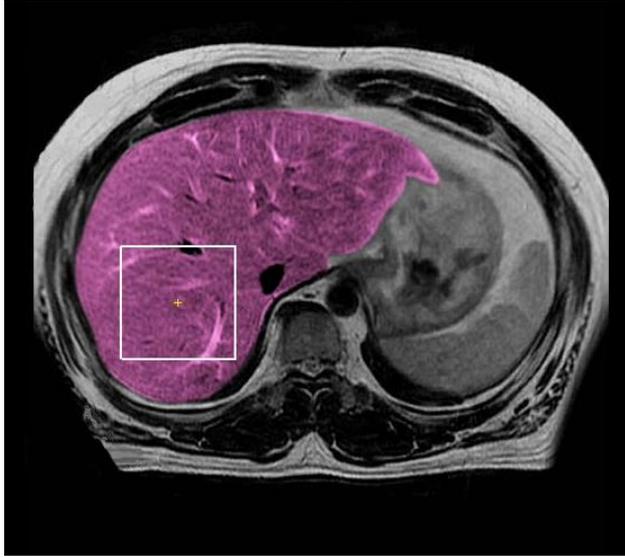
Weighted Blending



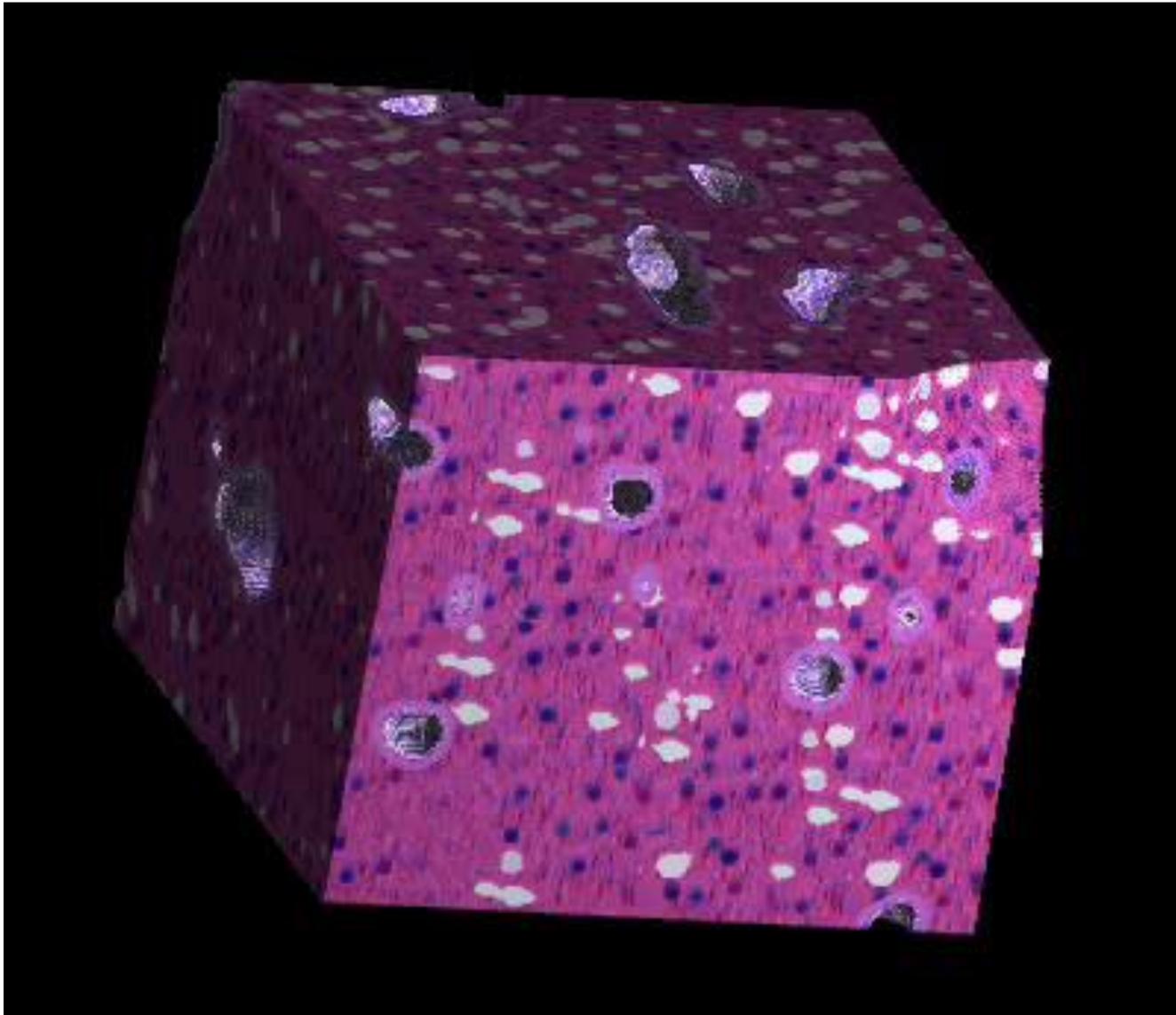
Semantic Zoom: The Infinite Microscope



Semantic Zoom: The Infinite Microscope Still Frames



3D Microscope Semantic Zoom



Semantic Zooms: Information Visualization

Could show different levels/aspects of information

