

Visual data analysis

Matti Gröhn

Matti.Grohn@ttl.fi



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

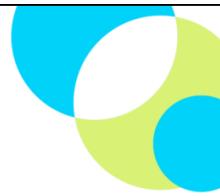


Richard Hamming:" The purpose of scientific computing is insight not numbers."



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Outline

Background
Goals of Visualization
General visualization techniques
Representation of multidimensional data
Introduction to R
Animation
Other data representation methods
Future



Background

Visual perception
~200 different hues
~20 levels of saturation
~300-400 levels of brightness

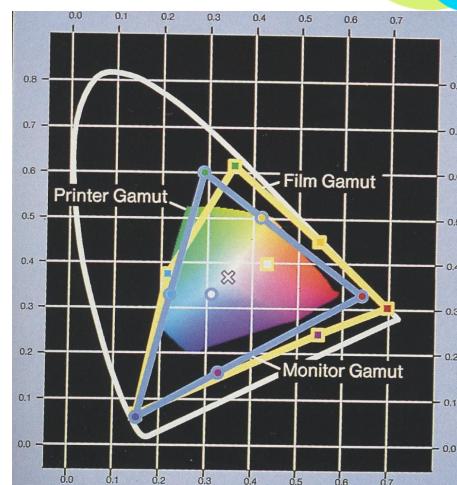
Common problem: red-green color blindness
(8%-12% of males of European origin, and 0.5%
of females)

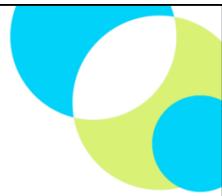
Background

- The color reproduction varies from monitor to monitor, and printer to printer.
- The monitors and printers use different color spaces (RGB for the monitors, CMY(K) for the printers)
- They don't cover all the possible colors and not even the same colors.

Different color spaces

- Printer, film and monitor color spaces in comparison with our color perception.





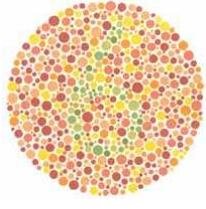
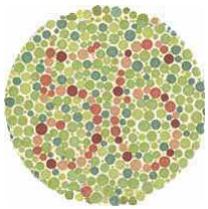
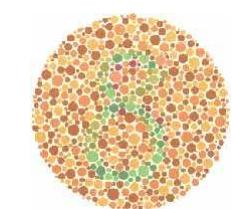
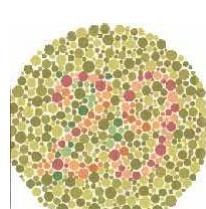
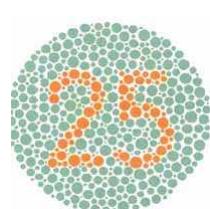
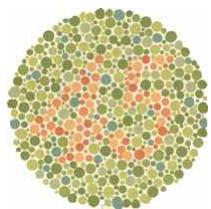
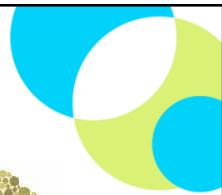
Red-green color blindness

- Common problem: (8%-12% of males of European origin, and 0.5% of females)



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



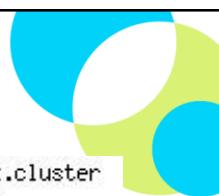
Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

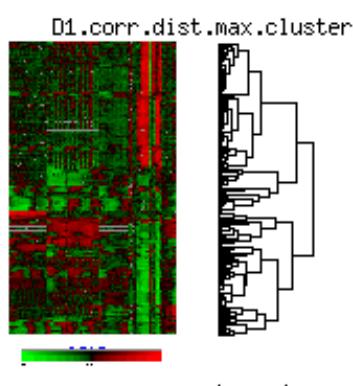


Color blindness and visualization

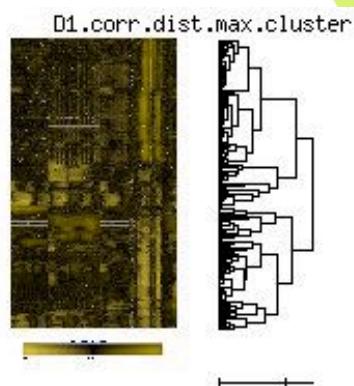
- The 'true' color of something may be irrelevant, but the fact it is different from the surroundings is very important.



Microarray data



• Normal



as seen by redgreen-blind



Definition

- Visualization: The use of computer-supported, interactive, visual representations of data to amplify cognition.



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Background

From symbolic information to geometric information.

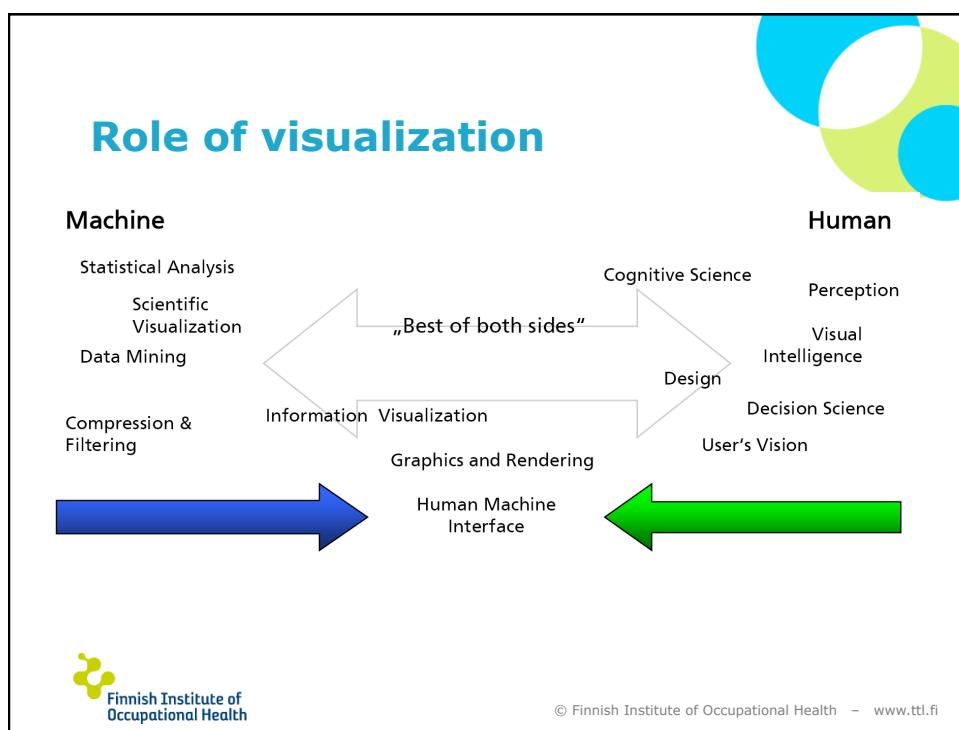
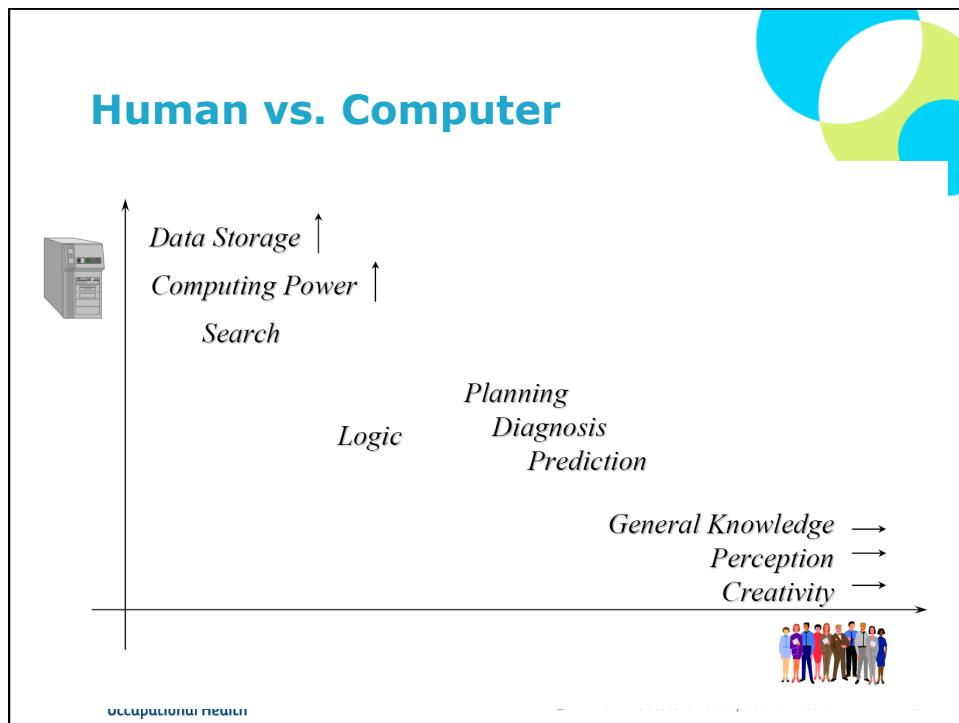
Development of computer graphics enables 3D visualizations and animations.

