VizSEC 2007
Workshop on Visualization for Computer Security

Georgia State University
Department of Computer Science

Measuring the Complexity of Visualization for Computer Security

Xiaoyuan Suo, Ying Zhu, G. Scott Owen
Department of Computer Science
Georgia State University

Motivation

- We propose a method to analyze the complexity of a information visualization design
- The method evaluates a set of factors that influence the efficient comprehension of visual information
- The method is useful during the visualization design stage before any user study can take place

Case Studies

- Two open source security visualization software are used as case studies
  - TNV: http://tnv.sourceforge.net/
  - Rumint: http://www.rumint.org/
- The sample dataset is obtained from http://wiki.ethereal.com/CIGI

Outline

- Complexity Analysis
- Case Studies (Rumint)
- Conclusion and future work

Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

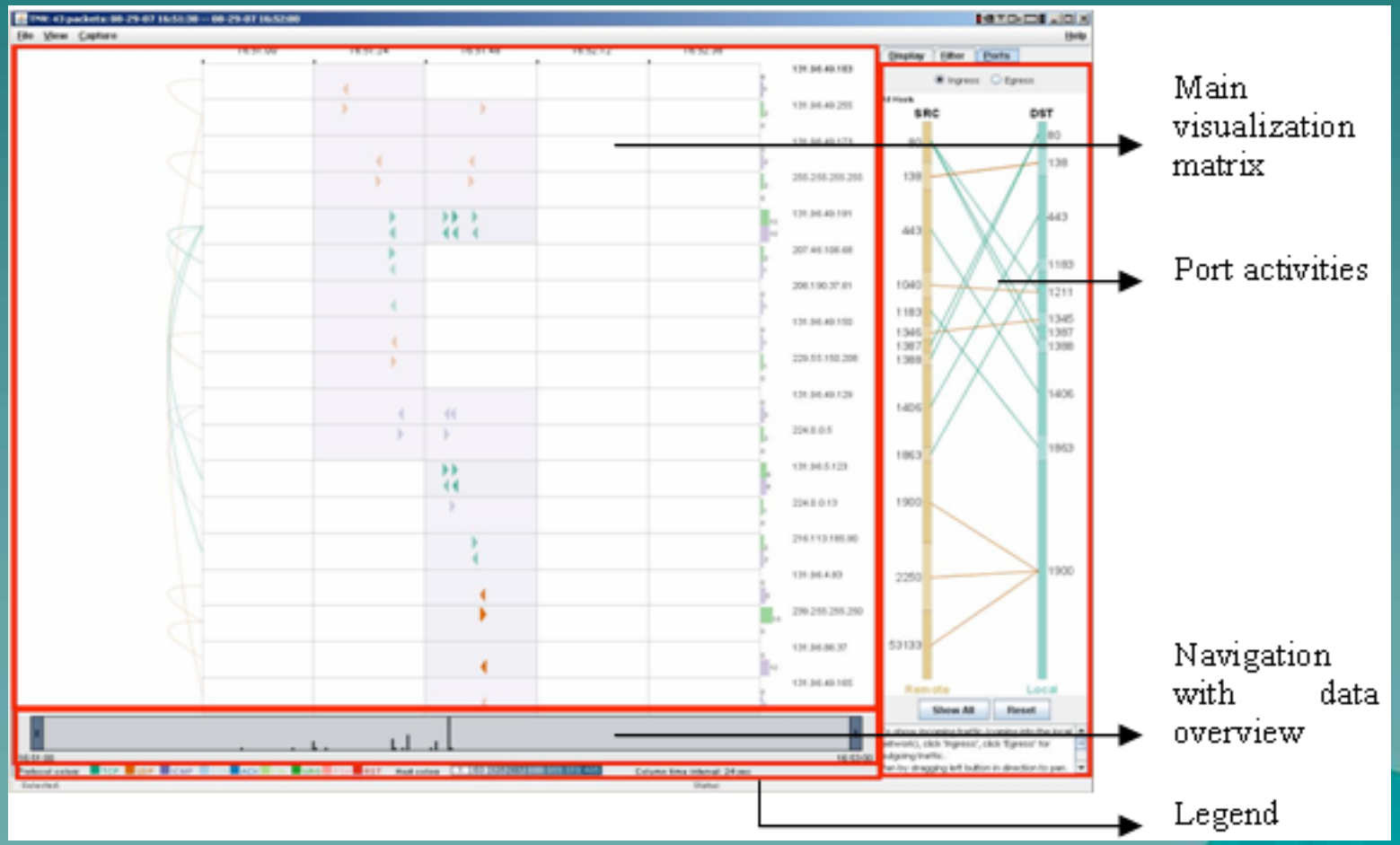
Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

Hierarchical analysis of data visualization

- Workspace
  - One or more visual frames that are designed for a specific purpose.
- Visual frame
  - A visual frame is a window within a workspace and contains multiple visual patterns.
- Visual patterns
  - A visual pattern is a set of visual units that are readily perceived as a group
- Visual units
  - Some examples of visual units such as: point, line, 2D shape (glyph), 3D object, text, and image.
- Visual attributes
  - Each visual unit is defined by seven visual attributes [10]: position, size, shape, value, color, orientation, and texture.

