Interactive Information Visualization

How to make "a picture [...] worth a thousand words"

Renaud Blanch <blanch@imag.fr>

Université Grenoble Alpes, Polytech'Grenoble & UFR IM^{2}AG

November 2022

Course objectives

After following this course, you will be able to:

- **know** the scientific foundations of InfoVis;
- **analyze** data sets using visualization techniques; and
- **build** visualization that convey information and ideas.

Planning

18 hours together (12 $\times$ 1.5h), lectures and practical works.
Six lectures, i.e., an introduction to the domain:

- Introduction, Human visual perception;
- The visualization pipeline, Data types, Seminal works;
- Trees and graph visualization;
- Tabular data and time series;
- InfoVis toolkits; and
- Visualization evaluation.

Online ressources

Home page for the class:

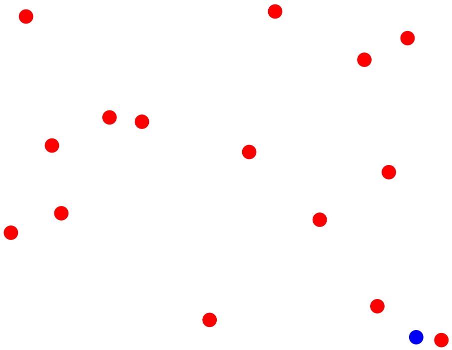
<https://mosig.imag.fr/ClassNotes/UIS-IV>

Problem

Q: What is the **best channel to convey information** to a human?
A: **Vision** because:

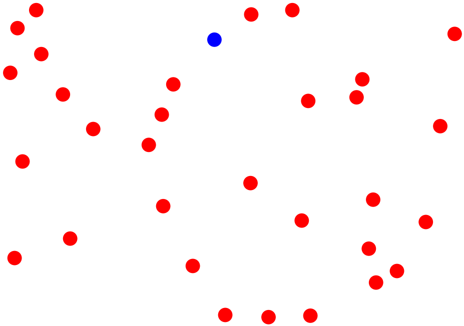
- **highest bandwidth** sense ($\approx 100\,\mathrm{MBs}^{-1}$, then ears $< 100\,\mathrm{bs}^{-1}$);
- **extends** memory and cognition;
- people **think visually.**

"Pre-attentive" perception



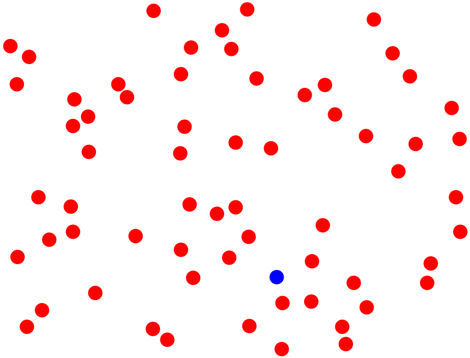
Find the **blue** dot…

"Pre-attentive" perception (cont.)



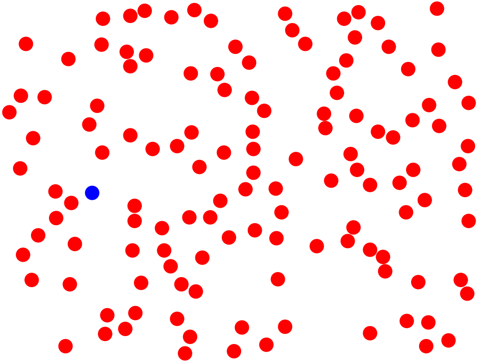
Find the **blue** dot. . .

"Pre-attentive" perception (cont.)



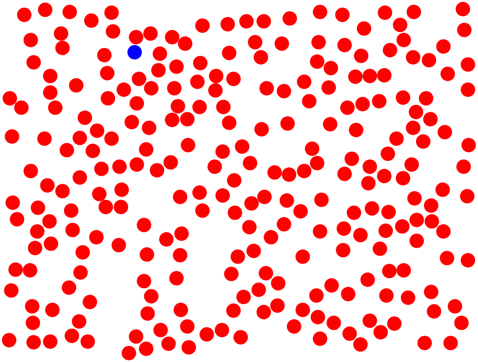
Find the **blue** dot. . .

"Pre-attentive" perception (cont.)



Find the **blue** dot…

"Pre-attentive" perception (cont.)



...in **constant time**, no matter the number of red dots!

Anscombe's Quartet (1973)

Four **data sets**:

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.50
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Anscombe's Quartet (1973) (cont.)

…having the exact **same statistical properties**:

- number of observations (n): 11
- mean of the x's ($\bar{x}$): 9.0
- mean of the y's ($\bar{y}$): 7.5
- equation of regression line: $y = 3 + 0.5x$
- sums of squares of $x - \bar{x}$: 110.0
- regression sums of squares: 27.50 (1 d.f.)
- residual sums of squares of y: 13.75 (9 d.f.)
- Multiple R^2: 0.667

Anscombe's Quartet (1973) (cont.)

