Concept Demonstration
Metrics for Effective Information Visualization

Richard Brath, Visible Decisions Inc., 200 Front St. W. #2203, Toronto, Canada; rbrath@vdi.com

ABSTRACT

Metrics for information visualization will help designers create and evaluate 3D information visualizations. Based on experience from 60+ 3D information visualizations, the metrics we propose are:
- number of data points and data density;
- number of dimensions and cognitive overhead;
- occlusion percentage; and
- reference context and percentage of identifiable points.

KEYWORDS

visualization, metric, occlusion, context, density, dimensionality.

INTRODUCTION

Metrics will aid the development of effective 3D information visualizations; just as Edward Tufte's metrics [Tuf83] have improved the effectiveness of 2D information graphics.

Expert knowledge accumulated during the development of over 60 3D information visualizations has resulted in guidelines [Bra97] that are the basis for the set of metrics we propose.

PRIOR WORK

Tufte's metrics [Tuf83] are specific to 2D representations, although Tufte does explore 3D in his recent works [Tuf90, Tuf96]. The intuitive extension of these metrics into 3D has resulted in some exceptions [Bra97]. For example, Tufte's principle of erasing redundant data-ink extended into 3D becomes the erasing of redundant geometry – however, redundant geometry within a 3D information visualization may aid interpretation if the scene is partially occluded or may aid the viewer's recall of the data mapping to the visual attributes.

Bertin [Ber83] and Harris[Har96] classify and characterize some 3D information graphic types. Information visualization research has characterized relations between data dimensions and visual representations [Rot94, Zho96] as well as interactions [Chu96].

Cleveland's methods of visual analysis relate the statistical model to 2D and 3D visualizations incorporating many guidelines into the application Trellis Display [Cle93, Bec94].

Information visualizations rely on combinations of visual attributes to convey multidimensional information. The classification of combinations of visual attributes which can be processed by a viewer in almost instantly (i.e. preattentively) vs. those which require a slower visual search can result in quantified guidelines, for example [Hea96].

METHODOLOGY AND ANALYSIS

The 60+ proprietary visualizations, from which our metrics are derived, were developed through an iterative process with frequent client reviews as well as occasional peer reviews within the company. Some of the visualizations were published as case studies e.g. [Wri95, Wri97]. As the number of applications grows, some techniques have recurred and been identified as successful, and some visual difficulties recur.

To quantify these differences, similar visualizations were compared, differences discussed, models proposed and then measured. The metric was applied to other internal visualizations where possible to test generalization.

These metrics assume:
- A static image of the scene. Although 3D visualizations can be viewed from many angles, users create printouts and images attached to emails, thus losing the benefits of 3D scene navigation. Further, novices tend to use 3D navigation sparingly until they become comfortable with the navigation interface and the representation. We therefore advocate using the "best view on the scene" for calculating metrics. This permits published images of other visualizations to be used in testing and generalizing metrics.
- No interaction. Interaction, such as animation, filtering, drill-down, etc., can enhance the usability of the information visualization, but cannot be used with still images for the same reasons as described above. Attempting to include interaction factors in the metrics adds significant complexity to them.

RESULTING METRICS

Number of Data Points and Data Density

Tufte's metric of data density is applicable to 3D information visualization. In general, Tufte assumes that the greater amount of data represented per square centimeter of print, the more effective the resulting representation. From our experience, this has held true in 3D visualization as well.

Since most of our visualizations are displayed on high resolution displays of similar resolutions (e.g. 1280 x 1024) we often ignore the denominator and focus only on the number of data points.

Number of Data Points: number of discrete data values represented on screen at an instant.

For example, a typical 2D time series line chart with 20 intervals along the time axis has 40 data points (20 2-coordinate pairs).

Data Density: number of data points / number of pixels in the display where number of pixels does not include the pixels in the window borders, menus, etc.

Bounds

Our visualizations range from 200 to 100,000 data points. The table (at the end) shows values from some of our visualizations.

Lower:
Visualizations with less than 500 data points are questionable visualizations - the same information can be represented in other simpler forms. For example, a time series with 250 time intervals and 250 corresponding measurements can be represented in a 2D line chart. Or, a 12 x 12 array can be represented as a 2D array of bars or "heat map" (a 2D grid of squares with each square's hue mapped to a data value).