- on a map, show either parking lots, bars, or restaurants
- zoom in by price range (cheap first, then more expensive...)
- zoom in by preference (favorite food first, then less favorite...)
- may combine these criteria into a preference function

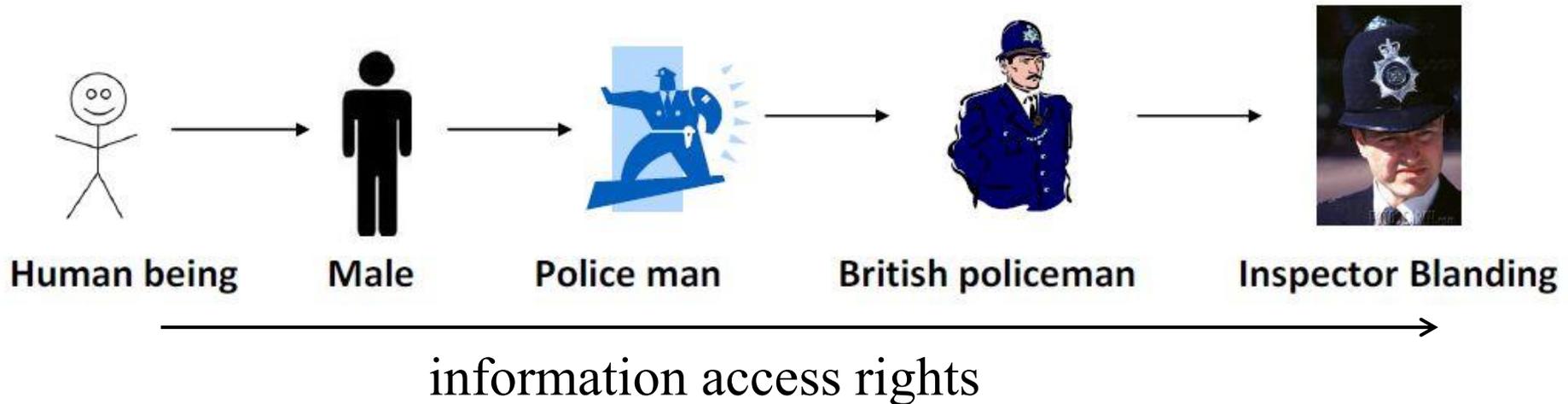
Semantic Zooms: Information Visualization