Visualizations give us better insight.

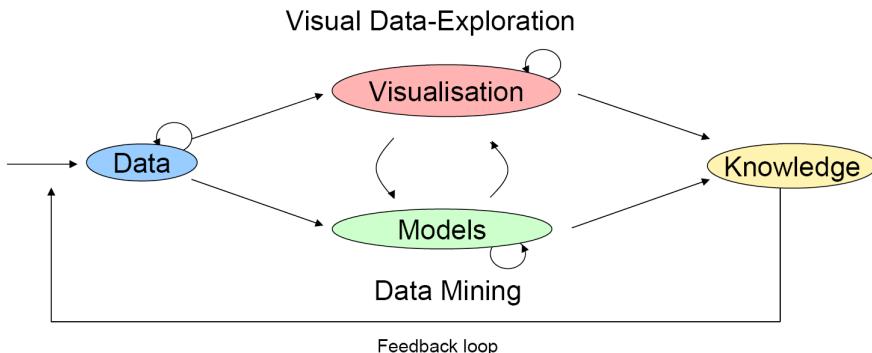


Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



From data to knowledge



© Finnish Institute of Occupational Health - www.ttl.fi

Motivation 1

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

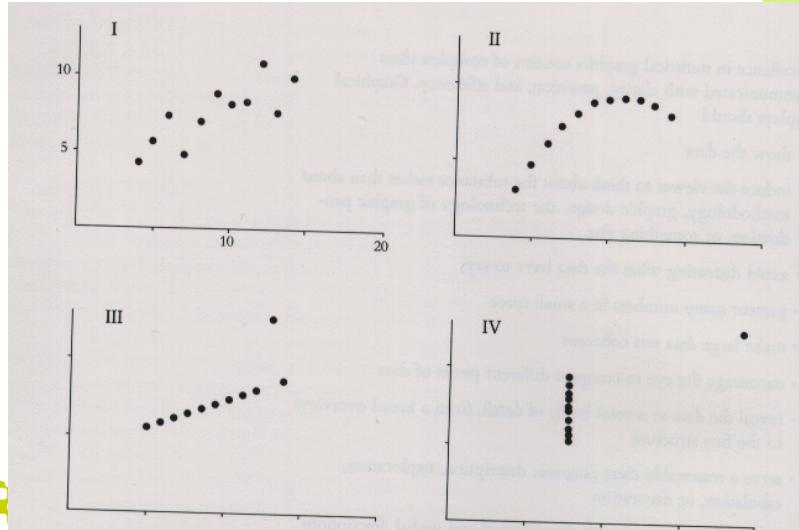
N = 11
mean of X's = 9.0
mean of Y's = 7.5
equation of regression line: $Y = 3 + 0.5X$
standard error of estimate of slope = 0.118
 $t = 4.24$
sum of squares $\sum (X - \bar{X})^2 = 110.0$
regression sum of squares = 27.50
residual sum of squares of Y = 13.75
correlation coefficient = .82
 $r^2 = .67$

Anscombe's quartet (Tufte: The Visual Display of Quantitative Information)



© Finnish Institute of Occupational Health - www.ttl.fi

Motivation 2



www.ttl.fi

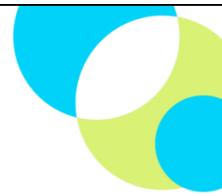
Goals of visualization 1/3

Explorative analysis

Starting point: data without hypotheses about the data

Process: interactive, usually undirected search for structures, trend etc.

Result: visualization of the data, which provides hypotheses about the data



Goals of visualization 2/3

Confirmative analysis

Starting point : hypotheses about the data

Process : goal-oriented examination of the hypotheses

Result : visualization of the data, which allows the confirmation or rejection of the hypotheses



Goals of visualization 3/3

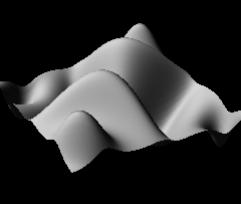
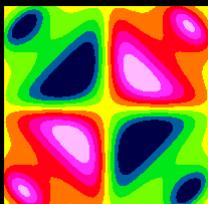
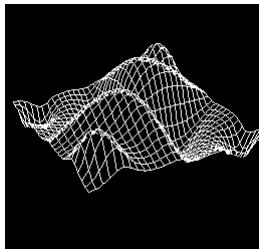
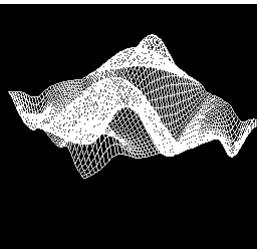
Presentation

Starting point : facts to be presented are fixed a priori

Process : choice of appropriate presentation technique

Result : high-quality visualization of the data presenting the facts.