This is not a fixed lower bound. The lower bound threshold depends on:
- The use of the application. Applications required to be quickly comprehended, such as, an executive information system, real-time alarm system, or marketing application tend to have lower numbers of points compared to analysis applications. The lower bounds for these applications can be lower than 500.
- Interaction with application. Although a snapshot may have few data points, techniques, such as animation, effectively increase the number of data points delivered to the user. For example, a visualization we produced for bank executives has 410 data points at any instant, from an interactive animation of 20,000+ data points. The user's short-term memory enables the user to comprehend more than the 410 data points.

Upper:
Visualization can deliver overwhelming amounts of information. We have not been able to establish a simple upper bound. We have seen applications with 50,000 to 100,000 data points (e.g. 100 time series each with 500 time intervals) perceived as effective while other applications with 2,000 data points (e.g. an application with 500 objects with 5 values each) perceived as too complex. Thus we have derived other metrics for complexity:

Complexity

A complex visualization is more difficult to comprehend than a simple one. An effective visualization should seek to reduce complexity.

Visualizations frequently represent multivariate data. The greater the number of dimensions which are displayed in the visualization, the greater the cognitive complexity for the user. For example, a simple 2D chart representing many different variables may use many different colors and/or patterns resulting in more cognitive effort on the part of the user to remember the mapping between the data dimensions and the representation.

Number of simultaneous dimensions

A simple measure of cognitive complexity is the number of different data dimensions displayed simultaneously.

Our visualizations have ranged from 3 to 150 data dimensions simultaneously displayed in the visualization. Our experience has been that higher dimensional problems are difficult to design suitable representations for; and thus the lowest number of dimensions which solves the task is desirable.

Example: An early visualization representing 50 different simultaneous data dimensions (e.g. sales, margin, area, etc.) was not effective since the dimensions had a weak correlation to each other and thus the different dimensions were not comparable to each other. The result was a complex visualization. The visualization was later redesigned to place all data elements with the same data dimension in the same area, effectively making 50 independent visualizations (i.e. small multiples).

Maximum of the number of dimensions from each separable task representation

Many of our visualizations have a higher number of dimensions than we intuitively thought when we reviewed this metric against past visualizations. This resulted from a heavy reliance on compound visualizations. A compound visualization results from two (or more) spatially distinct different data representations, each of which can operate independently, but can be used together to correlate information in one representation with that in another. Tufte's example of dot-dash-plot [Tuf83, p.133], which combines the bivariate distribution (scatterplot) in the center with marginal distributions along the axes, is an example of a compound visualization, as is Tufte's example of a Java railroad timetable [Tuf90, p.24].

This measure calculates the number of dimensions for each separable representation based on the task (i.e. the typical task, or the most complex task if the visualization solves many tasks).

For example, a six dimensional data set can be represented with 3 separate scatterplots. If only the two dimensions within a scatterplot are compared at any one time, the resulting maximum score is 2. A task requiring comprehension of all six dimensions simultaneously requires correlating between the 3 scatterplots simultaneously, which should thus have a score of 6. Thus, the "Maximum of the number of dimensions from each separable task representation" is dependent on the definition of task.

Appropriate representations + dimensional score

Neither of the above "dimensional counts" consider the effectiveness of the representation. An effective mapping of data dimensions into a visual representation requires little explanation. A poor mapping results in a visualization that requires repeated explanation.

We have made visualizations worse as a result of a redesign because we did not consider appropriate representations.

We have created a simple model of cognitive complexity with a scoring scheme. We evaluate the effectiveness of the mapping of each data dimension into a visual dimension (e.g. length, width, hue, position, orientation, shape, etc.). This simplistic generalization does not take into account many factors but:
- is quick to evaluate,
- does not require classification of data dimensions and visual dimensions, and
- does not require a large correlation matrix of scores.

Dimensional Score:

The scoring is based on measuring the effectiveness of the mapping from the data to the visual dimension. A mapping fits into 4 categories (from worst to best):
- n-to-1 mapping,
- 1-to-1 general mapping,
- 1-to-1 intuitive mapping, and
- preexisting understood representation.

Each mapping has a associated score, based on a simple model of the cognitive effort for that mapping. For example, the 1-to-1 general mapping has a cognitive score of 2: 1 point for the user to recall the data dimension and 1 point for the user to recall the visual dimension that it maps to. The sum of all the scores for all the dimensions within a visualization is the dimensional score. The desired result is the lowest possible score.

