Interactive Visual Analysis with different levels of complexity

Helwig Hauser, Delft, 2010-06-24
Introduction

- **Vis.** good for exploration/analysis & presentation
  - exploration: visualize to detect the unexpected
  - analysis: check hypotheses with visualization
  - presentation: show/communicate what you’ve found

- **Focus** here: interactive visual analysis, IVA

- **Goal** today: understanding IVA
  - levels of IVA
  - patterns of IVA
  - …

- Personal background: ≈10 years of IVA research
  - SimVis, IVA of …, etc., applications, …
  - VisMaster
Basis: Data Model, domain & range

- **Data model** \(d(x,t)\), independent/dependent variables

- **Domain**: space \(x\), time \(t\) (and ...)
  - where? when?
  - 2D/3D space
    1D time
  - parameters

- **Range**: attributes \(d_i\)
  - what?
  - several–many \(d_i\) (can be dozens)
  - often somehow coherent across space-time locations (continuous, distributions)

- **Example**:
  - for each of one million grid cells
    at each of hundreds of time steps
  - the simulated values of temperature, pressure, flow velocity, etc., are given
Interactive Visual Data Analysis

Main idea:
- enabling interactive \& iterative exploration / analysis of complex (multi-variate) and often also large data
- allowing for visual information drill-down, cf. Ben Shneiderman’s visual information seeking mantra
- built around a notion of the user’s interest – subjective \& current degree of interest per data item

Different “levels” of IVA:
- show \& brush (tightest IVA loop)
- relational analysis (“reading between the lines”)
- complex analysis (“joining forces”)
- and more …

... an example of IVA (level 1) first ...
First Example: IVA of Simulation Data

- Important preliminary:
  - multiple views onto the data (here flow simulation data)
  - often at least one for domain variables (here 3D space) and one for range variables (here pressure & velocity)

... this is the initial visualize step ...
First Example: Linking & Brushing

Next: first IVA loop
- visualization leads to “I see (something)!” effect
- user brushes “this something” (literally!)
- linked visualization reveals insight!

... this is the 1st brushing (&...) step ...

highlighted: fast, rather high pressure
Show & Brush (IVA level 1)

- **Tightest IVA loop**
  - show data (explicitly represented information)
  - one brush (on one view, can work on >1 dims.)

A typical (start into an) IVA session of this kind:
- bring up multiple views
  - at least one for x, t
  - at least one for d_i
- I see (something)!
- brush this “something”
- linked F+C visualization
- first insight!
Show & Brush

- Tightest IVA loop
  - show data (explicitly represented information)
  - one brush (on one view, can work on >1 dims.)

- Requires:
  - multiple views (≥2)
  - interactive brushing capabilities on views (brushes should be editable)
  - focus+context visualization
  - linking between views

- Allows for different IVA patterns (wrt. domain & range)

A typical (start into an) IVA session of this kind:
- bring up multiple views
  - at least one for $x, t$
  - at least one for $d_i$
- I see (something)!
- brush this “something”
- linked F+C visualization
- first insight!

... leads to...
... requires...
... is realized via ...

degree of interest

(next slide)
IVA: Multiple Views

- One dataset, but multiple views
- Scatterplots, histogram, 3D(4D) view, etc.

[Doleisch et al., '03]
IVA: Interactive Brushing

- Move/alter/extend brush interactively
- Interactively explore/analyze multiple variates

[Doleisch et al., '03]
IVA: Interactive Brushing

- Move/alter/extend brush interactively
- Interactively explore/analyze multiple variates

[Doleisch et al., ’03]
IVA: Interactive Brushing

- Move/alter/extend brush interactively
- Interactively explore/analyze multiple variates

[Doleisch et al., '03]
IVA: Interactive Brushing

- Move/alter/extend brush interactively
- Interactively explore/analyze multiple variates

[Doleisch et al., ’03]
IVA: Interactive Brushing

- Move/alter/extend brush interactively
- Interactively explore/analyze multiple variates

[Doleisch et al., ’03]
IVA: Focus+Context Visualization

- Traditionally space distortion
  - more space for data of interest
  - rest as context for orientation

- Generalized F+C visualization
  - emphasize data in focus (color, opacity, …)
  - differentiated use of visualization resources

[Hauser... 2001, 2003]

Alternatives...
F+C Visualization in IVA Views

- Colored vs. gray-scale visualization
- Opaque vs. semi-transparent visualization

In a scatterplot (left) or histogram (right): brushed data in red...