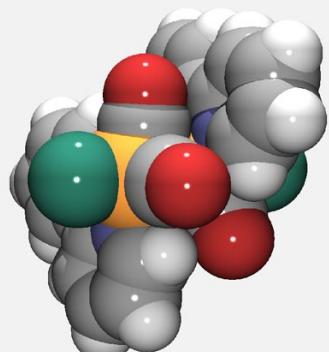
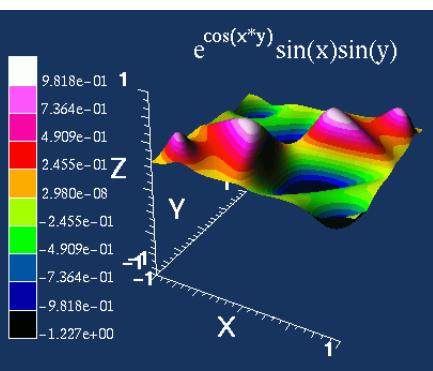
Visualizations for explorative and confirmative analysis



Finnish Institute of
Occupational Health

© Finnish Institute of

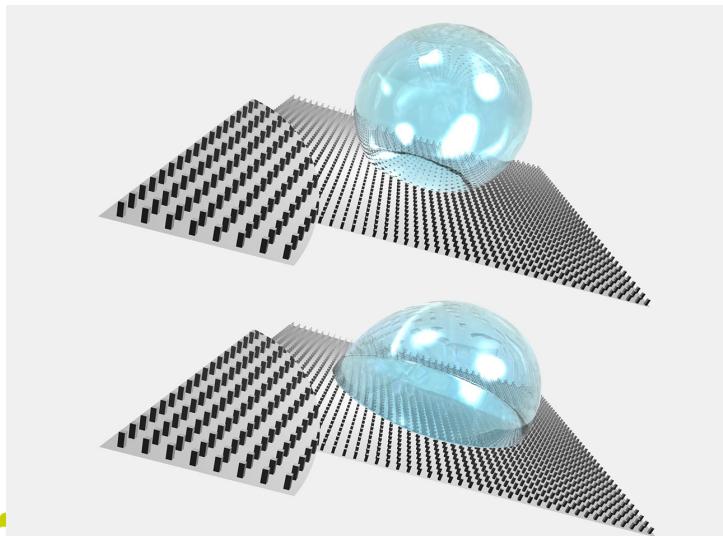
Visualizations for presentations



Finnish Institute of
Occupational Health

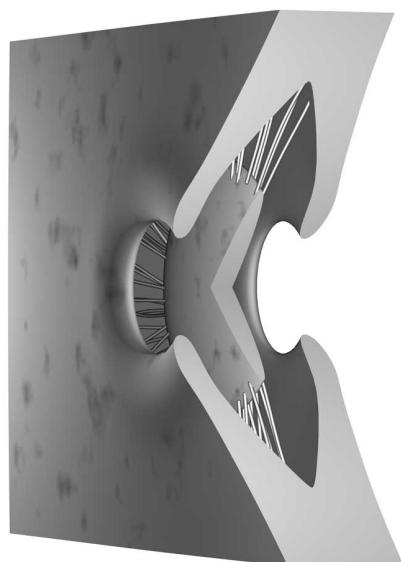
© Finnish Institute of Occupational Health – www.ttl.fi

Visualizations for presentations



occupational health

© Finnish Institute of Occupational Health - www.ttl.fi



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health - www.ttl.fi





Visualizations for presentations

- <http://www.csc.fi/english/research/sciences/visualization/gallery>



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Something to consider:

Emphasize your facts

The viewer has not the same insight than you.

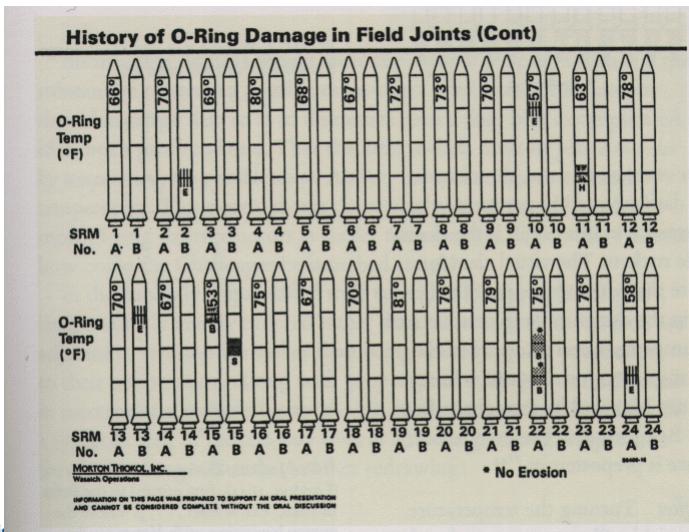
Visualize only the essential information



Finnish Institute of
Occupational Health

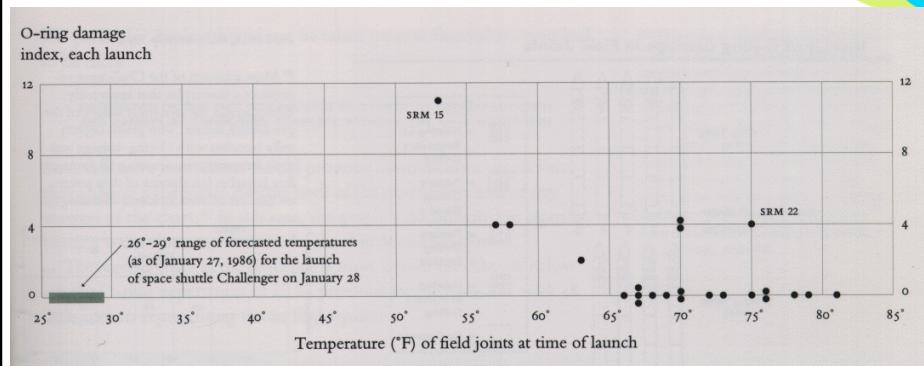
© Finnish Institute of Occupational Health – www.ttl.fi

Original shuttle visualization



- www.ttl.fi

More revealing visualization



Tufte: Visual Explanations



© Finnish Institute of Occupational Health - www.ttl.fi



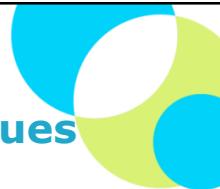
Tufte (Visual explanations):

There are right ways and wrong ways to show data; there are displays that reveal the truth and displays that do not.



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



General visualization techniques

false color maps
surfaces
isocontours
isosurfaces
arrows

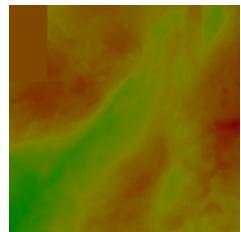
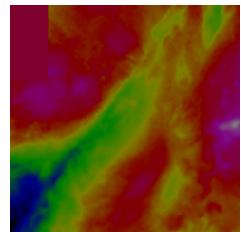
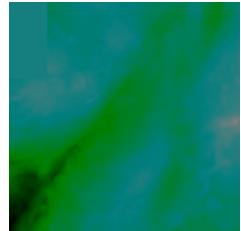
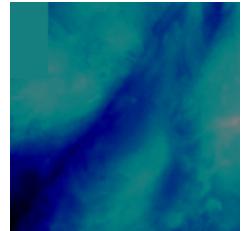
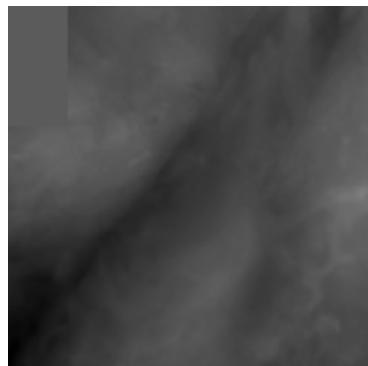
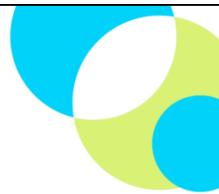
particles
graphs
glyphs
volume rendering
plots