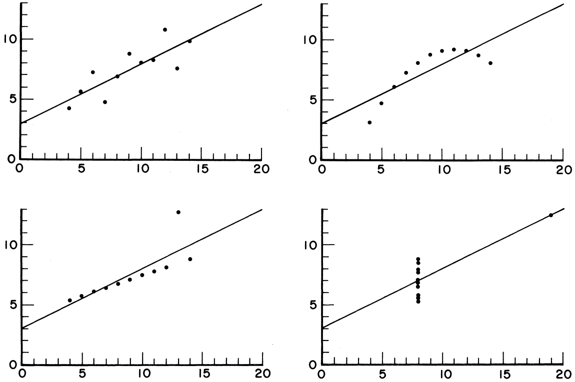


figure 1 – Anscombe's quartet plotted.

Visual thinking

"A picture is worth a thousand words"

— anonymous, 1911

"Un petit dessin vaut mieux qu'un long discours"

— Napoléon Bonaparte, 18xx

Visual thinking (cont.)

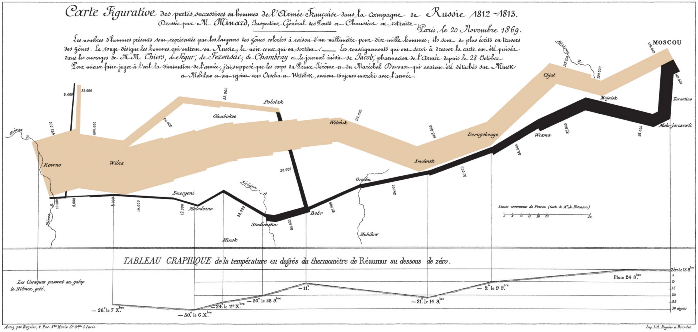


figure 2 – Charles Minard's 1869 chart showing the number of men in Napoleon's 1812 Russian campaign army, their movements, as well as the temperature they encountered on the return path.

Communication

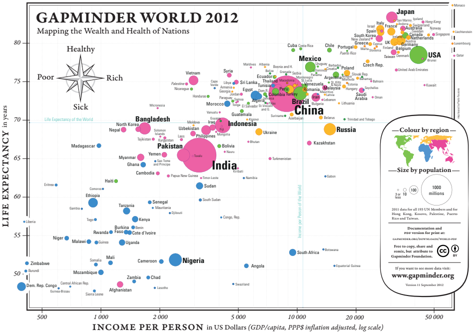


figure 3 – Hans Rosling's <Gapminder world> (e.g. showing <new insights on poverty>).

Information Visualization

InfoVis

*"The use of **computer-supported, interactive visual representations** of data to **amplify cognition**"*

— Card, Mackinlay & Shneiderman, 1998

Many fields are involved:

- **graphics** (millenniums of history)
- **cognitive psychology** (centuries of history)
- **Human-computer interaction** (decades of history)

Scientific Visualization

SciViz
Visualization of data sets captured from real world, having a **given spatialization.**

Key differences:

- continuous math vs. discrete math
- limited set of application domains
- smaller design space

Data Visualization

DataVis

No established definition yet:

- often assumed to use **web-based** technologies
- the **communication** part of visualisation

Challenges

- **diversity**: what is *information*?
- **scale**: what is a *large* dataset?

The Visual Information-Seeking Mantra

The **Visual Information-Seeking Mantra** [Shneiderman, 1996]:

- **Overview first**,
- **Zoom and filter**, then
- **Details-on-demand.**

- Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proc. Visual Languages, 336–343, 1996.

Simplified model

Visual perception is a **two stage process**:

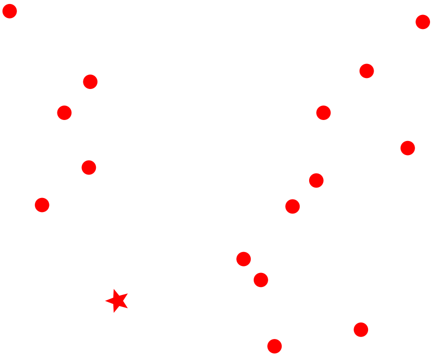
- a **parallel** extraction of **low-level** properties; then
- a **sequential goal-directed** processing.

Pre-attentive processing

Parallel processing by the retina (bottom-up) of:

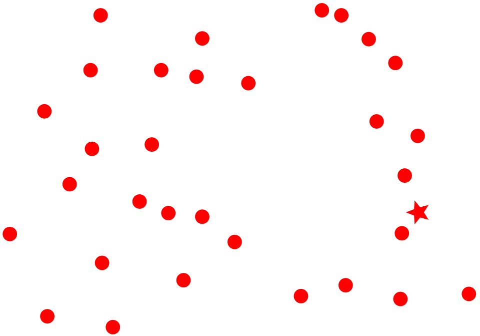
- orientation;
- color;
- texture;
- movement;
- etc.

Pre-attentive processing (cont.)



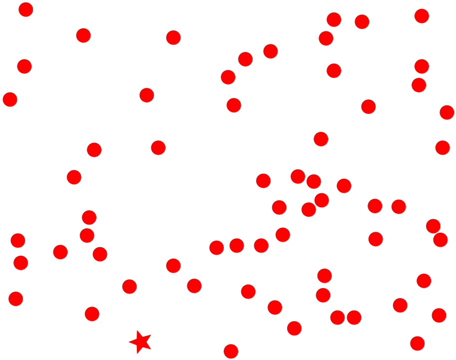
Find the **star** dot...