TNV Interface



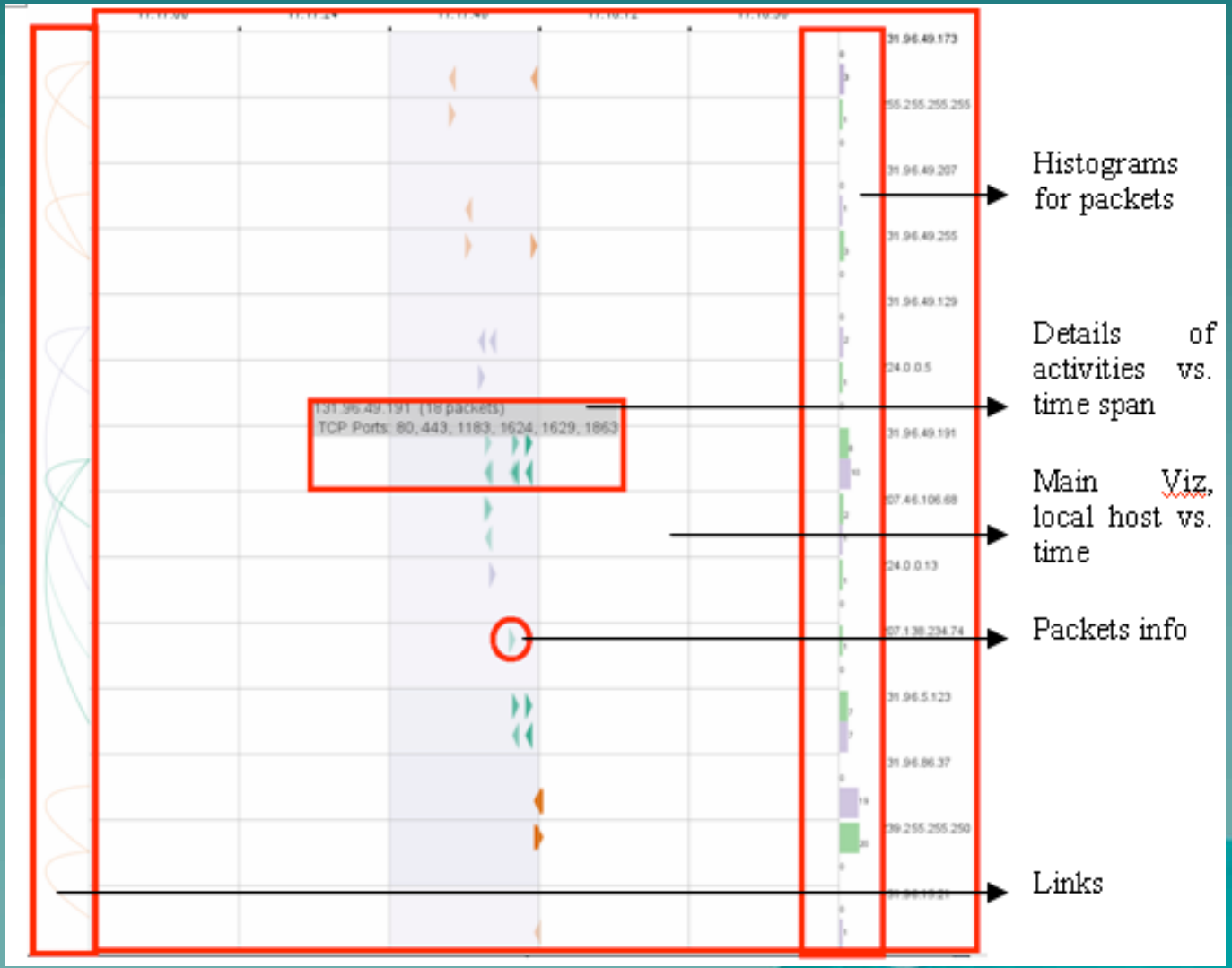
Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