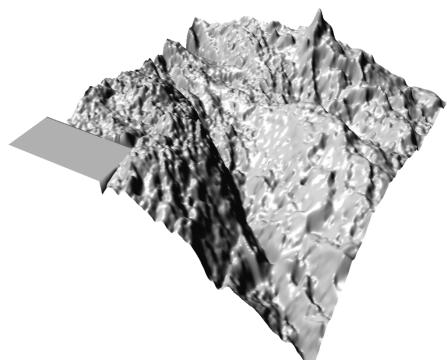
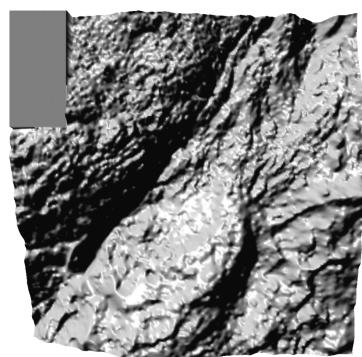
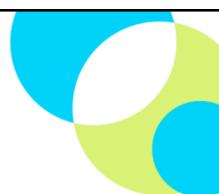
Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

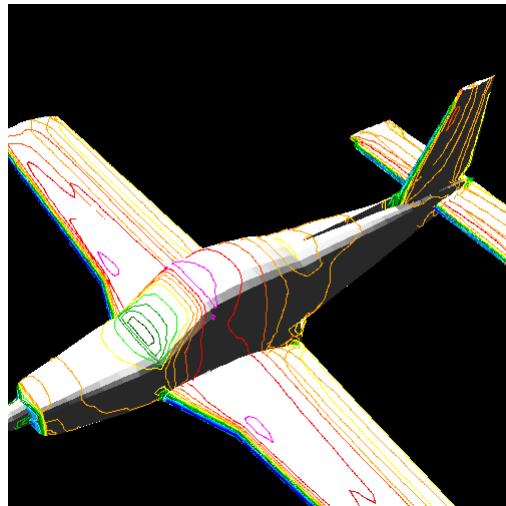
False color maps



Surfaces



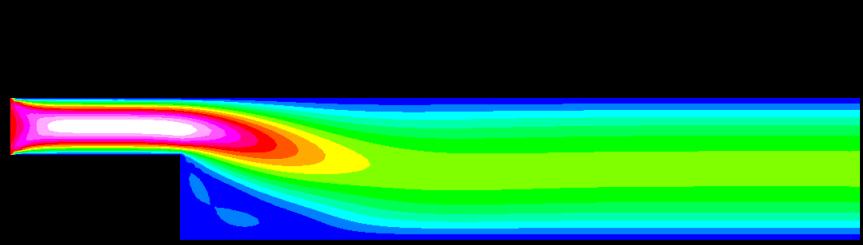
Isocontours



© Finnish Institute of Occupational Health – www.ttl.fi



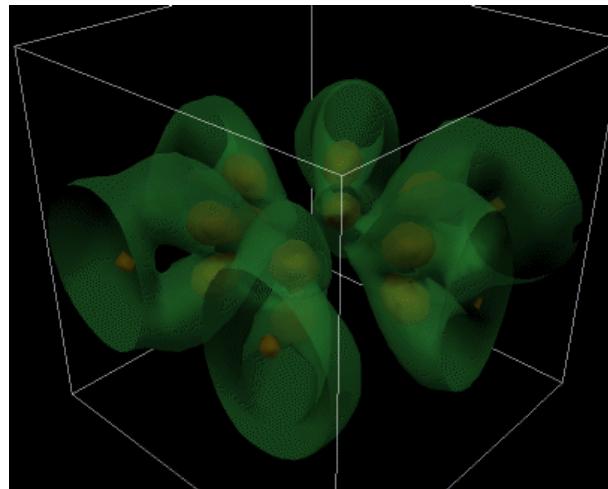
Segmented color map as isocontours



© Finnish Institute of Occupational Health – www.ttl.fi



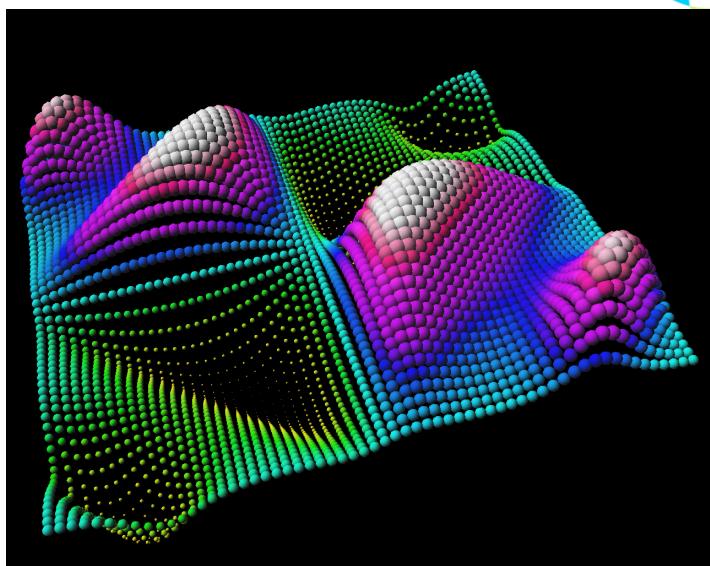
Isosurfaces



 Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

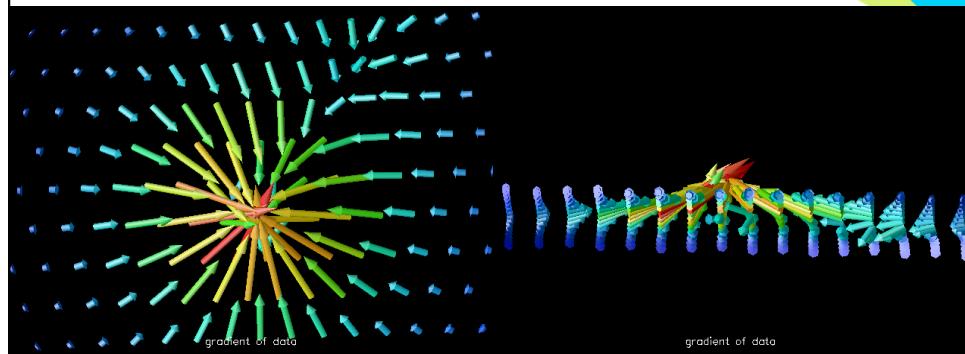
Glyphs



 Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

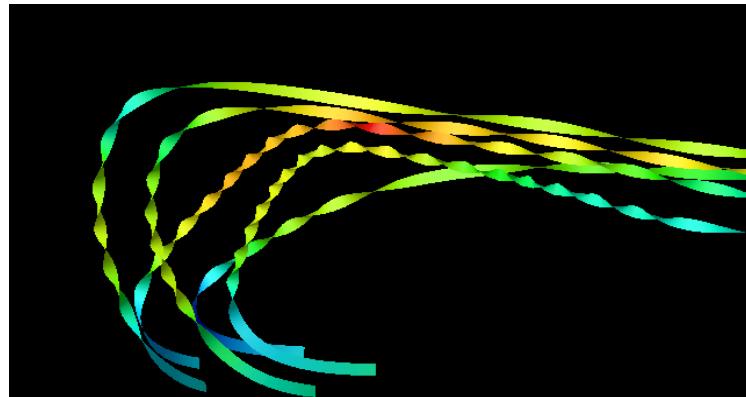
Arrows 2



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Streamlines



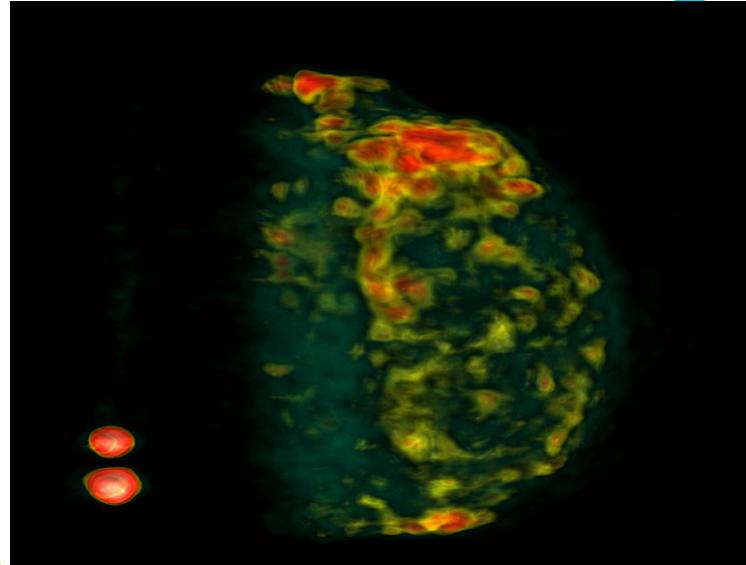
data courtesy of NCSA, University of Illinois at Urbana-Champaign



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Volume rendering



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Combinations 1

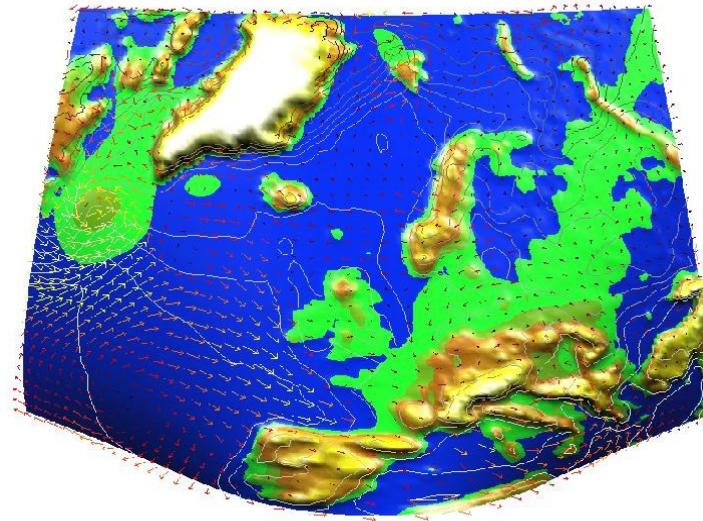


crank -7.64000
max X 9.270
max Z 62.450

Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