Pre-attentive processing (cont.)



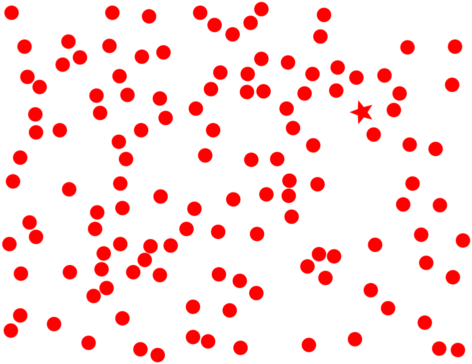
Find the **star** dot. . .

Pre-attentive processing (cont.)



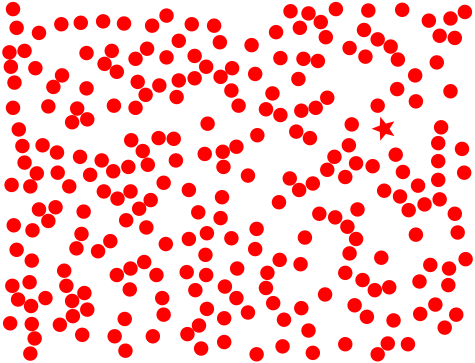
Find the **star** dot. . .

Pre-attentive processing (cont.)



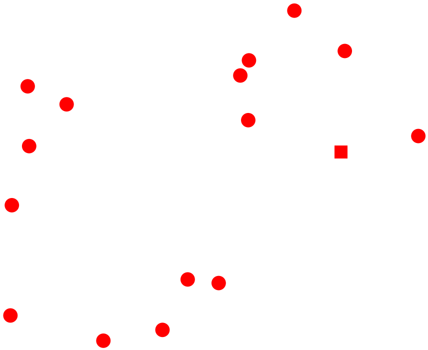
Find the **star** dot...

Pre-attentive processing (cont.)



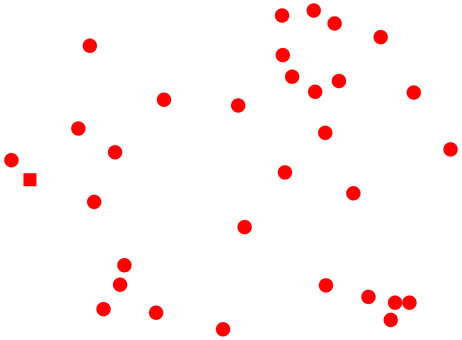
. . . in **constant time**, no matter the number of round dots!

Signal detection



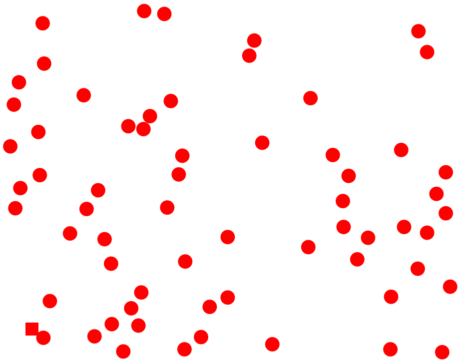
Find the **square** dot. . .

Signal detection (cont.)



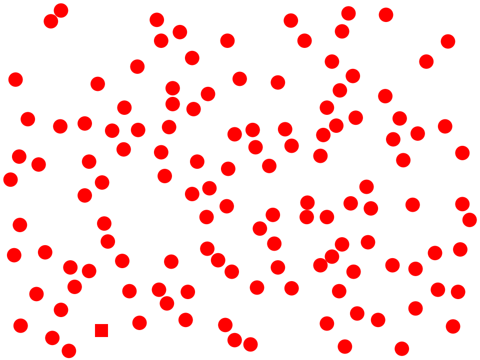
Find the **square** dot. . .

Signal detection (cont.)



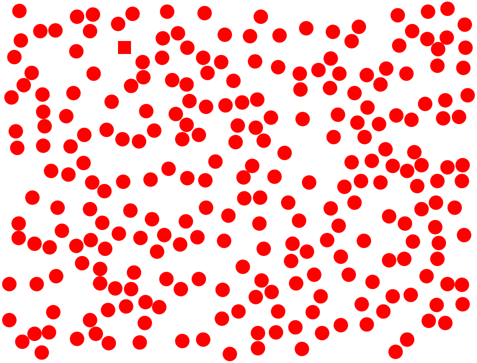
Find the **square** dot. . .

Signal detection (cont.)



Find the **square** dot. . .

Signal detection (cont.)



. . . in **constant time**, no matter the number of round dots, really?

Signal detection (cont.)

The **Feature-Integration Theory of Attention** [Treisman & Gelade, 1980] can be seen as a **limit case** of **Visual Search and Attention: a Signal Detection Theory Approach** [Verghese, 2001].

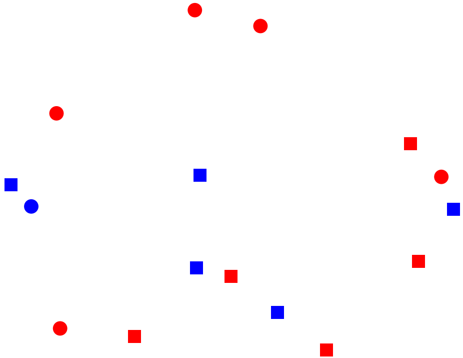
- A. Treisman and G. Gelade. A Feature-Integration Theory of Attention. In Cog. Psycho., vol. 12, 97–136, 1980.
- Preeti Verghese. Visual Search and Attention: a Signal Detection Theory Approach. In Neuron, vol. 31, 523–535, 2001.