Could show different levels/aspects of information

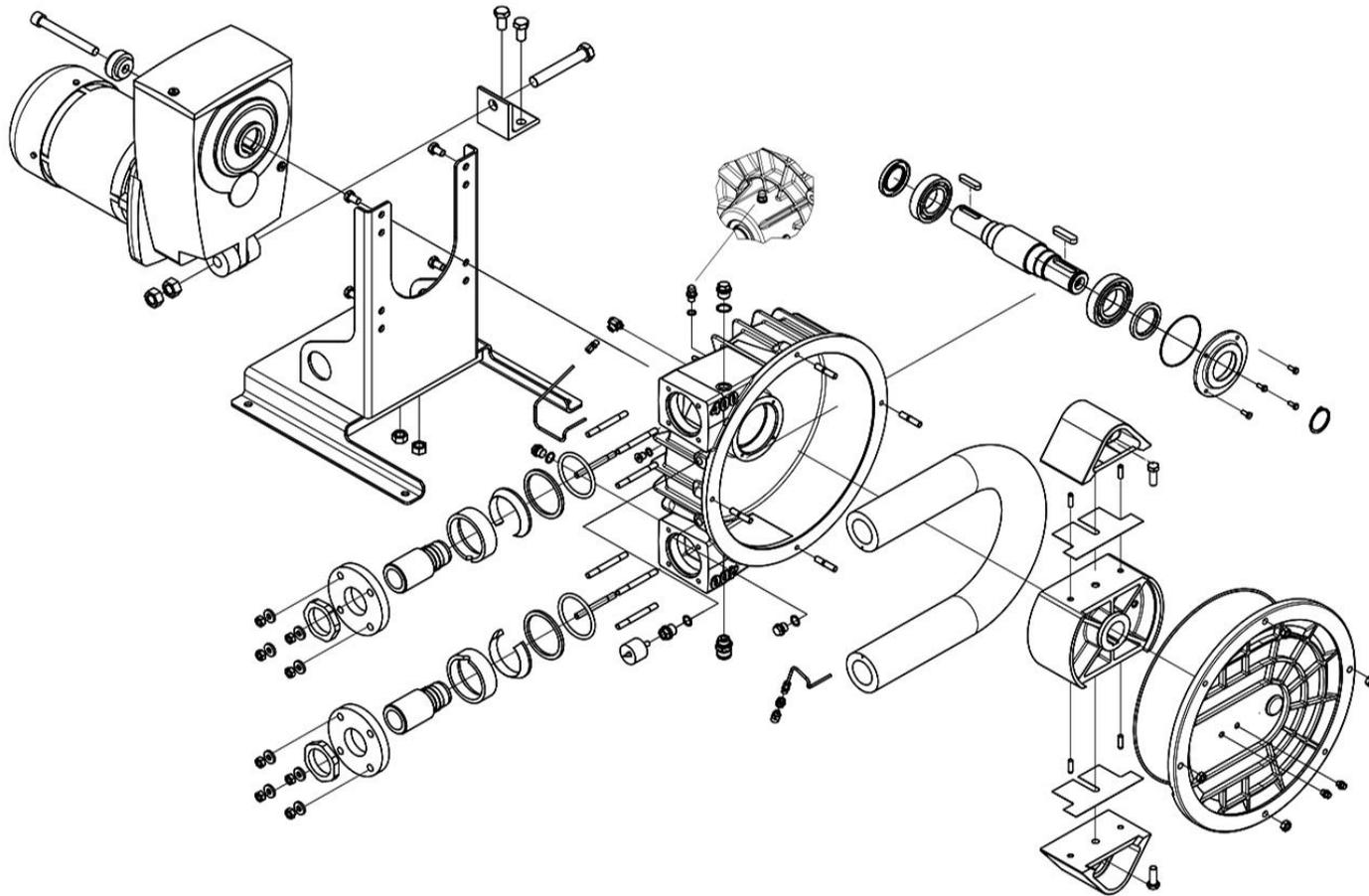
- on a map, show either parking lots, bars, or restaurants
- zoom in by price range (cheap first, then more expensive...)
- zoom in by preference (favorite food first, then less favorite...)
- may combine these criteria into a preference function