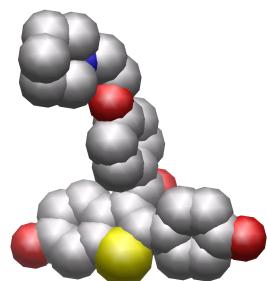
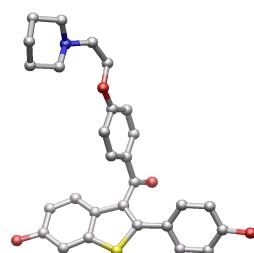
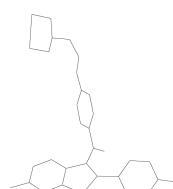
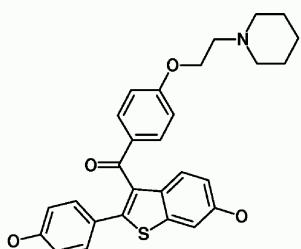
Combinations 2



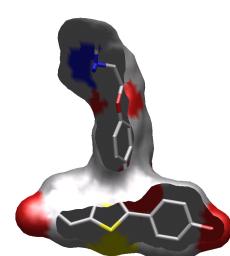
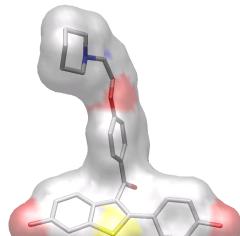
 Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

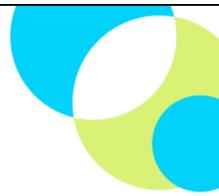
Molecule visualizations



Occupational Health



© Finnish Institute of Occupational Health – www.ttl.fi



Representation of arbitrary multidimensional data

Classification

- Geometric techniques
- Icon-based techniques
- Pixel-oriented techniques



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Geometric techniques

Visualization of geometric transformations and projections of the data

Techniques:

- Scatterplots
- Landscapes
- Prosection views
- Hyperslice
- Parallel Coordinates

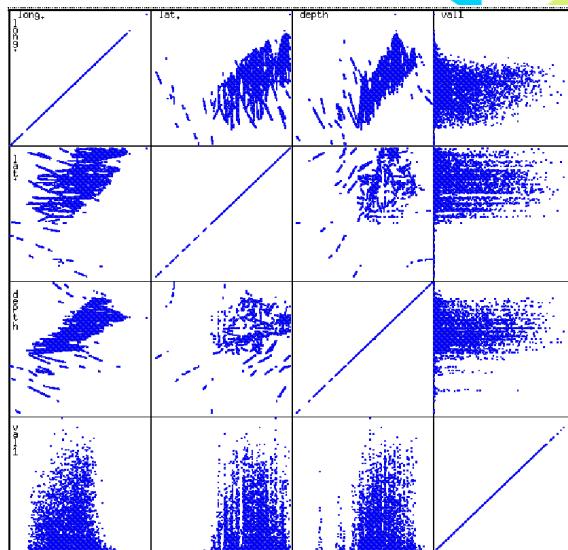


Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Scatterplot-matrices

- Matrix of scatterplots (x-y-diagrams)

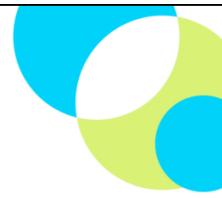


Parallel Coordinates

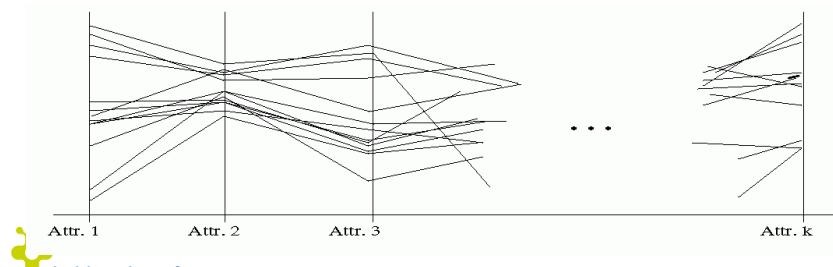
n equidistant axes for n attribute

The axes are scaled to the min-max range of the corresponding attribute.