Many-to-one mapping:
Some of our visualizations overloaded the color dimension in the visualization. Different colors would represent positive and negative values, different data types, etc. Reliance on only one visual dimension to convey multiple data dimensions has a severe negative impact on the effectiveness of the visualization - in some cases requiring us to redesign the mapping.
We assign the n-to-1 mapping a score of 3 x n, assuming a simple cognitive model of:
- 1 point to recall the data dimension,
- 1 point to recall the visual dimension, and
- 1 point to recall the mapping rule for that dimension

multiplied by the number of data dimensions, since this recall is required for each data dimension.

One-to-one general mapping:
This is a common mapping found in visualizations. For example, price maps to height, or priority maps to color, etc. The score is 2: 1 for data dimension recall, 1 for visual dimension recall.

One-to-one intuitive mapping:
Some data types map more appropriately to some visual dimensions than others. We assume that data dimensions mapped to visual dimensions with similar connotations are easier to recall. For example, a data value "size" maps to the visual dimension "size" well. Or spatial data dimensions map well to spatial visual dimensions. We assign this mapping a score of 1: the visualization dimension is implied by the data dimension and does not need to be recalled explicitly.

Preexisting understood representation:
A widely understood visual representation with its frame of reference is a representation that is automatically understood within the task domain. For example, latitude and longitude overlaid on a map is automatically understood. Depending on the task and the user community, we also find other representations are automatically understood:
- rooms and other direct spatial representations (e.g. many molecular, medical and fluid dynamic visualizations),
- bar charts (e.g. histogram),
- time-series charts (e.g. stock price over time),
- pie charts (e.g. market share),
- hierarchies (e.g. org charts) and
- some scatterplots or other industry specific graphics (such as a yield-curve in the financial markets).

We assign these a score of 0 - e.g. there is no cognitive complexity to remember that a map is map.

Scoring Summary:

Mapping	Score
n-to-1	3 x n
1-to-1 general	2
1-to-1 intuitive	1
preexisting representation	0

Stock Visualization Example: A stock visualization was redesigned to reduce redundant geometry. We showed both visualizations to various people with a reasonably consistent reaction. Most people comprehend the information in stock visualization A (fig.1) and are able to pick out meaningful patterns without prompting, while many people viewing visualization B (fig.2) struggle to maintain the mapping:

Data	Visualization A (2 cubes + line)	Score A	Vis. B (1 cube)	Score B
stock price	height of line	1	height	1
buy price	vertical position of red cube	1		
sell price	vertical position of blue cube	1		
buy size	size of red cube	1		
sell size	size of blue cube	1		
total volume			width	2
block volume			depth	2
Spread	distance bet. cubes	1	hue	2
Liquidity	size of both cubes	1	brightness	2
Stock	location	2	location	2
Total score		**9**		**11**

Even though the two visualizations did not contain the same data, the result was counter to our original intuitions. Visualization B contained less data and less geometry yet was more difficult to understand. Even if one considers the common components between the two visualizations, the score for B is worse (7 vs. 5). We believe the result in B was more difficult for the user to understand because simple cognitive mappings had been replaced with constraint to minimize the extraneous geometry resulting in a complex cognitive mapping.

This measure does not adequately capture complexity across visualizations – the ordering that results from this score does not correspond with subjective interpretations of complexity for the same visualizations: thus it cannot be generalized (i.e. it is not a metric [Fen96]). However, we find it useful for comparing tradeoffs between alternative designs for the same visualization.

Occlusion Percentage

Complete occlusion of data objects hides information from the user and is undesirable.

Occlusion percentage: number of data points completely obscured / number of data points

In most of our visualizations, we try to create optimal viewpoints with minimal complete occlusion. We aim for 0% complete occlusion, and typically have values ranging from 0 to 10% occluded (see table at end). The actual threshold depends on the application - a system control visualization requires close to 0%

occlusion, while a volume visualization will always have a back side occluded - 50% or more may be occluded.

Figure 3 shows a bar chart using cubes where length, width, height and color of each bar are data dependent (with the bars sorted by height). The result is a visualization with little complete occlusion and a lot of partial occlusion. The use of cubes results in redundant geometry which from many viewpoints provides enough partial information to establish the data.

This metric does not address partial occlusion. Figure 4 shows a relationship diagram. The endpoints are visible in all cases, but as the obscuring object becomes larger, the certainty of the relationships degrades. A weighted measure based on percentage of partial occlusion would capture the effect in fig. 4 but would also capture the partial occlusion of redundant geometry in fig. 3.