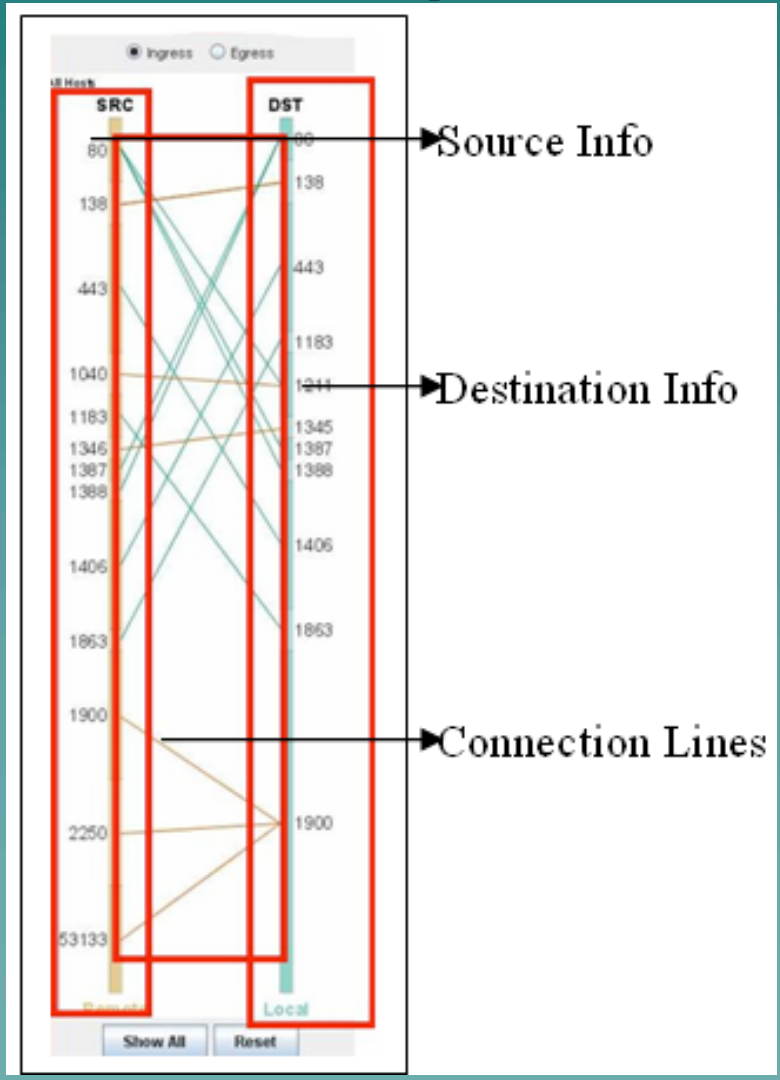
Criterion of separable dimensions

- X and Y coordinates (position)
- Shape, Size, Color,
- Value (gray scale),
- Orientation

Five different dimensions in TNV main visualization matrix.



Three different dimensions in port visualization



- (a) Source and destination information is categorized based on the coordinate (vertical axis).
- (b) Connection lines between the two axes are categorized based on shape/color

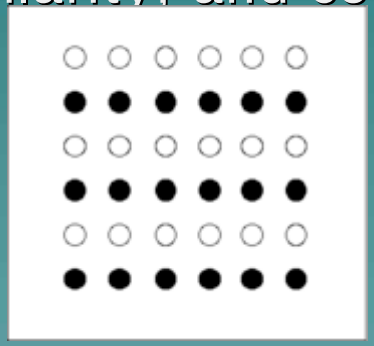
Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

Complexity of visual integration

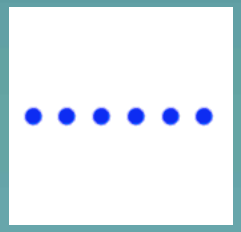
- For each visual frame, we identify the visual patterns in that frame based on four Gestalt laws: proximity, good continuation, similarity, and common fate.