[Matković et al., ’09]
F+C Visualization in IVA Views

[Novotný & Hauser, ’06]

In parallel coordinates (above): brushed data in red & over ...
F+C Visualization in IVA Views

[Muigg et al., ’07]

In 3D (above): less transp. & colored, in illustrative context ...
IVA: Linked Views

- **Brushing**: mark data subset as especially interesting
- **Linking**: enhance brushed data in linked views consistently (F+C)

[Doleisch & Hauser, ’02]
IVA: Degree of Interest (DOI)

- $\text{doi}(.)$: data items $tr_i$ (table rows) → degree of interest
  $\text{doi}(tr_i) \in [0, 1]$
  - $\text{doi}(tr_i) = 0 \Rightarrow tr_i$ not interesting ($tr_i \in \text{context}$)
  - $\text{doi}(tr_i) = 1 \Rightarrow tr_i$ 100% interesting ($tr_i \in \text{focus}$)

- Specification
  - explicit, e.g., through direct selection
  - implicit, e.g., through a range slider

- Fractional DOI values: $0 \leq \text{doi}(tr_i) \leq 1$
  - several levels (0, low, med., …)
  - a continuous measure of interest
  - a probabilistic definition of interest

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>$d1$</th>
<th>$d2$</th>
<th>$\text{doi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>17.20</td>
<td>-0.22</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>12.10</td>
<td>0.10</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>7.70</td>
<td>0.45</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2.10</td>
<td>0.90</td>
<td>0.00</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>24.10</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>21.90</td>
<td>0.36</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>15.50</td>
<td>0.67</td>
<td>0.74</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>11.10</td>
<td>1.29</td>
<td>1.00</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>27.20</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>24.10</td>
<td>0.67</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>17.30</td>
<td>1.35</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>12.10</td>
<td>2.20</td>
<td>0.60</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>35.50</td>
<td>0.67</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>30.90</td>
<td>1.30</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>24.50</td>
<td>2.10</td>
<td>0.10</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>20.80</td>
<td>2.90</td>
<td>0.00</td>
</tr>
</tbody>
</table>

(cont’d on next slide)
IVA: Smooth Brushing → Fractional DOI

- **Fractional DOI values** esp. useful wrt. **scientific data**: (quasi-)continuous nature of data ↔ smooth borders

- Goes well with gradual focus+context vis. techniques (coloring, semitransparency)

- **Specification: smooth brushing** [Doleisch & Hauser, 2002]
  - “inner” range: all 100% interesting (DOI values of 1)
  - between “inner” & “outer” range: fractional DOI values
  - outside “outer” range: not interesting (DOI values of 0)
Fuzzy Classification

DOI \in [0,1] – 0 … not interesting

\begin{align*}
R & = a \land b \quad \text{if combination, we use} \\
& = \min(a, b) \\
R & = a \lor b \quad \text{if combination, we use} \\
& = \max(a, b) \\
R & = \neg a \quad \text{such as} \\
& = 1 - a
\end{align*}

Matches the smooth nature of the data

Goes well with \( + \)

\[ \text{opacity varies gradually with DOI} \]

2002 Helwig Hauser

http://www.VRVis.at/

SimVis: Interactive Visual Analysis of Simulation Data

http://www.SimVis.at/
Fuzzy Classification

- DOI \( \in [0,1] \) – 0 … not interesting

\( R_{fuzzy} \) for combination, we use

\[ c = a \land b \iff c = \min(a, b) \]

\[ c = a \lor b \iff c = \max(a, b) \]

\[ \neg a \iff c = 1 - a \]

- Matches the smooth nature of the data

- Goes well with FCG visualization, e.g.,

Opacity varies gradually with DOI

Helwig Hauser 2002

http://www.VRVis.at/

http://www.SimVis.at/
Three Patterns of IVA

1. Preliminary: domain \( x \) & range \( d \) visualized (≥2 views)

   - **Brushing on domain visualization**, e.g., brushing special locations in the map view

2. **Brushing on range visualization**, e.g., brushing outlier curves in a function graph view

3. Relating multiple range variates

   - **Multi-variate analysis**
Traffic Sensor Network
(Minneapolis, St. Pauls)

- 12 weeks (84 days)
- 564 sensors
- daily data, aggregated from measurements all 30 secs.
- \(2 \cdot 47376 = 94752\) graphs
  (144 \(f(t)\)-values each, one per 10 mins.)
... 2 sensors, 2 days (Sun+Mon)!
… 293 outliers (out of 47376 ≈ 0.6%)!

… negative(!) volume-values brushed...
... again the same sensor!

... just 1 outlier!