Goal-directed processing

Sequential processing by upper level in the brain (top-down):

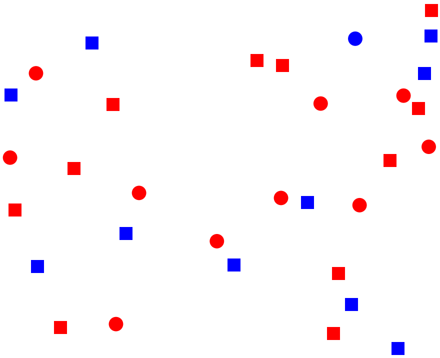
- object segmentation;
- object identification;
- etc.

Goal-directed processing (cont.)



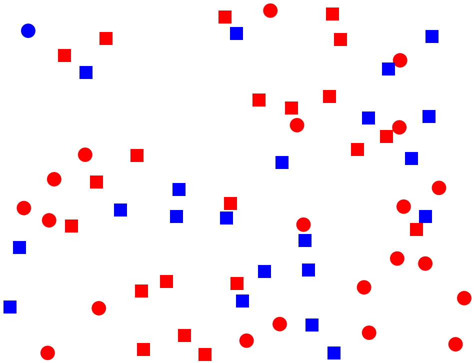
Find the **blue** dot. . .

Goal-directed processing (cont.)



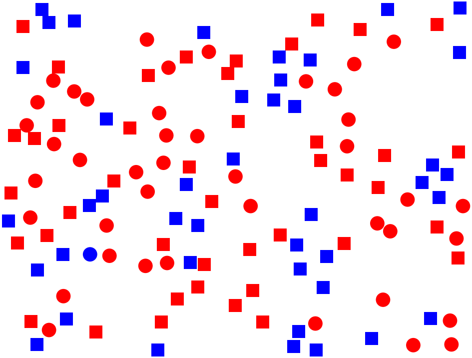
Find the **blue** dot. . .

Goal-directed processing (cont.)



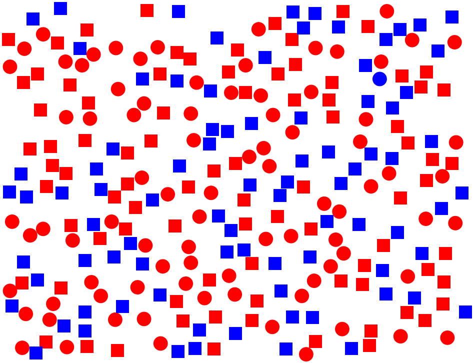
Find the **blue** dot…

Goal-directed processing (cont.)



Find the **blue** dot. . .

Goal-directed processing (cont.)



. . . in **linear time** if mixed with red/blue squares!

Eye anatomy

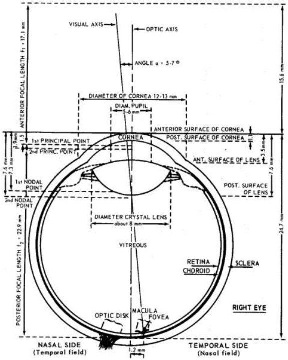


figure 4 – Right eye.

Acuity

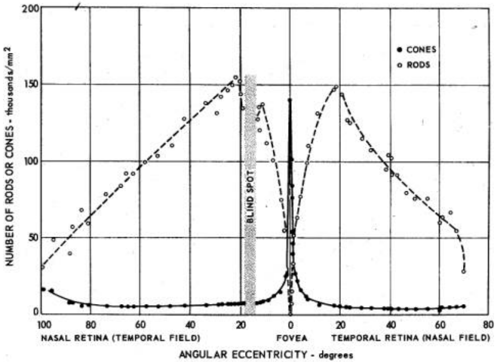


figure 5 – Density of receptors.

Visual field

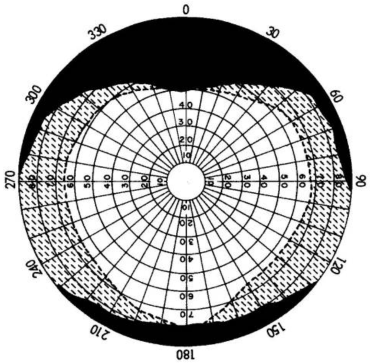


figure 6 – Binocular visual field.

Summary

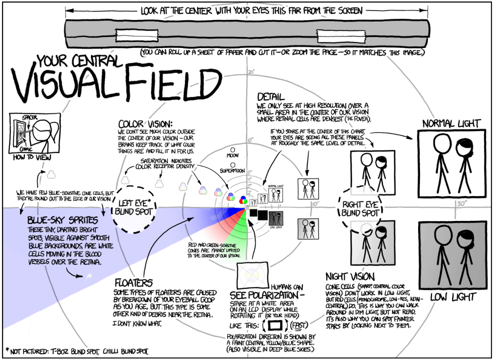


figure 7 – Your central visual field <http://xkcd.com/1080/>.