Zoom levels may require access rights

- members only
- big wallets only
- classified information



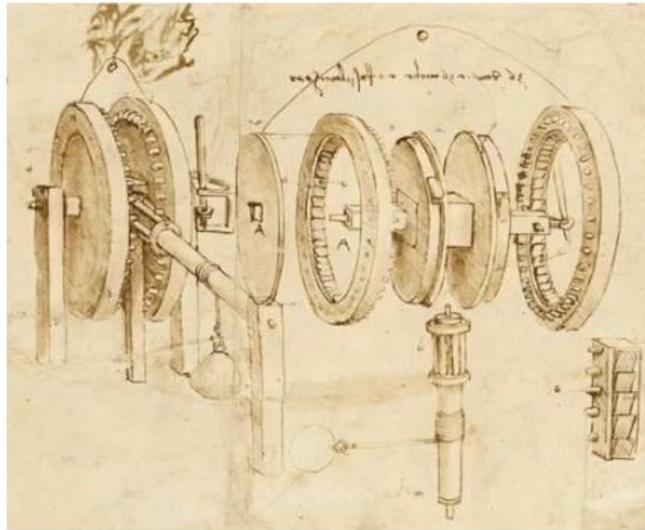
Exploded Views



S, Mahmood, K. Mueller, “An Exploded View Paradigm to Disambiguate Scatterplots,” Computers & Graphics, 2018

Exploded View Applied to Infovis

An Exploded View Paradigm to Disambiguate Scatterplots



[video](#)

Brushing and Linking

Interactive technique

- Highlighting
- Brushing and Linking

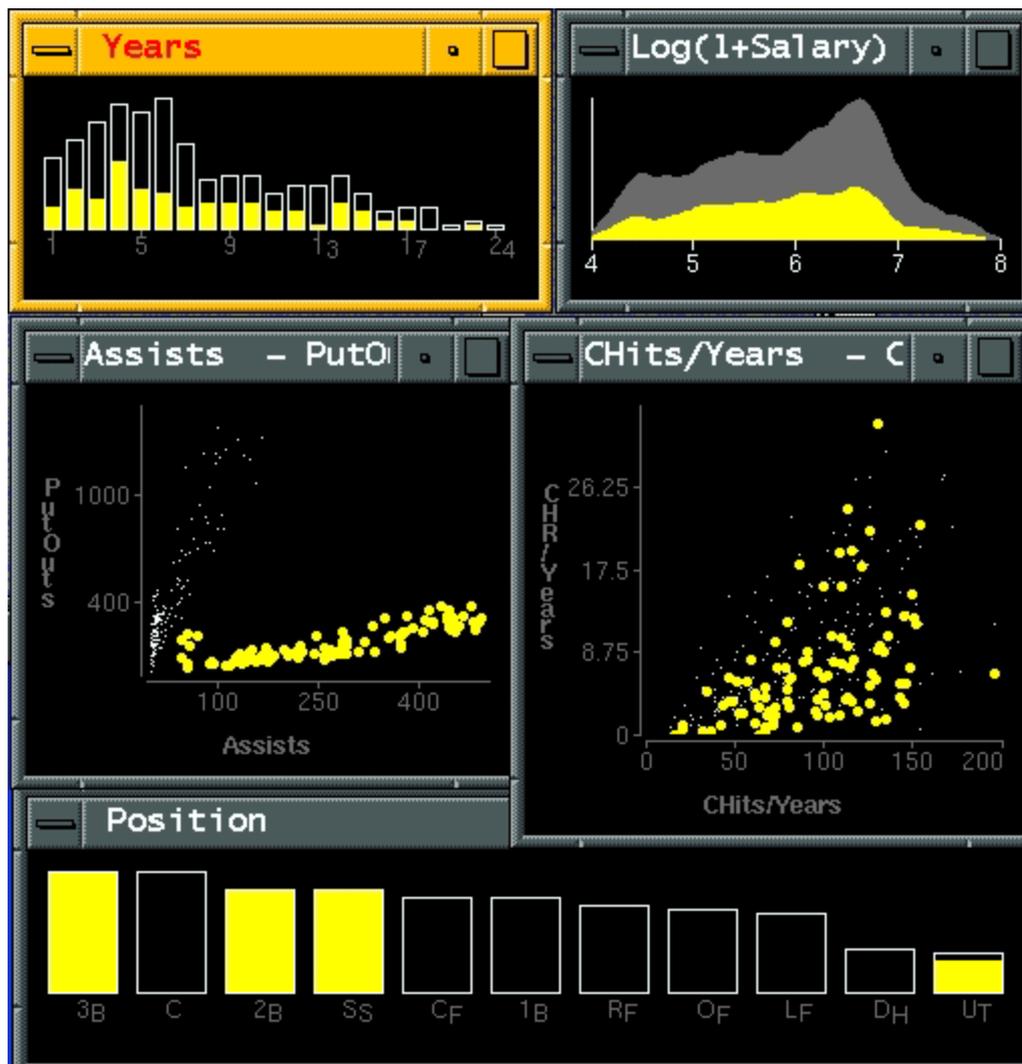
At least two things must be linked together to allow for brushing