Parallel Coordinates

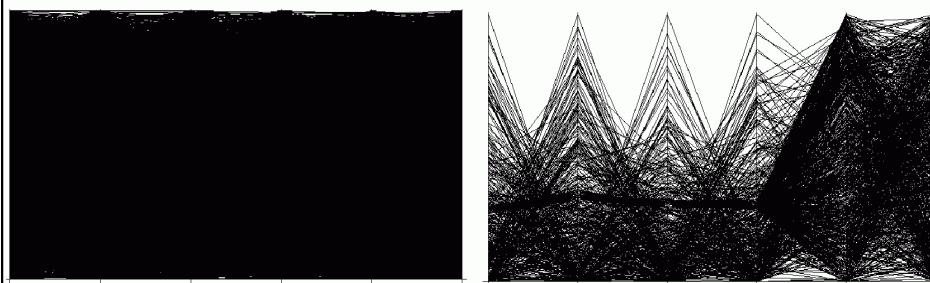


- Every data item is represented as a polygonal line, which intersects each of the axes at the point which corresponds to the value for the attribute.

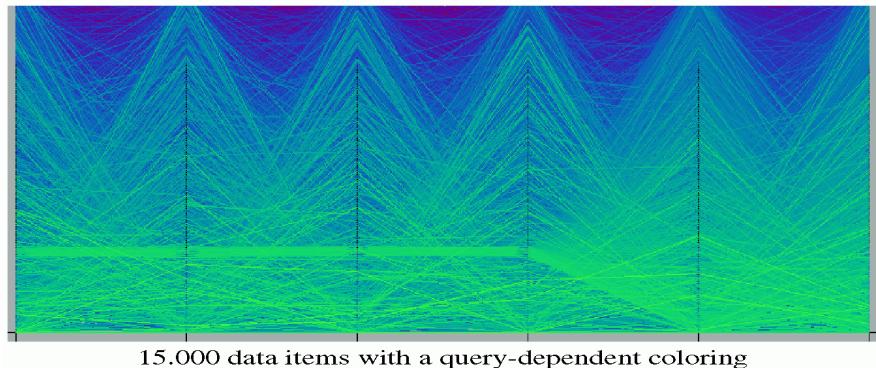


© Finnish Institute of Occupational Health – www.ttl.fi

Parallel Coordinates



Parallel Coordinates



Icon-based Techniques



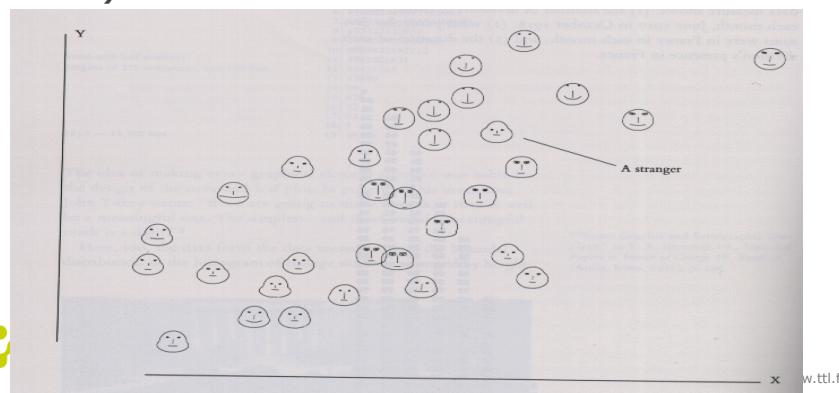
Visualization of data values as
feature of icons

Techniques

- Chernoff-Faces
- Stick figures
- Shape coding
- Color-icons

Chernoff-Faces

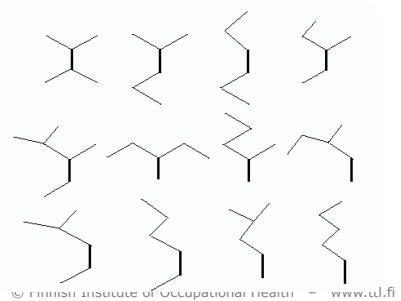
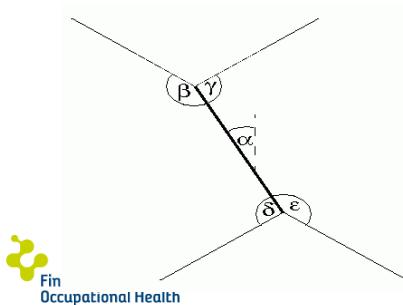
Visualization of the data using the properties of a face icon (shape of nose, mouth, eyes, and the face itself)



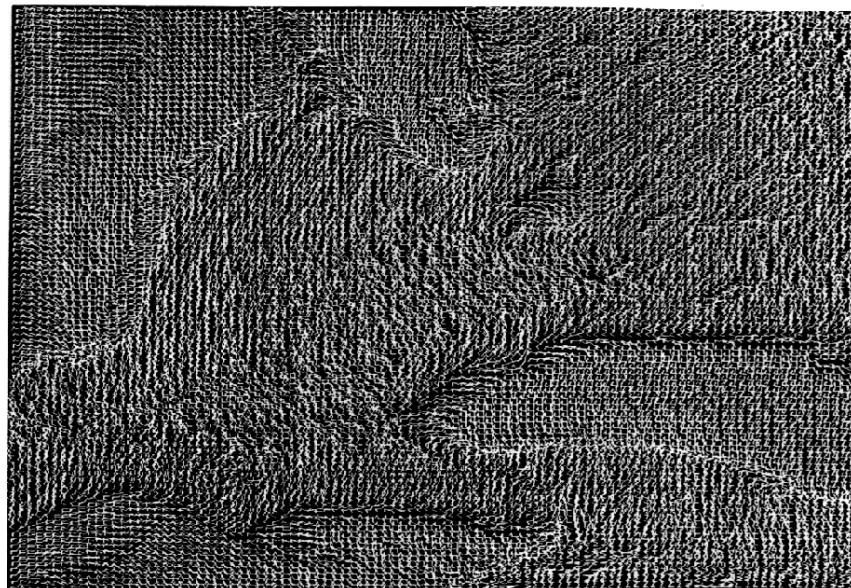
Stick figures

Two attributes are mapped to the display axes and the remaining attributes are mapped to the angle and/or length of the limbs.

Texture patterns in the visualization show certain data characteristics



Stick figures

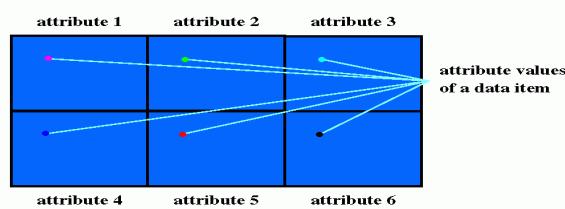


.fi

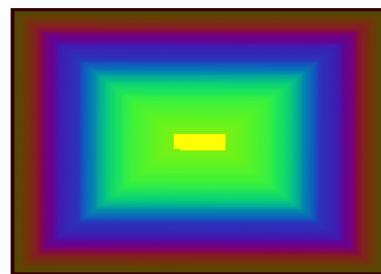
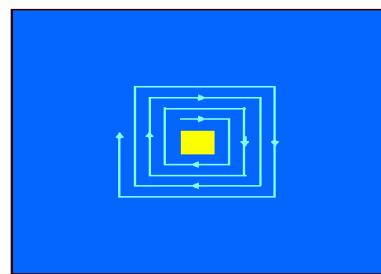
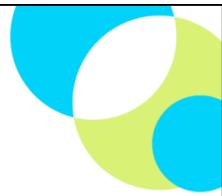
Pixel-oriented techniques

Each attribute value is represented by one colored pixel.