Gestalt psychology (192X)

figure 8 – "My wife and my mother-in-law." (1915)

Gestalt psychology (192X) (cont.)

The *Gestalt* psychology is a **theory of perception** that is often summed up by:

> *"The whole is other than the sum of the parts"*
>
> — *Kurt Koffka (1922)*

Gestalt psychology (192X) (cont.)

The *Gestalt* psychology notably describes the **perception of forms** by the visual system. It relies on four principles:

- **Emergence**;
- **Reification**;
- **Multistability**; and
- **Invariance.**

It also describes our visual perceptions by a **set of laws.**

Emergence

figure 9 – emergence (picture from [Gregory, 1970])

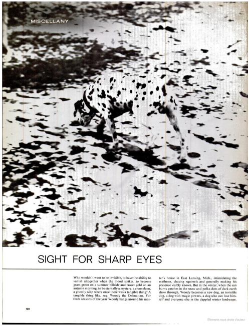
• R. L. Gregory. The intelligent eye. McGraw-Hill, New York, 1970.

Emergence (cont.)

Emergence

The **global perception** can **not** be **explained by** the sum of **its parts**.

Emergence (source)


MISCELLANY

SIGHT FOR SHARP EYES

Who wouldn't want to be invisible, to have the ability to vanish altogether when the mood strikes, to become grass green on a summer hillside and russet gold on an autumn morning, to be eternally a mystery, a chameleon, a ghostly wisp where once there was a tangible thing? A tangible thing like, say, Woody the Dalmatian. For three seasons of the year Woody hangs around his master's house in East Lansing, Mich., intimidating the mailman, chasing squirrels and generally making his presence visibly known. But in the winter, when the sun burns patches in the snow and polka dots of dark earth show through, Woody becomes a new dog, an invisible dog, a dog with magic powers, a dog who can lose himself and everyone else in the dappled winter landscape.

120

Éléments sous droits d'auteur

figure 10 – original picture [James, 1965]

- Ronald C. James. Life Magazine, 19 feb. 1965.

Emergence (source) (cont.)

figure 11 – The Hallucinogenic Toreador [Dali, 1970]

- Gala Salvador Dali. The Hallucinogenic Toreador, 1970.

Reification

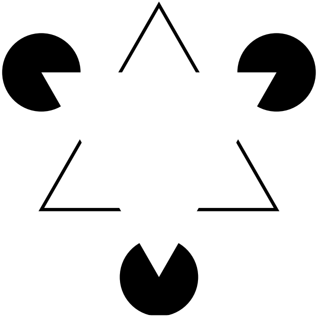


figure 12 – reification

Reification (cont.)

Reification

The **perception** contains **more** spatial **information that the stimulus** on which it is based: part of the perception is generated.

Multistability

figure 13 – multistability

Multistability (cont.)

Multistability

Ambiguous stimuli can generate different perceptions but they can not coexist simultaneously.

Invariance

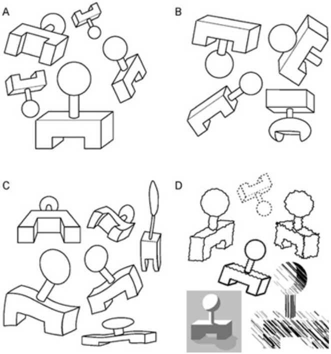


figure 14 – invariance

Invariance (cont.)

Invariance
Objects are recognized independently of various variations, such as geometrical transformations, lighting, etc.

Gestalt laws of grouping

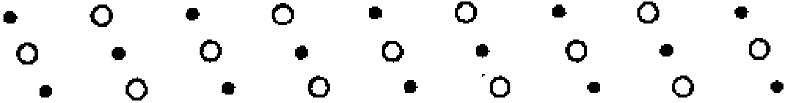


figure 15 – grouping of dots (illustration from <Laws of Organization in Perceptual Forms (1923)>).

The **laws of grouping** state how **low-level perceptions** are grouped **into higher-level objects**.

Gestalt laws of grouping (cont.)

Good *Gestalt* (*Prägnanz*)

We tend to order our experience in a manner that is regular, orderly, symmetric, and simple.

Proximity

Objects that are close tend to be perceived as a group.

Similarity

Objects that are similar (in shape, color, shading, etc.) tend to form a group.

Closure

The perception fills gaps in stimuli.

Gestalt laws of grouping (cont.)

Symmetry

Objects with symmetric disposition tend to be perceived as forming a whole.

Common Fate

Objects evolving together are perceived as a group.

Continuity

Ambiguous stimuli are perceived preferentially with the interpretation that is the most continuous.

Past Experience

We group things we have learned to group (e.g. letters)

The Information Visualization Pipeline

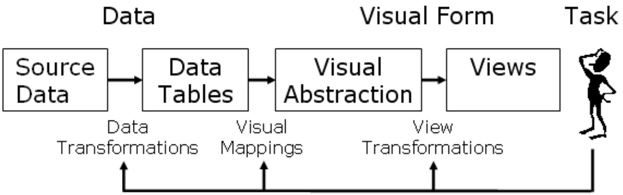


figure 16 – InfoVis pipeline [Chi & Riedl, 1998].