- select a subset of points
- see the role played by this subset of points in one or more other views

Example systems

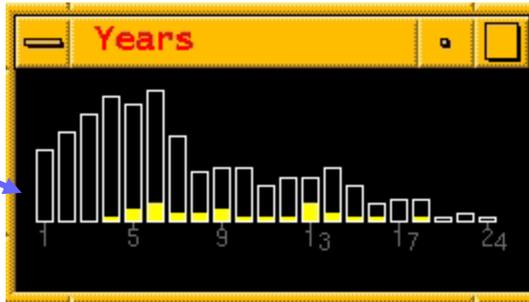
- Graham Will's EDV system
- Ahlberg & Sheiderman's IVEE (Spotfire)

Linking Types of Assist Behavior to Position Played

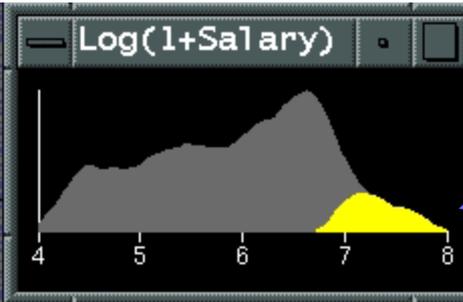


Baseball Data: Scatterplots and Histograms and Bars

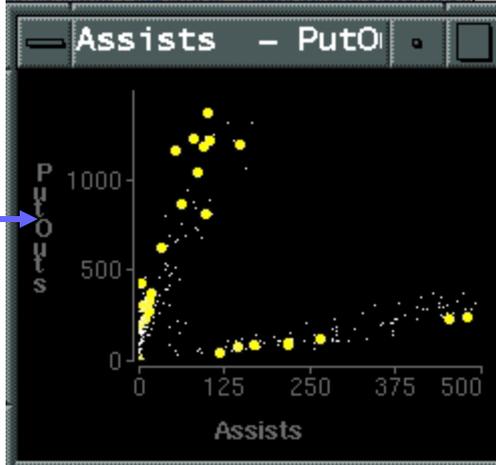
how long
in majors



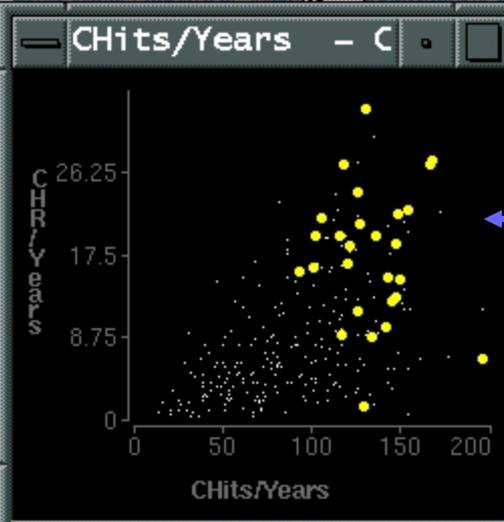
select high
salaries



avg. assists (x)
vs.
avg. putouts (y)
(fielding ability)



avg. career
home runs (y)
vs.
avg. career hits (x)
(batting ability)