[TVCG 2006]
whole day no cars?!

why (just) here???
A Layered Information Space

- Metaphor of a “sea of information”
- Explicitly represented information (the data) on top, implicitly represented information below (in layers)

show & brush

user, task

data

relational information

complex information

features in application terms

temp.   vel.   ...
Relational Analysis

- Extended IVA loop
  - iterative & relational exploration/analysis
  - read between the lines (implicitly represented inform.)

A typical continuation of an IVA session:
- bring up more views
- add/combine brushes
- focus the analysis, drill deeper!

Requires:
- multiple views & selections
- feature definition language
- show combination of multiple levels of F+C views.

All allows drill-down information...
Relational Analysis (IVA level 2)

- Extended IVA loop
  - iterative & relational exploration/analysis
  - read between the lines (implicitly represented inform.)

- Requires:
  - multiple views
  - feature definition language
  - multiple levels of F+C vis.

A typical continuation of an IVA session:
- bring up more views
- add/combine brushes
- focus the analysis, drill deeper!

... the example of IVA again (now level 2) ...

- Allows deep(er) information drill-down
IVA: Feature Definition Language

- Explicitly represented degree of interest (DOI)
  DOI: additional (synthetic) data dimension(s)

- Brushing results in DOI attribution(s), relational analysis through multiple brushes

- Tree structure through logical operators
  - root, level 0 (OR node): set of features
  - level 1 (AND node(s)): feature specifications
  - ... (individual brushes)

- Compare to:
  - natural language
  - DB query

- In/out: XML

Example:

interesting are ... flow regions where pressure is high AND velocity is high
IVA: Four Levels of F+C Visualization

- **In show & brush:** one brush (focus), rest is context
  - data in focus: colored, less transparent
  - data in context: gray-scale, in background

- **In relational analysis:**
  - multiple features in a feature set (below top node)
  - multiple views define a feature

⇒ advanced F+C visualization
- here: three views, two (a, b) for one feature, one (c) for one other…

[Muigg et al., ’08]
IVA: Four Levels of F+C Visualization

- **gray**: overall context
- **green**: other features (in the set)
- **red**: feature defined with this view (and possibly with others, too)
- **yellow**: brushed in this view, but not part of feature (only in combination with other views)

advanced F+C visualization

- here: three views, two (a,b) for one feature, one (c) for one other

but how to color in combination with smooth brushing?

[Muigg et al., ’08]
IVA: Coloring Complex DOI Combinations

- How to combine colors, when fractional DOI values overlap?
- How to combine colors, when many data items make one pixel?

Simple color mixing (convex combination of RGB values) does not do it!

why brown??

0 . . . . . . . . . . . . . . . . . . . . 1
IVA: Coloring Complex DOI Combinations

- **Color:**
  - red over (green over (yellow over gray))
  - most important!

- **Blending weights** $\alpha_f$, $\alpha_s$, and $\alpha_c$ with
  - $\alpha_f$ = feature DOI
  - $\alpha_s$ = set DOI – feature DOI
    - // only the rest here!
  - $\alpha_c$ = max(comp. DOI – set DOI, 0)
    - // only non-hidden parts!
IVA Example (repeated from level 1)

- Two views (domain, range), one brush
  - simple feature localization
  - here: fast, mid-large pressure

focus on fast, rather high pressure
IVA Example, going level 2

- More views, more brushes, logical combinations

here: two foci

+ focus on rather slow flow
IVA Example, level 2

- Three brushes, complementary F+C vis., FDL

+ focus on high turbulence
Iterative Exploration/Analysis

SciVis (overview, orientation)
New View (InfoVis)
Adjust Projection (InfoVis)
Brush InfoVis-view
Linked Visualization
Alter Brush
Add Brush (→ compound brush)
Add View (InfoVis)
(presentation, report, …)

all steps
IVA: Iterative Exploration / Analysis

SciVis (overview, orientation)

New View (InfoVis)

Adjust Projection (InfoVis)

Brush InfoVis-view

Linked Visualization

Alter Brush

Add Brush (→ compound brush)

Add View (InfoVis)

(presentation, report, …)

the start
IVA: Iterative Exploration / Analysis

SciVis (overview, orientation)

New View (InfoVis)

Adjust Projection (InfoVis)

Brush InfoVis-view

Linked Visualization

- Alter Brush
- Add Brush (→ compound brush)
- Add View (InfoVis)