- E. Chi and J. Riedl. An Operator Interaction Framework for Visualization Systems. In proc. InfoVis'98, 63–70, 1998.

Data tables

Data sets are to be organized into tables with:

- **attributes** (or dimensions, . . .) as columns
- **items** (or observations, subjects, . . .) as rows

Visual mapping

Each data **item** will be projected onto a **mark** in the visualization space.
The mark will have its appearance set according to the data **attributes**.

Data

Several taxonomies of **data types**, e.g., [Card & Mackinlay, 1997]:

- **Nominal**
- **Ordered**
- **Quantitative** (differences, ratios)

differing by how you can **compare** and **aggregate** values.

- S. Card and J. Mackinlay. The Structure of the Information Visualization Design Space. In proc. InfoVis'97, 92–99, 1997.

Nominal data

examples

- names, gender, country, ID number . . .

comparisons

- (not) equal

aggregations

- frequency distribution, mode
- groups if there is a taxonomy over the values

Ordered data

examples

- sizes (XS, S, M, L, XL), ranks, grades on non-numeric scales ...

comparisons

- (not) equal, (not) less than

aggregations

- histogram, quantiles, median

Quantitative data

examples

- interval: date, temperatures (ºC, F), coordinates ...
- ratio: size, weight, population ...

comparisons

- (not) equal, (not) less than, magnitude of the difference

aggregations

- distribution, moments (mean, variance, etc.)

Graphic variables

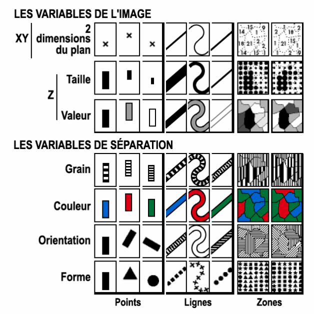


figure 17 – Graphic variables [Bertin, 1967].

- Jacques Bertin. Sémiologie graphique. 1967.

Graphic judgements

Association (≡)

Allows to represent different levels without inducing a visual hierarchy between those levels, and thus not interfering with the perception of other variables.

Graphic judgements (cont.)

Selection ($\neq$)

All items sharing a level (e.g. same color) can be perceived as a group and considered for inspection without the items of other levels interfering.

Graphic judgements (cont.)

Order (O)

Two items can be ordered according to this variable, without relying on a lookup to a legend.

Quantity (Q)

The difference between two items can be quantified.

Position

Position ($\equiv, \neq, O, Q$)
can be parametrized using various coordinates systems:

- **Cartesien** (x, y)
- **polar** (ρ, ϕ)
- hyperbolic plane, log scales, etc.

Size

Size ($\neq, O, Q$)

provides Q but beware of Stevens'Law:

$$\frac{p(x_1)}{p(x_0)} = \left(\frac{x_1}{x_0}\right)^\beta$$

- $\beta \in [0.9, 1.1]$ for 1D (length)
- $\beta \in [0.6, 0.9]$ for 2D (area)
- $\beta \in [0.5, 0.8]$ for 3D (volume)

A note on depth

Depth

is perceived mainly because it impacts (apparent) **size.**

Using the **third dimension** introduces **ambiguity.**
Something can **appear small** for two causes that can mix up:

- it is **far**, or
- it has actually a **small size.**

Value

Value ($\neq$, O)
is the **lightness** of the color (e.g. the gray level with the same brightness).

Texture

Texture ($\equiv, \neq, O$)

is the scale of a fixed pattern.

This is the only variable that provides both:

- **association** ($\equiv$); and
- **order** (O).

Color

Color ($\equiv, \neq$)

is the **hue** part of the color, i.e. the tint, independently of the lightness.

A note on color

Color is a 3D space, with different parametrizations.
The **color opponent process model** [Ware, 2000] is the most "psychophysical":

- **black-white** (luminance $=$ red $+$ green)
- **red-green**
- **yellow-blue** (luminance $-$ blue)

• C. Ware. Information visualization. 2000.

Orientation

Orientation $(\equiv, \neq)$
is the **angle** of the mark (modulo shape symetries).

Shape

Shape (≡)
is the **form** of the mark.
This is the only variable providing **association** (≡) that do not provide **selection** (≠).

Guidelines for mapping

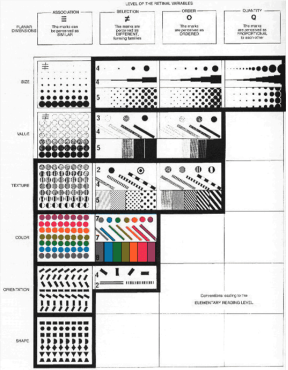


figure 18 – Variable properties [Bertin, 1967].

Guidelines for mapping (cont.)