distribution
of positions
played

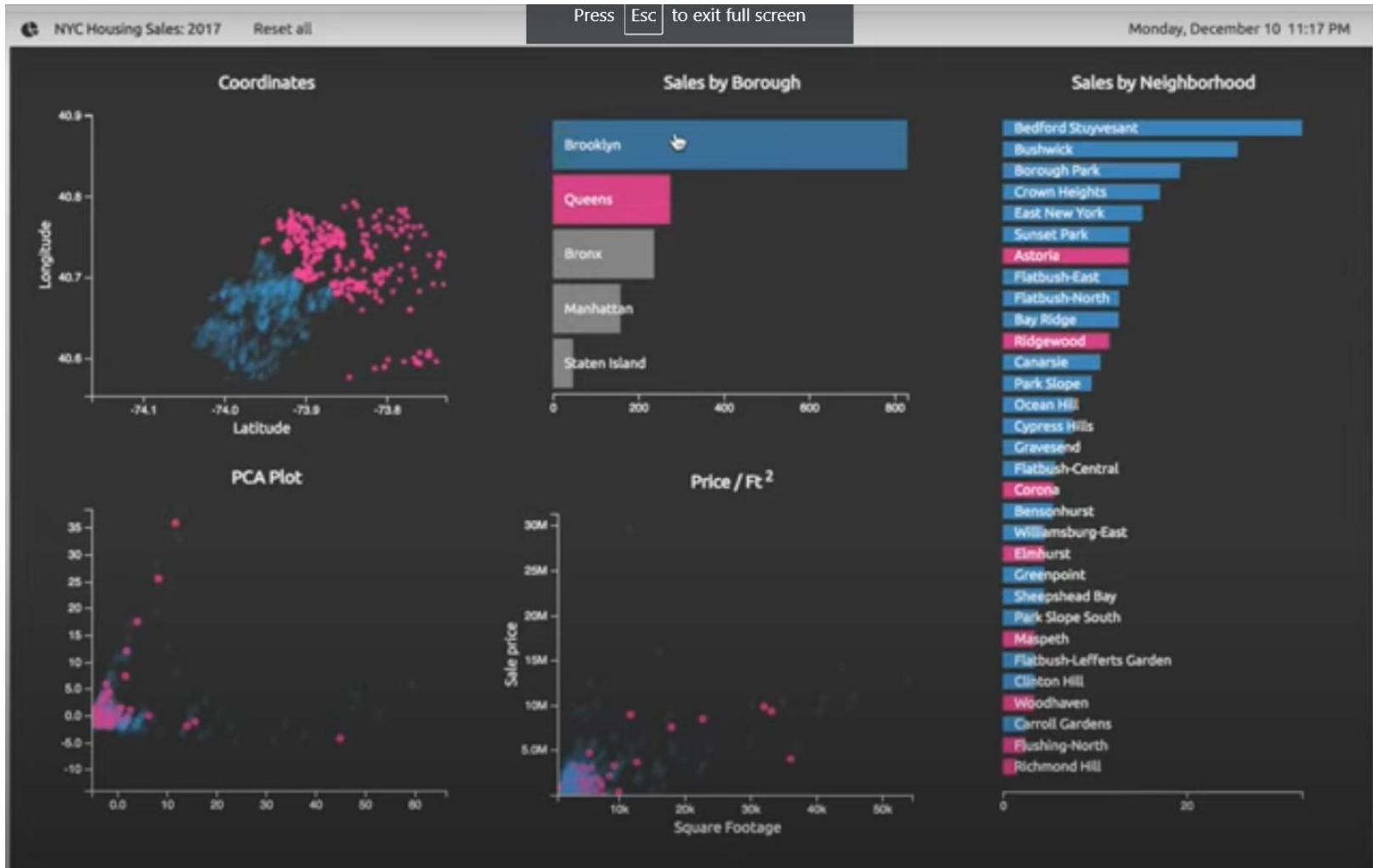


anything interesting?

What was Learned from Interaction w/ the Baseball Data?

- Seems impossible to earn a high salary in the first three years
- High salaried players have a bimodal distribution (peaking around 7 & 13 yrs)
- Hits/Year a better indicator of salary than HR/Year