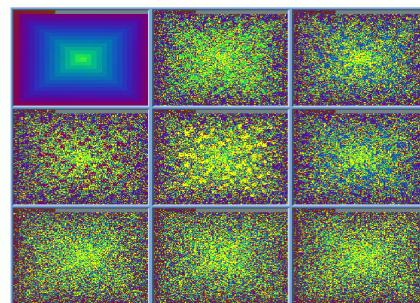
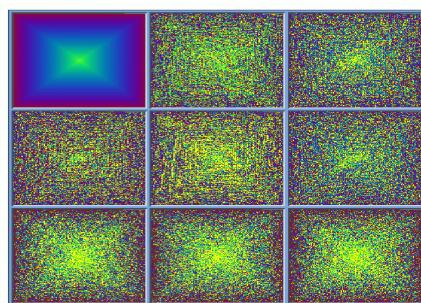
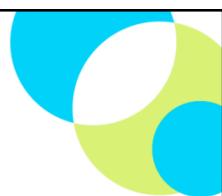
The attribute values for each attribute are presented in separate subwindows.



Pixel-oriented techniques



Pixel-oriented techniques



Other visual data mining techniques

Hierarchical techniques

Graph-based techniques

Hybrid techniques

Distortion techniques

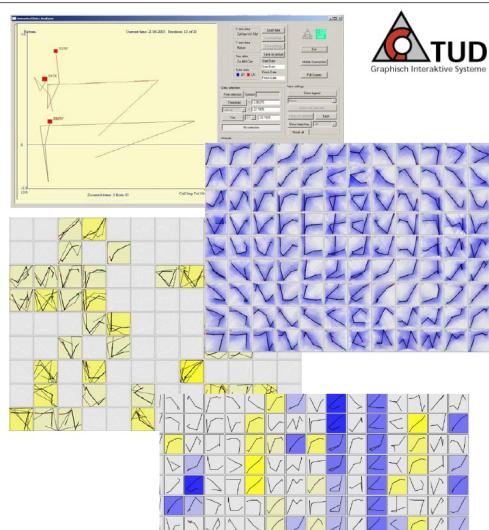
Daniel A. Keim: Visual Data Mining



t1

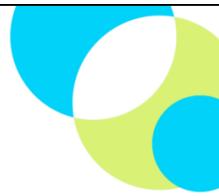
FinExplorer – Multidimensional Visual Analysis of Time Series

- Financial market analysis in 2D
- Visualisation and analysis of market dynamics
- Identification of interesting patterns and relationships
- Automatic analysis aspects
- Interactive presentation
- Application domains
 - Financial market analysis
 - Macro-economic analysis
 - Insurance Analysis
 - and more



Introduction to R

- Hands-on exercise
 - scatterplots
 - parallel coordinates



Finnish Institute of
Occupational Health

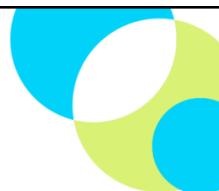
© Finnish Institute of Occupational Health – www.ttl.fi

Animation

Motivation

Animations could reveal information, that is hard or impossible to perceive from static images

Exploration of complex time-series data



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



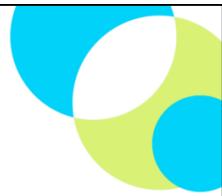
Definition

- Animate: 1. having life; lively 2. enliven: give life to
- Animated: 1. lively, vigorous, having life 2. characterized by animation
- Animation: 1. vivacity, ardour 2. being alive 3. photographing successive drawings or positions of puppets to create illusion of movement
 - (The Oxford Dictionary of Current English)



Illusion of movement

- Small changes between images
- TV frame rate 25 Hz
- Movie film frame rate 24 Hz
- Real time computer animation 8 - 60 Hz

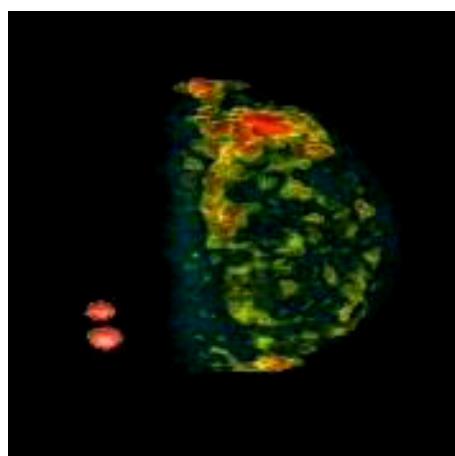


Animation

could use for
causality
transformation of structures
time-varying data
changes in parameter values
scaling of time
perceiving the 3D structure of complex objects



Animation





Other data representation methods

Haptic feedback

Sonification

Virtual reality

Augmented reality



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Haptic feedback

Provides physical feedback to user

Force-feedback

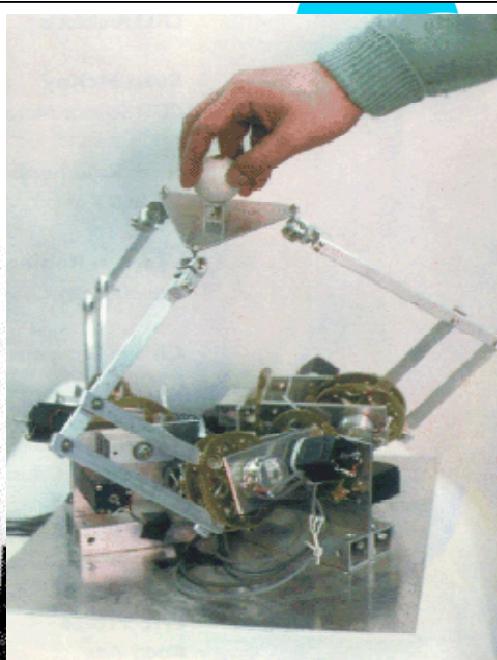
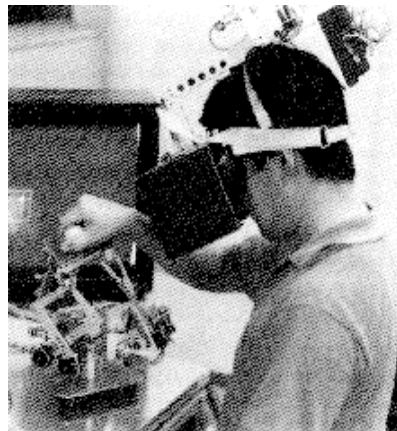
Tactile-feedback



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Force-feedback



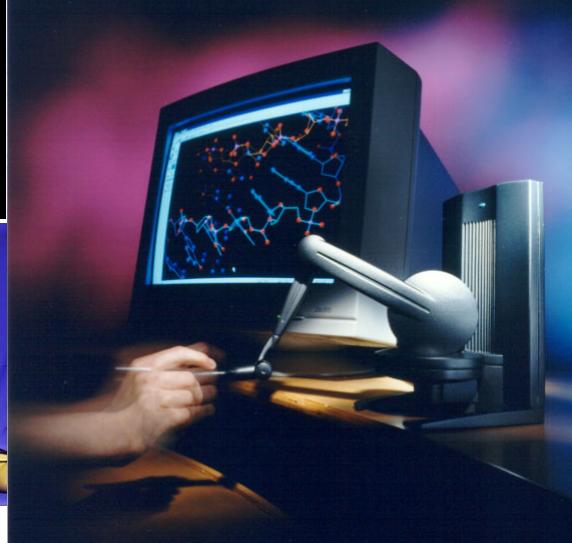
Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Haptic feedback



Finnish Institute of
Occupational Health





Sonification

Short definition: The use of nonspeech audio to convey information

Longer definition: Sonification is the transformation of data relations into perceived relations in an acoustical signal for the purposes of facilitating communication or interpretation



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Benefits

Bring out new features of the data
Perception of time dependent changes
Repetitive patterns are perceived
Decrease visual overload



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Virtual Reality

Computer generated multi-modal world with which the user can interact.