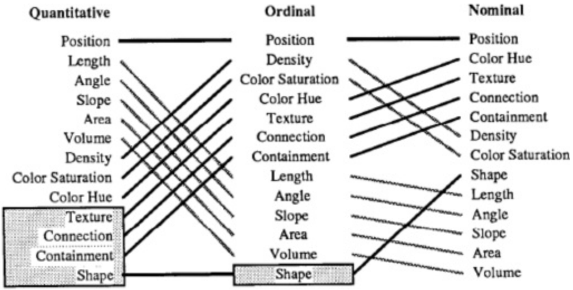
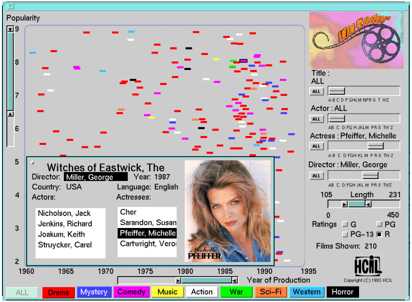


figure 19 – Suitability of variables [Mackinlay, 1986].

- J. Mackinlay. Automating the Design of Graphical Presentations of Relational Information. ACM Trans. Graph. 5(2): 110–141, 1986.

Design space for mapping



Name	D	F	D'	X	Y	Z	T	R	—	[]	CP
Year	Q	>	Q	P							
Quality	Q	>	Q		P						
Type	N	>	N					C			
Title	O	sl>									
Actor	O	sl>									
Actress	O	sl>									
Director	O	sl>									
Length	Q	br>									
Rating	N	br>									

figure 20 – Characterization of Film finder [Card & Mackinlay, 1997].

A tour through the visualization zoo

Let’s take a `<tour through the Visualization zoo>` [Heer et al., 2010]!

- J. Heer, M. Bostock, V. Ogievetsky. A Tour Through the Visualization Zoo. Communications of the ACM 53(6): 59–67, 2010.

Multi-dimensional data

Data attributes

- **quantity** $q1$
- **quantity** $q2$
- **quantity** $q3$
- **quantity** $q4$
- ...

Problem

We do **not** have **enough** Q **variables**.

Multi-dimensional data (cont.)

Solutions

- scatter plot matrix
- parallel coordinates [Inselberg, 1985]
- ...

Idea

The two visualizations **reuse axes** in different ways.

- A. Inselberg. The Plane with Parallel Coordinates. Visual Computer. 1(4): 69–91, 1985.

Scatter plot matrix (SPLOM)

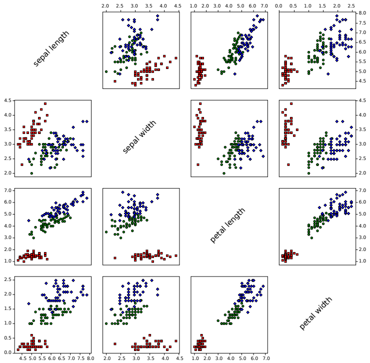


figure 21 – Scatter plot matrix visualisation of the Iris dataset.

Parallel coordinates

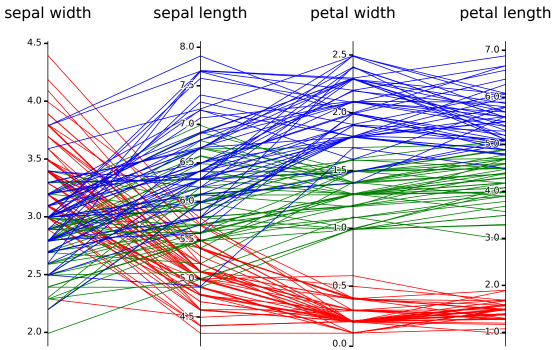


figure 22 – Parallel coordinates visualisation of the Iris dataset.

Time series

Data attributes

- **timestamp** t
- **quantity** q

encoded with:

Graphic variables

- **X** position
- **Y** position

Horizon Graph

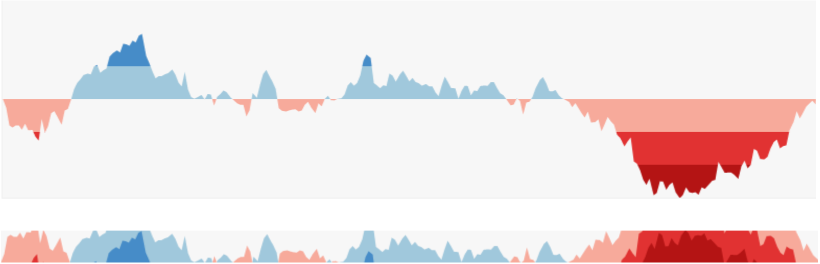


figure 23 – Horizon Graph visualisation technique for time series, adapted from [Reijner, 2005].

- H. Reijner. The Development of the Horizon Graph.
  proc. VisWeek workshop "From theory to Practice", 2005.

Horizon Graph (cont.)

Encoding

- $t \rightarrow \text{X}$
- sign $\boldsymbol{q} \rightarrow$ color
- $|\boldsymbol{q}|$ div $V \rightarrow \text{Y}$
- $|\boldsymbol{q}| \bmod V \rightarrow$ value

Composite visual mapping

Composite Visual Mapping

A **visual mapping** is said **composite**
if (at least) **one data attribute**
is mapped onto (at least) **two graphic variables**
[Jabbari et al., 2018]**.**

- A. Jabbari, R. Blanch, S. Dupuy-Chessa Composite Visual Mapping for Time Series Visualization In proc. of PacificVis 2018, 116-124, 2018.