(presentation, report, …)

iterative refinement

main direction of work

analysis
IVA: Iterative Exploration / Analysis

SciVis (overview, orientation)
New View (InfoVis)
Adjust Projection (InfoVis)
Brush InfoVis-view
Linked Visualization
Alter Brush
Add Brush (→ compound brush)
Add View (InfoVis)
(presentation, report, …)
IVA: Iterative Exploration / Analysis

- SciVis (overview, orientation)
- New View (InfoVis)
- Adjust Projection (InfoVis)
- Brush InfoVis-view
- Linked Visualization
- Alter Brush
- Add Brush (→ compound brush)
- Add View (InfoVis)

(presentation, report, …)
Iterative Exploration/Analysis

SciVis (overview, orientation)

New View (InfoVis)

Adjust Projection (InfoVis)

Brush InfoVis-view

Linked Visualization

Alter Brush

Add Brush (compound brush)

all steps

but how to capture features which cannot be grasped through data values or logical combinations of them?

show & brush

relational analysis

presentation report
Joining forces:
- integrate computational analysis
- extend brushing

Boosting IVA:
- derive information
- advanced brushes
- access a new level of exploration/analysis!

Combination:
- show
- attribute derivation
- multiple views & selections
- advanced brushing
## Complex Analysis (IVA level 3)

- **Joining forces:**
  - Integrate computational analysis
  - Extend brushing

- **Builds upon:**
  - Advanced brushing
  - Attribute derivation

- **Boosting IVA:**
  - Derive information
  - Advanced brushes
  - Access a new level of exploration/analysis!

- Very powerful analysis / exploration mechanism!
IVA: Advanced Brushing

- Two ways to get more out of IVA:
  - bring the data to the interaction (attribute derivation)
  - bring the interaction to the data (advanced brushing)
    - angular brushing [Hauser et al., 2002]
    - similarity brushing [Muigg et al., 2008]

- Attribute derivation + advanced brushing = access to complex features
Comprehensible ways to derive synthetic data dimensions from original data
- data transformations
  - linear transformations
  - to log scale
  - etc.
- derivative information (against the domain variables)
  - $\frac{d d_i}{d x}$ – gradient information (wrt. space)
  - $\frac{d d_i}{d t}$ – change over time
- relative information
  - data normalization
  - differences, ratios
- model-related derivations
  - according to known relations, e.g., $\text{div} = \nabla \cdot \mathbf{v}$
Curve Sketching

- Understanding function graphs:
  - special values of \( f(x) \): zero, extremes, etc.
  - relative properties – positive/negative change \( f'(x) \)
    local maxima/minima – \( f'(x) = 0 \)
  - double-relative properties: the change of change
e.g., local maxima \( \iff f'(x) = 0 \) & \( f''(x) < 0 \)
infection point – \( f''(x) = 0 \)

- Remember your days in school:

http://www.nipissingu.ca/calculus/tutorials/curves.html
$\Delta$soot

soot$^{(0)}$ vs. soot$^{(1)}$

soot$^{(2)}$ vs. soot$^{(1)}$

$\Delta^2$soot

soot$^{(0)}$ vs. soot$^{(2)}$
t=10s
t=15s

smallest $\Delta^2$ soot only…
t=15s, smallest $\Delta^2$soot

- red: 0
- purple: -1
- blue: -2
- cyan: -3
- green: -4
- yellow: -5
Without slow changing...

red: 0;
purple: -1;
blue: -2;
cyane: -3;
green: -4;
yellow: -5;
wo/ slow changing, wo/ (almost) done..
t=30s, without slow changing

- red: 0
- purple: -1
- blue: -2
- cyan: -3
- green: -4
- yellow: -5
t=30s, without (almost) done

red: 0;
purple: -1;
blue: -2;
cyan: -3;
green: -4;
yellow: -5;
t=40s

quite „neg.“ Δsoot, slowest changing..
t=40s, “quite” negative $\Delta$soot

red: 0;
purple: −1;
blue: −2;
cyan: −3;
green: −4;
yellow: −5;
t=40s, slowest changing

red: 0;
purple: -1;
blue: -2;
cyan: -3;
green: -4;
yellow: -5;
$t=60s$

red: 0;
purple: -1;
blue: -2;
cyan: -3;
green: -4;
yellow: -5;
IVA: Attribute Derivation

- Further interesting opportunities
  - re-projecting the data
    - f.i. according to PCA
    - data shear
  - statistical analysis
    - moments of data subsets
    - data in relation to moments, e.g., z-score
  - scale-space repr.

- Attribute derivation
  + advanced brushing
  = access to complex features
Considering “scientific” data $f(x)$, i.e.,
- some measured/simulated(modeled) data $f$, e.g., $f$ being temperature, pressure, velocity, etc.,
- wrt. some domain $x$, with $x$ being 2D or 3D space, time, parameters, etc.