Law of Proximity



Law of Similarity

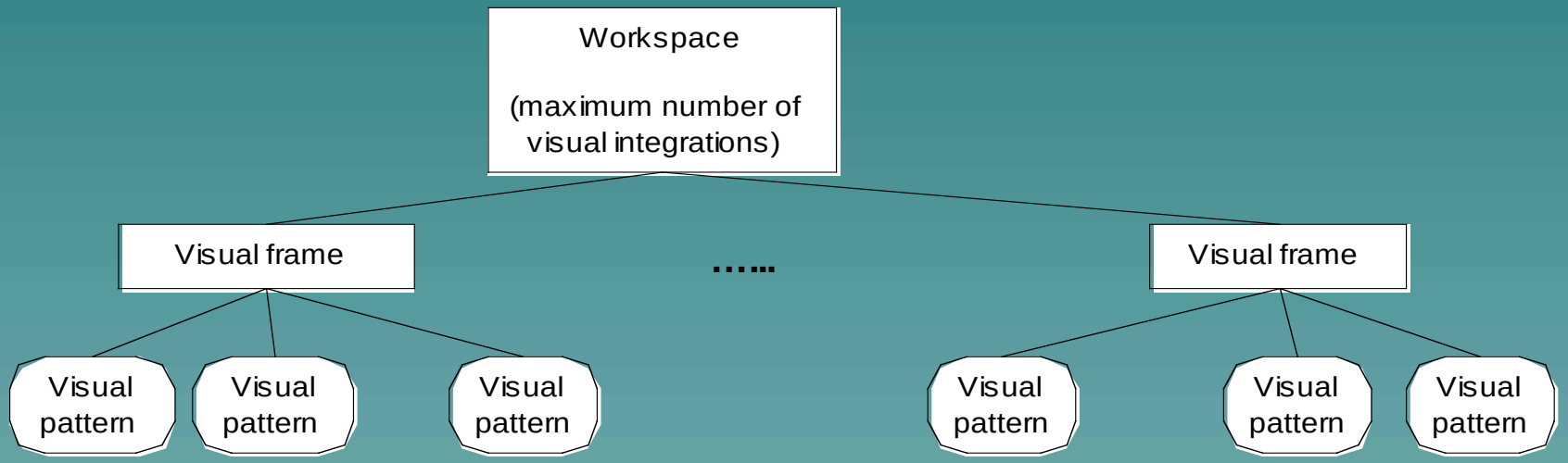
Law of Continuation



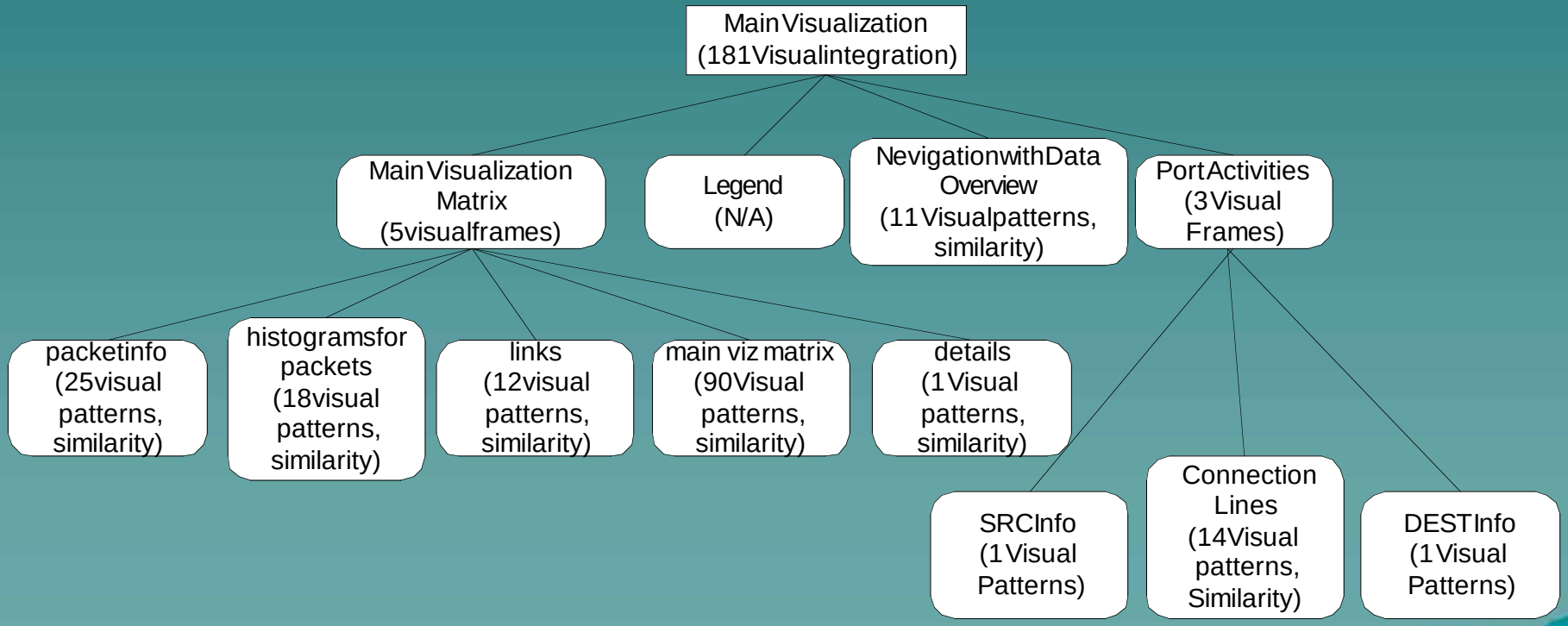
Law of Common Fate

- The number of nodes on the visual integration complexity represents the upper bound of visual integration a reader might perform.