Structured data

More zoos:

- a `<tree visualization reference>` [Schulz, 2011];
- a `<survey on set visualization>` [Alsallakh et al., 2014];

- H. Schulz. Treevis.net: a Tree Visualization Reference.
  IEEE Computer Graphics and Applications 31(6): 11–15, 2011.
- B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, P. Rodgers.
  Visualizing Sets and Set-typed Data: State-of-the-Art and Future Challenges.
  In proc. EuroVis'14, State of The Art Reports.
  .

Structured data (cont.)

- a <survey on multifaceted scientific data visualisation> [Kehrer & Hauser, 2013];
- a <survey of text visualization techniques>;

- J. Kehrer, H. Hauser.
  Visualization and Visual Analysis of Multifaceted Scientific Data: a Survey.
  IEEE Transactions on Visualization and Computer Graphics 19(3): 495–513, 2013.

Time dependent data

Even more Zoos:

- a `<visual survey of visualization techniques for time-oriented data>` [Aigner et al., 2011];
- a `<review of temporal visualizations based on space-time cube operations>` [Bach et al., 2014];

- W. Aigner, S. Miksch, H. Schumann, C. Tominski.
  Visualization of Time-Oriented Data. 2011.
- B. Bach, P. Dragicevic, D. Archambault, C. Hurter, S. Carpendale.
  A Review of Temporal Data Visualizations Based on Space-Time Cube Operations.
  In proc. EuroVis'14, State of The Art Reports.
  .

Time dependent data (cont.)

- a `<state of the art in visualizing dynamic graphs>` [Beck et al., 2014].

- F. Beck, M. Burch, S. Diehl, D. Weiskopf. The State of the Art in Visualizing Dynamic Graphs. In proc. EuroVis'14, State of The Art Reports.

Surveys

And zoos of zoos!

- a <visualization of visualization tools and books>

The Information Visualization Pipeline (bis)

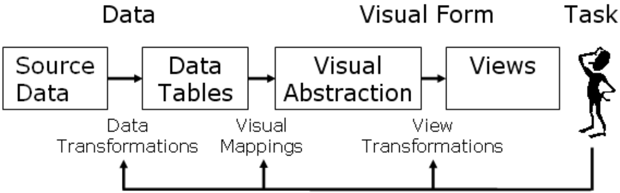


figure 24 – InfoVis pipeline [Chi & Riedl, 1998].

Multi-dimensional data

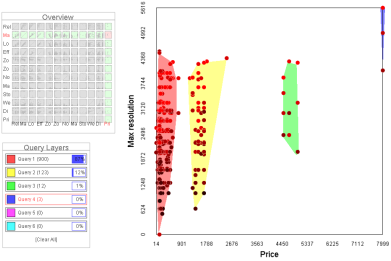


figure 25 – <ScatterDice> [m4v] [Elmqvist et al., 2008].

- N. Elmqvist, P. Dragicevic, J.-D. Fekete. Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation.
  In Proc. InfoVis 2008, 1141-1148, 2008.

Zoomable treemaps

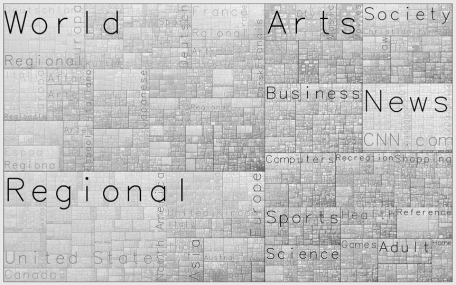


figure 26 – `<Zoomable Treemaps>` `[mov]` [Blanch & Lecolinet, 2007].

- R. Blanch and É. Lecolinet. Browsing Zoomable Treemaps: Structure-Aware Multi-Scale Navigation Techniques. In Proc. of InfoVis 2007, 1248–1253, 2007.

Hybrid visualization of graphs

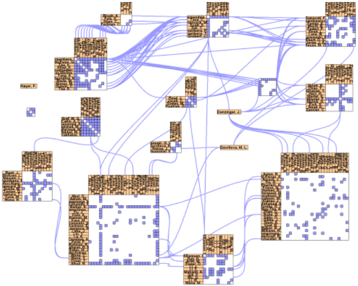


figure 27 – `<NodeTrix>` `[mov]` [Henry et al., 2007].

- N. Henry, J.-D. Fekete, M. J. McGuffin. NodeTrix: A Hybrid Visualization of Social Networks. In Proc. of InfoVis 2007, 1302-1309, 2007.

Hybrid tree/matrix visualizations

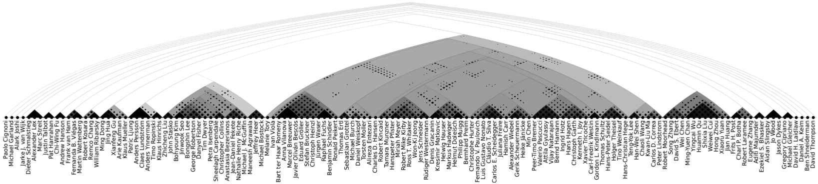


figure 28 – `<Dendrogramix>` `[mov]` [Blanch et al., 2015].

- R. Blanch, R. Dautriche and G. Bisson. Dendrogramix: a Hybrid Tree-Matrix Visualization Technique to Support Interactive Exploration of Dendrograms. In Proc. of PacificVis 2015, 31-38, 2015.

Thank you!

Thank you for your attention!