If $x$ is high-dimensional (>3), then “low-level” IVA is hard
- example: 100 runs of time-dependent 3D sim. data
- reducing the dimensionality can help

Means to reduce the dimensionality
- selection, e.g., through sampling
- aggregation, e.g., by averaging
- etc.
Integrating Statistics and IVA

- Statistics allow to assess distributional characteristics of sets of data, e.g., along one data dimension.

Examples:
- Map showing the average temperature in ten years
- Accumulated sea ice in summer 2008

Statistics can be reintegrated into IVA through attribute derivation:
- Mean, variance
- Median, 1\textsuperscript{st} & 3\textsuperscript{rd} quartile, IQR
- Min, max, min–max range
- Etc.
Example: multi-run climate simulation data

- 10 • 10 = 100 runs of time-dependent (250 time steps) ocean simulation (3 2D sections: Atlantic, Indian, Pacific)

Considering statistics wrt. the multiple runs

- derivation on demand
- visualization, e.g., glyph-based (↔)
- basis for complex analysis (next slide)
Brushing “Boxplots” [J. Kehrer et al., submitted]

- Analyzing outliers
  - derivation of IQR / (max–min)
    - large, i.e., ≈1: no outliers
    - small, i.e., ≈0: some far outliers
  - derivation of upper/lower outlier range,
    \[ UOR = \frac{(max–q_3)}{(max–min)} \]
    \[ LOR = \frac{(q_1–min)}{(max–min)} \]
    and \[ ULR = UOR–LOR \]
    - positive, if max far away
    - negative, if min far away
  - scatterplot of both and brushing

- The according IVA loop
  - show, derive, show, brush, ...
  - very powerful analysis approach
IVA beyond Complex Analysis (level 4)

- Of course there’s more:
  - approaches that “leave” the field
  - specialized feature extraction [Post et al., 2003]
  - etc.

- A lot of good literature available

- Much can be embedded within IVA, also!
  - code as field [Bürger et al., ’07]
Discussion of IVA Levels (1)

- Show & brush:
  - satisfies KISS principle
    - one brush
    - simple linking
    - conceptually simple
  - solves (maybe) 80% of all problems (Pareto rule)
  - implemented in many cases

Example:
Discussion of IVA Levels

- Show & brush
- Relational analysis:
  - coherent data / interaction metaphor space,
  - but allows for more complex queries
  - logical combinations match natural language

Example:
Discussion of IVA Levels

- Show & brush
- Relational analysis
- Complex analysis:
  - extends data / interaction metaphor space
    - additional data derivatives
    - additional brushing opportunities
- combination of computational and interactive analysis very powerful!

Example(s):
Discussion of IVA Levels

- Show & brush
- Relational analysis
- Complex analysis
- Approaches beyond...
  - “container” for approaches that go beyond DOI-based IVA

Really attractive to have all in one (IVA) framework!
Applications

- IVA of time-dep. 3D CFD data (engineering)
- IVA of industrial ensemble simulation data
- IVA of medical perfusion scans (3D+time)
- IVA of meteorological / climate research data
- IVA of sensor network data
- IVA of customer relation management (CRM) data
- ...

...
Conclusions

- IVA useful in many application scenarios
  - follows common patterns \((x \Leftrightarrow d, d \Leftrightarrow x, d \Leftrightarrow d)\)

- IVA enables a visual dialog with the data
  - from data to information / knowledge (and back)

- Iterative concept enables steered analysis
  - conquering the unexpected
    - both in terms of findings,
    - but also in terms of analysis approaches
  - facilitates reasoning, leads to additional learning,
    empowers the expert user (even makes experts!)

- IVA as useful exploratory research methodology
  - hypothesis generation
  - analysis prototyping
Acknowledgements

- **You!** Thank you for your attention! ;–)

- **Funding agencies, cooperating companies/…**, FFG, AVL List GmbH, EU(SemSeg, VisMaster), …

- **Many bright colleagues & students**
  Helmut Doleisch, Krešimir Matković, Robert Laramee, Raphael Fuchs, Philipp Muigg, Martin Gasser, Martin Ilčík, Johannes Kehrer, Matej Novotný, Matej Mlejnek, Ivan Viola, Zoltan Konyha, Markus Hadwiger, Steffen Oeltze, …

- **The Meister!**

**final plug: we are hiring!**

... see [www.ii.UiB.no/vis](http://www.ii.UiB.no/vis)!!