Visual integration tree



Visual integration tree for TNV



Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

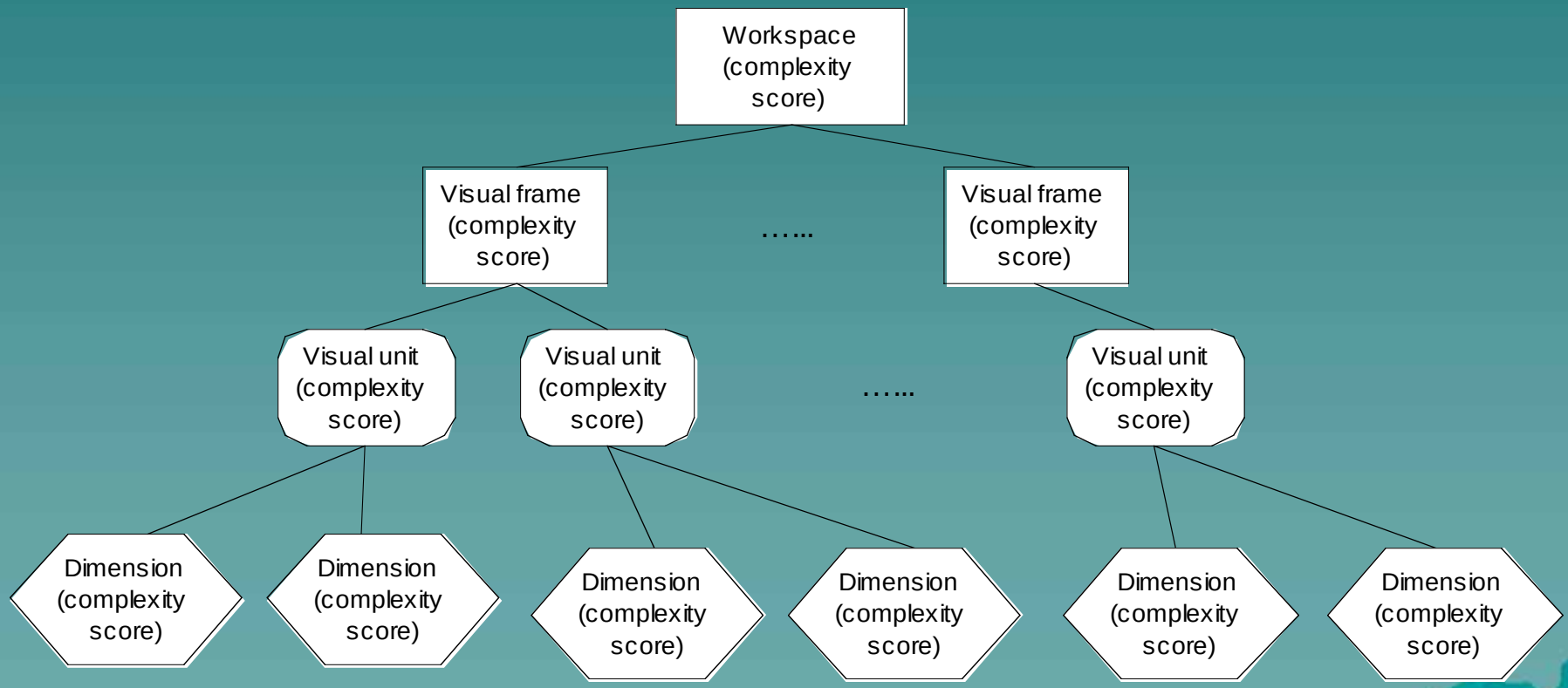
Complexity of visual mapping

- Readers not only need to remember the mapping between data parameters and visual attributes, they also need to interpret the value of these visual attributes.
- The mental effort for such interpretation is another source of extraneous cognitive load.
- For each separable dimension, we assign a score for the complexity of interpreting the values of the visual attribute based on the following criteria

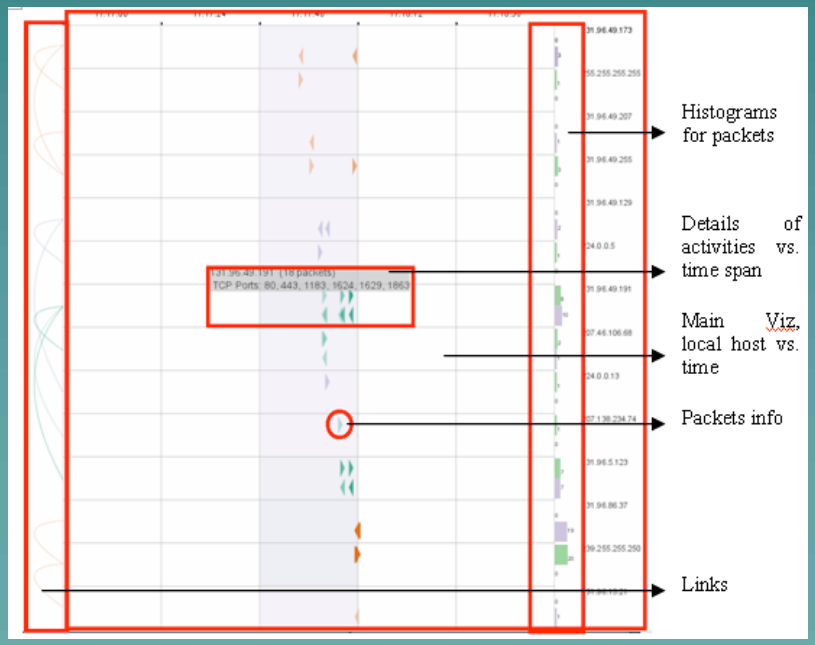
Complexity scores for interpreting visual units

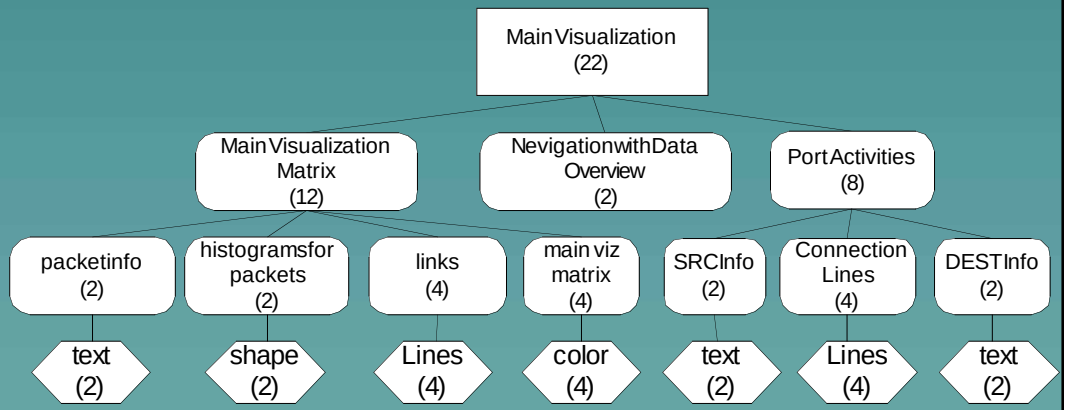
Complexity score	*Criteria*
5	Very difficult to interpret. There is no legend. A typical reader has to memorize the mapping between the value of the visual attribute and the value of the corresponding data parameter
4	More difficult to interpret. A typical reader needs to frequently refer to a legend to interpret the value of the visual attributes
3	Somewhat difficult to interpret. A typical reader needs to refer to a legend from time to time.
2	Relatively easy to interpret. A typical reader only needs to refer to a legend occasionally.
1	Easy to interpret. This is based on common knowledge. There is no need to memorize or refer to a legend.