Components: stereo graphics, spatial audio, interaction devices...



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Virtual Reality

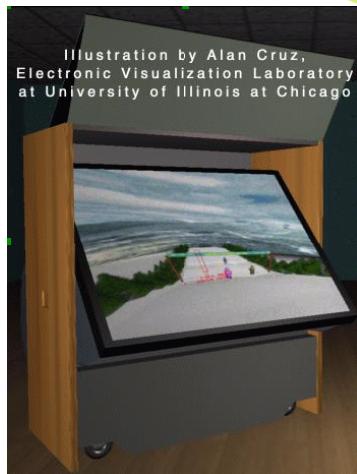
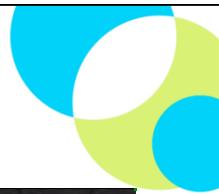
1. Head-Mounted Display (HMD)
2. Virtual Model Display (VMD)
3. Spatially Immersive Display (SID), virtual rooms



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

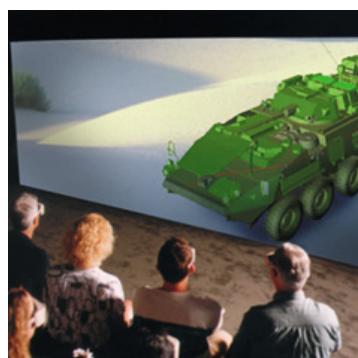
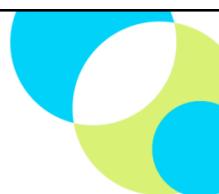
HMD and VMD



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Panoramic views

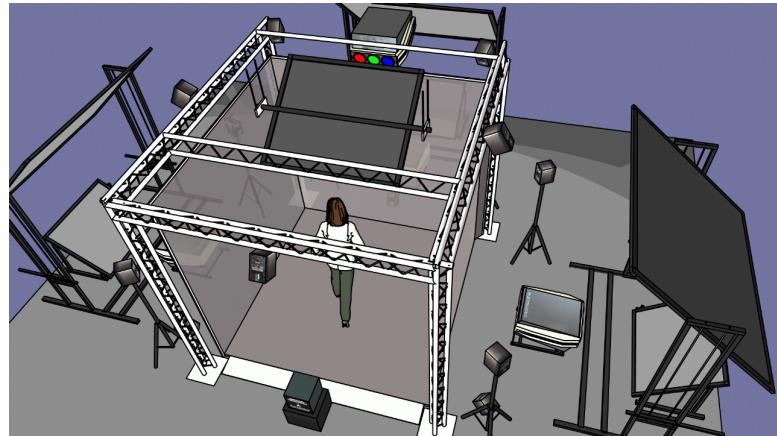


Multiple projectors
Edge-blended for one
panoramic view

Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Experimental Virtual Environment (EVE)



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi

Components of EVE



- Stereo graphics
- Spatial audio
- Advanced interaction techniques
 - trackers
 - two-handed interaction
 - speech recognition

Finnish Institute of
Occupational Health

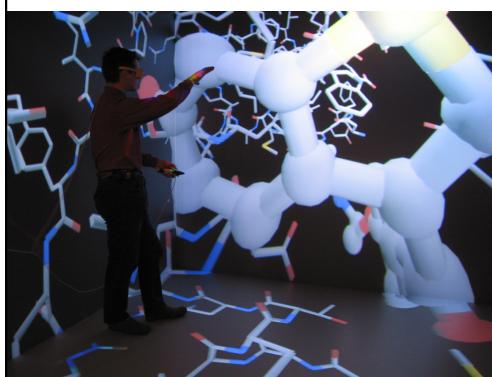
© Finnish Institute of Occupational Health – www.ttl.fi

Flow visualization in EVE



© Finnish Institute of Occupational Health www.ttl.fi

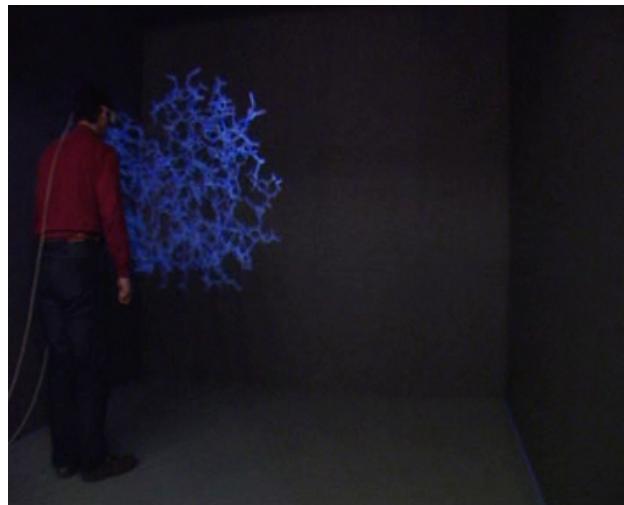
EVE (2)



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health www.ttl.fi

Examination of the protein-drug complex in EVE



 Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Future

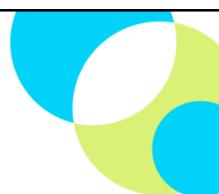
Hardware is no more the limit

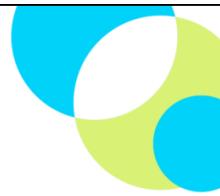
Usability of the programs will be the most critical issue

Integration of data analysis and representation

 Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi





Long term visions

Intuitive user interfaces

speech and gesture recognition

no more WIMP (30 year old technology...)

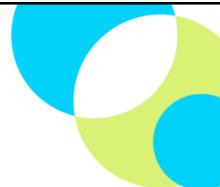
Wearable visualization devices (PDAs)

Personal virtual environments/augmented reality



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Examples

Interactive Visual Data Analysis

- <http://www.gapminder.org/>

Virtual reality

- EVE-video

New technology

- CastAR
<http://www.youtube.com/watch?v=AOI5UW9khoQ>
- Myo
- https://www.youtube.com/watch?feature=player_embedded&v=oWu9TFJjHaM



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



- Sonification

- [Earthquake: http://www.youtube.com/watch?v=3PJxUPvz9Oo](http://www.youtube.com/watch?v=3PJxUPvz9Oo)
- Distance and location sonification
- - video example

Future visions

- Productivity Future vision (~2:40-) video
- http://www.youtube.com/watch?v=6W8Q6wJ_Tt8



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi



Visual data analysis

Matti Gröhn

Matti.Grohn@ttl.fi



Finnish Institute of
Occupational Health

© Finnish Institute of Occupational Health – www.ttl.fi