- High paid outlier with low HR and medium hits/year. Reason: person is player-coach
- There seem to be two differentiated groups in the put-outs/assists category (but not correlated with salary) Why?

Example for Linking and Brushing

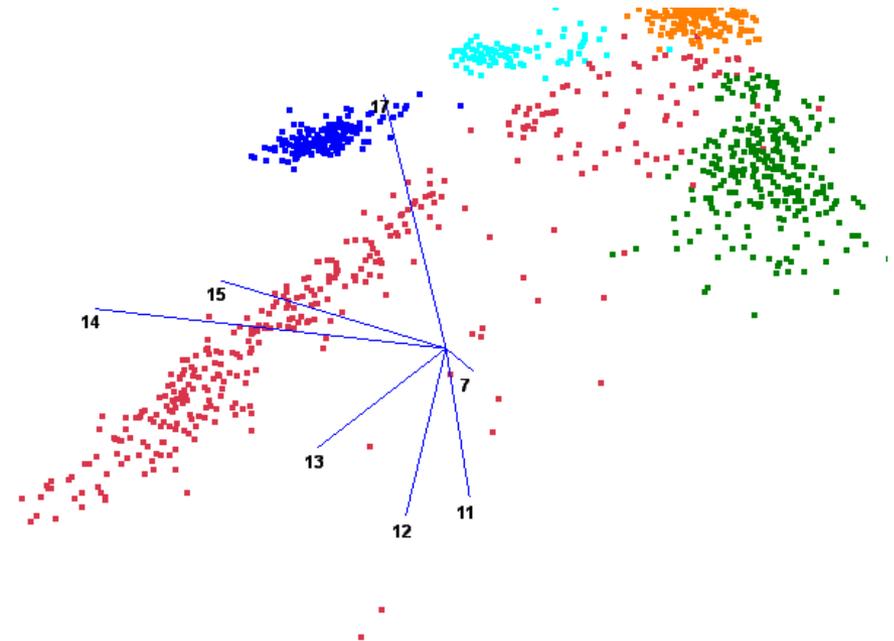
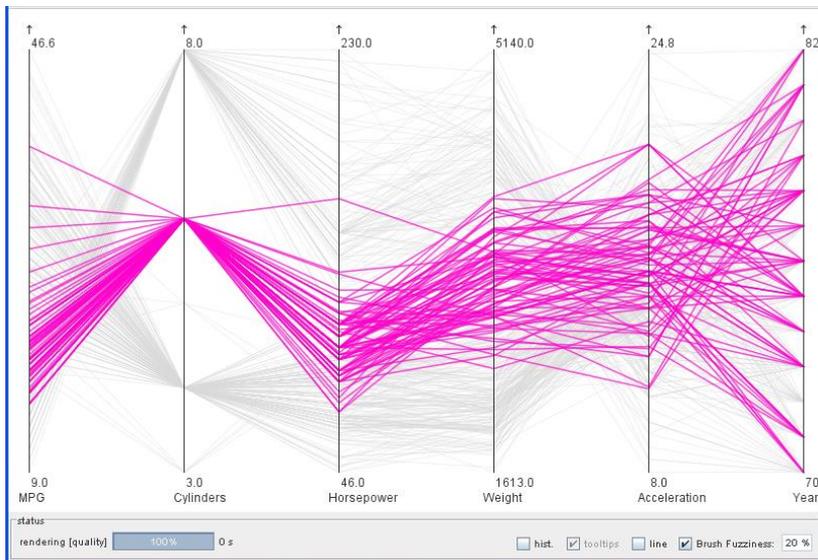