Visual mapping complexity tree



Visual mapping complexity tree for TNV





Complexity analysis steps

- Analyze the design space
- Separable dimensions of visual units
- Analyze the visual integration complexity
- Analyze the visual mapping complexity
- Analyze the visual search complexity

Complexity of visual search

- Target-distracter difference is the key to efficient visual search
- Four major factors affect the target-distracter differences

Color

$$c = \left(\left|\frac{\sum_{i=1}^{n} T_r}{N} - D_r\right| + \left|\frac{\sum_{i=1}^{n} T_g}{N} - D_g\right| + \left|\frac{\sum_{i=1}^{n} T_b}{N} - D_b\right|\right) / 765$$

motion

$$M = \frac{\left|\frac{\sum_{i=1}^{n} Tf}{N} - Df\right|}{F}$$

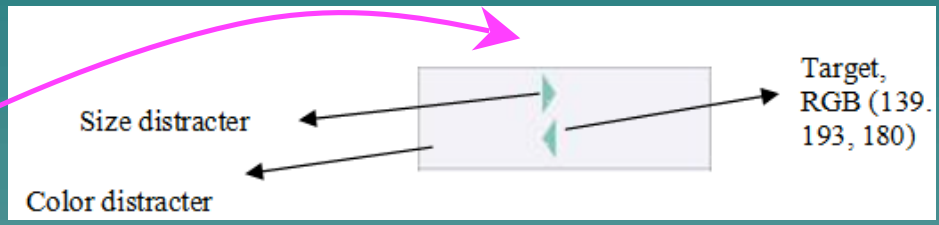
Size

$$s = \frac{\sum_{i=1}^{n} \frac{|Ts - Ds|}{s}}{N}$$

orientation

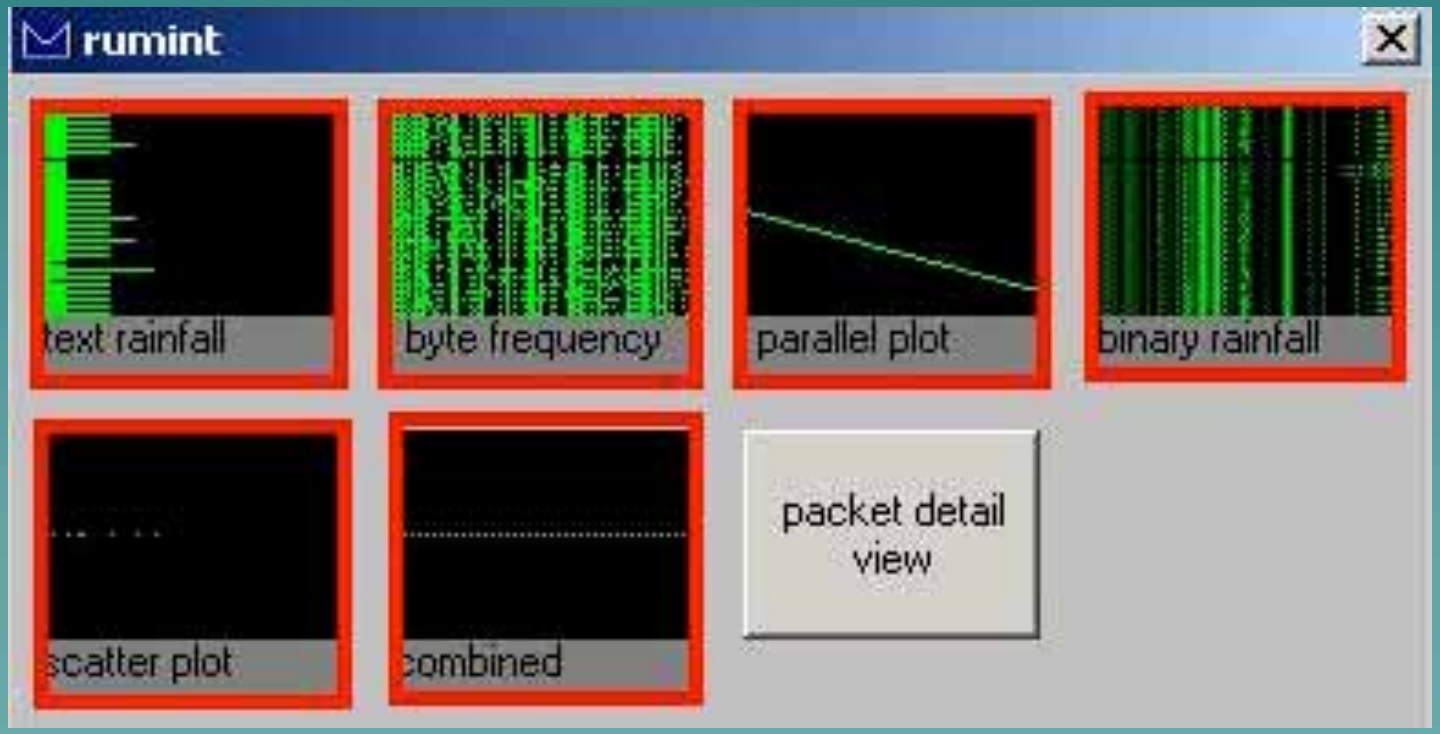
$$o = \frac{\sum_{i=1}^{n} \frac{|To - Do|}{N}}{180}$$

Target-distracter difference scores for TNV

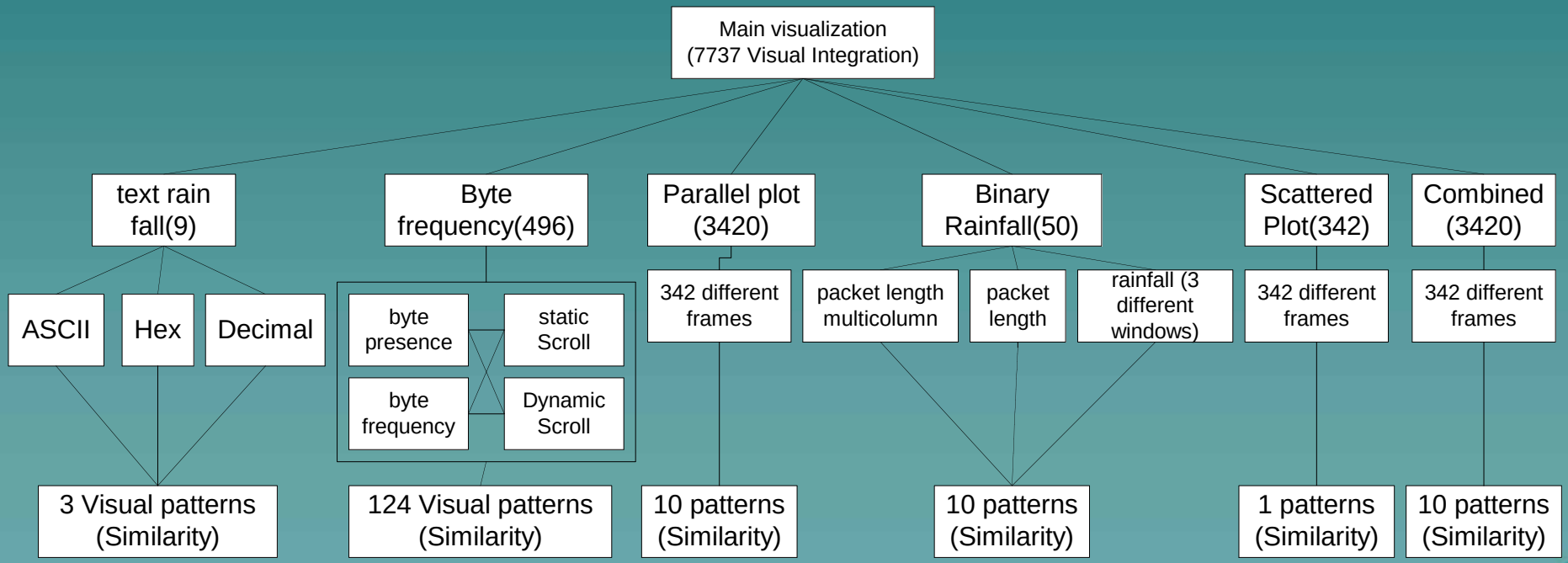


	Color	Motion	Size	Orientation
Target-distracter Difference scores	0.2850	N/A	0	1

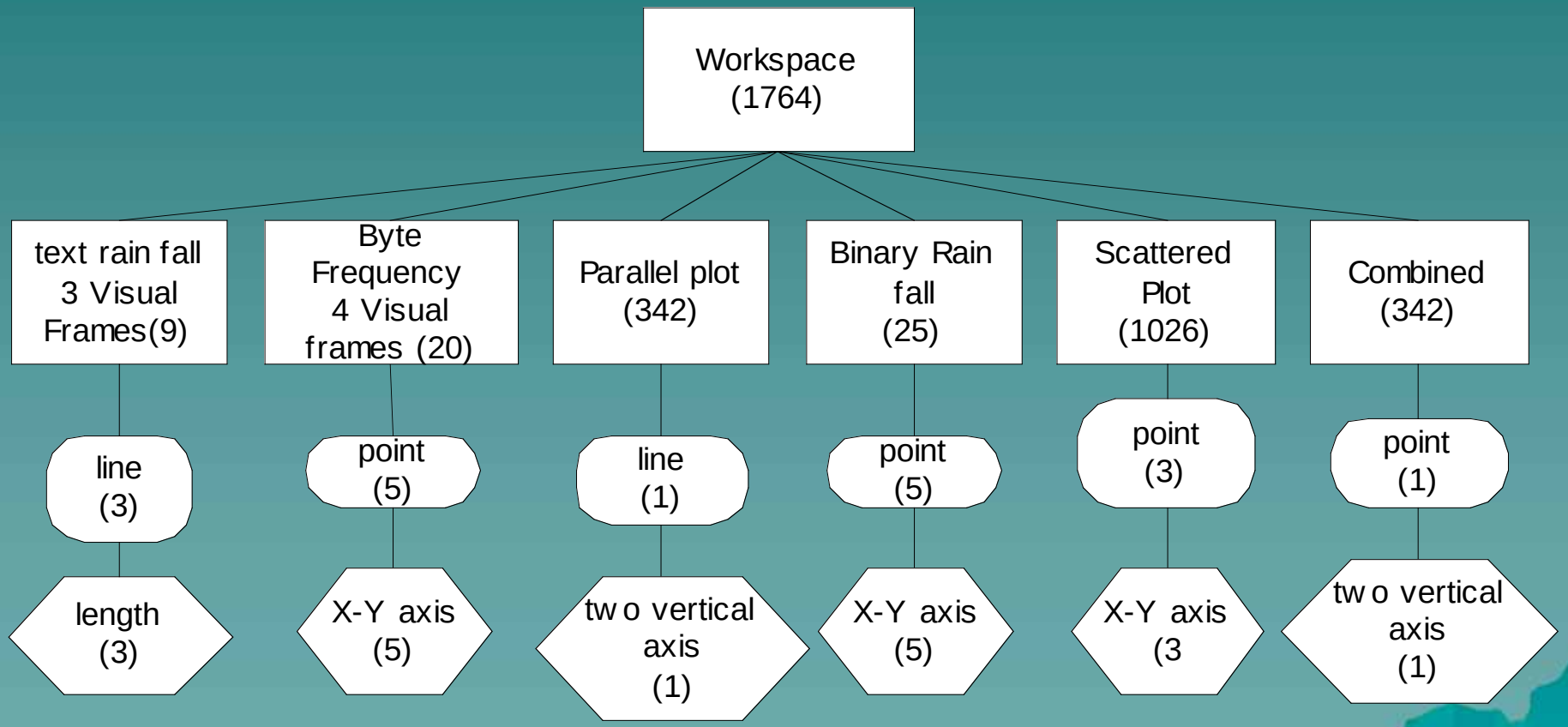
Rumint thumbnail overview



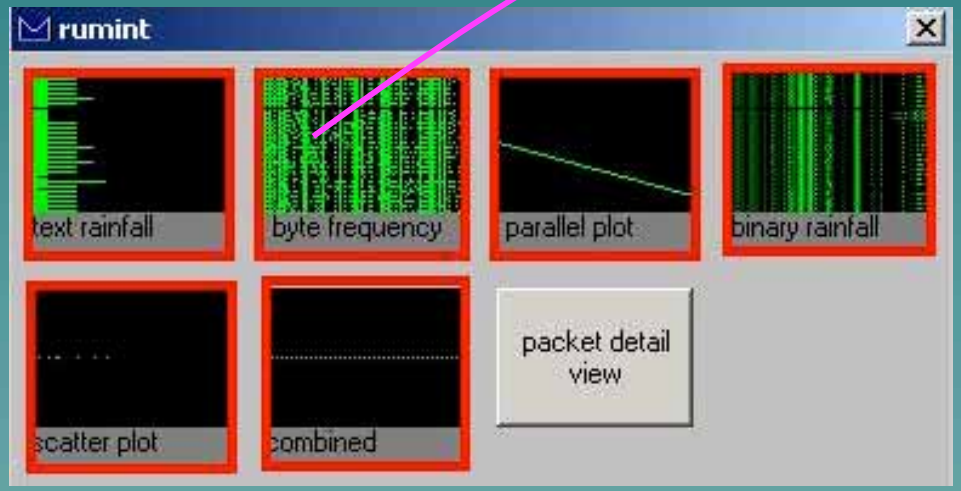
Visual integration tree for Rumint

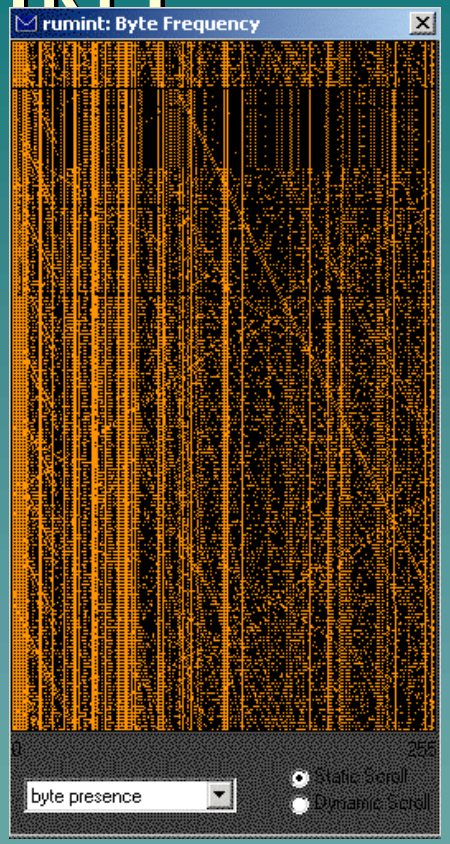


Visual mapping tree for Rumint



Visual search complexity metrics for RUMINT





	Color	Motion	Size (pix^2)	Orientation
Target-distracter difference scores	0.33333	n/a	0.32169	0

Comparison

- Rumint has a higher visual integration complexity score than TNV
- Rumint also has a higher visual mapping complexity score than TNV
- In general, Rumint is more complex than TNV, but it also presents more information than TNV

Conclusion and future work

- We have presented a method to analyze the complexity of information visualization design
- The complexity is measured in terms of
  - Analyze the design space
  - Analyze the visual integration complexity
  - Analyze the visual mapping complexity
  - Analyze the visual search complexity
- They indicate the amount of cognitive load involved in comprehending a particular visualization design

Conclusion and future work

- The complexity analysis is a quick way to review the visualization design before any user study can be conducted.
- We are currently applying and refining this complexity analysis method in the fields of computer security and bioinformatics.