[Video](#)

Brushing: Highlighting

Use mouse interaction to highlight points and lines in

- parallel coordinates
- scatterplots



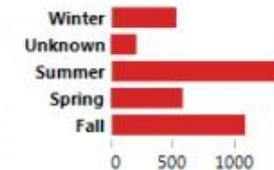
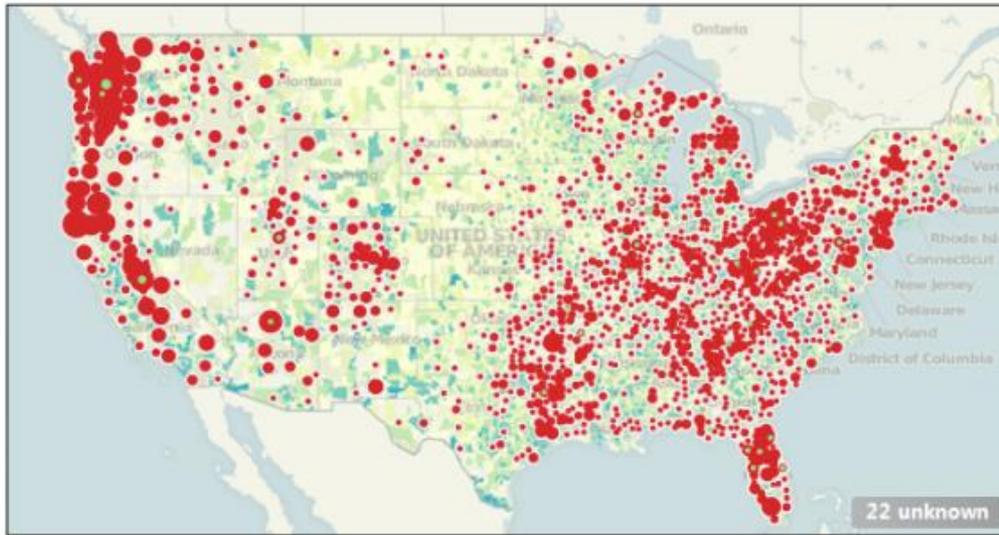
Interaction Video #1

Interaction in Parallel Coordinate

Accelerate Understanding

Dashboards should pass the 5-second test

Finding Bigfoot



Data gathered from the official website of the "Bigfoot Field Researchers Organization" (BFRO).

The data was attempted to be scrubbed and cleaned to attain some type of normalcy, unfortunately the BFRO data submission process has no validation and fields are often used arbitrarily by submitters.

BFRO does the "Finding Bigfoot" Animal Planet TV show.

Click on ANY element of the visualization (location, season, year, detail field) in order to filter by that item. Select the element AGAIN to go back to the full view.



The BFRO classifies sightings according to a system based on the sightings 'potential for misinterpretation'.

Total Sightings

3,806

Class A

1,951

Class B

1,696

Class C

31

Unclassified

128

Accelerate Understanding

Important rules:

- most important view goes on top or top left
- legends go near their views
- avoid using multiple color schemes
- use 5 views or fewer
- provide interactivity