Reference Context and Percentage of Identifiable Points

If the location of a graphic object cannot be identified, then the user may not be able to comprehend some of the values represented by that object. A visualization should provide a reference context which can be comprehended from most viewpoints. Techniques from other domains such as CAD are applicable, for example, shadow walls and anchor lines.

Percentage of identifiable points: number of visible data points identifiable in relation to every other visible data point / (number of visible data points^2)

Scatterplot Example: For the typical 3D scatterplot (fig. 5), it is difficult to establish the depth of different data points, even if multiple simultaneous viewpoints are available. Typically, this is overcome by interaction or animation - but this can be difficult for novice users (and not available in print). We typically draw a line from the data point to a reference plane (fig. 6). Before the line is added, the percent identifiable is close to 0; after the line, the percent is close to 100. (See also fig. 7 and fig. 8).

Blob Example: Blobby visualizations frequently have multiple blobs within the scene. All the points on a given blob are identifiable in relation to each other, but the blobs are not identifiable in relation to each other. For example, a visualization with 2 similar sized blobs will have a percentage of identifiable points near 50% since the points on one blob can not be adequately referred to the points on the other blob.

CONCLUSIONS AND FUTURE WORK

The creation of information visualization metrics will be a valuable aid to designers and users of visualizations. As a result of our experience, we have been able to discard assumptions (e.g. less geometry is better) and communicate our findings.

These metrics are immature and need to be tested against numerous visualizations to establish correctness and applicability – for example, the dimensional score does not generalize in its current form. It is our intention to test these metrics against a broader set of visualizations. The metrics should also be extended to accommodate interactions and navigation.

Table of Measures as Applied to Sample Visualizations

Name of Visualization	Number Data Points	Number Dimen-sions	Max. Num Dims	Mapping Score	Occlusion %	Identifi-able Points %
Risk Movies	410	9	5	11	0	100
Mold Series	3000	4	4	7	0	100
Mold 3space	3000	4	4	8	25	90
Pipeline	600	4	4	4	0	100
Ambiguity	700	5	5	12	20	0
Option Series	100000	3	3	6	25	90
Head Trader	2000	15	5	19	0	100

(see also http://employees.vdi.com/~rbrath/iv97/index.htm)

REFERENCES

[Ber83]   J. Bertin. *Semiology of Graphics*, University of Wisconsin Press, 1983.

[Bra97]   R. Brath. Interactive Information Visualization Guidelines. To be published in *Proceedings of HCI International '97* in August 1997

[Chu96]   M Chuah and S Roth. On the Semantics of Interactive Visualizations. In *Proceedings of Information Visualization '96*: 29-36. IEEE, 1996.

[Cle93]   W. S. Cleveland. *Visualizing Data*, Hobart Press, 1993.

[Bec94]   R. A. Becker, W. S. Cleveland and M. J. Shyu, *Trellis Graphics User's Guide*, 1994.

[Fen96]   N. E. Fenton, S. L. Pfleeger, *Software Metrics: A Rigorous and Practical Approach*, International Thomson Computer Press, 1996.

[Har96]   R. L. Harris. *Information Graphics*. Management Graphics, 1996.

[Hea96]   C. G. Healey, K. S. Booth, J. T. Enns: High-Speed Visual Estimation Using Preattentive Processing. *TOCHI* 3(2): 107-135. 1996.

[Rot94]   S F Roth, J. Kolojejchick, J. Mattis, J. Goldstein. Interactive Graphic Design Using Automatic Presentation Knowledge. In *Proceedings of CHI'94*: 112-117, 1994.

[Tuf83]   E. R. Tufte. *The Visual Display of Quantitative Information*, Graphics Press, 1983.

[Tuf90]   E. R. Tufte. *Envisioning Information*, Graphics Press, 1990.

[Tuf96]   E. R. Tufte. *Visual Explanations*, Graphics Press, 1996.

[Wri95]   W. Wright. Information Animation: Applications in the Capital Markets. In *Proceedings of Information Visualization'95*. IEEE, 1995.

[Wri97]   W. Wright, Information Visualization: The Fourth Dimension. In *Database Programming & Design*, Vol. 10, No. 4. Miller Freeman, April 1997.

[Zho96]   M Zhou and S Feiner. Data Characterization For Automatically Visualizing Heterogeneous Information. In *Proceedings of Information Visualization '96*: 13-20. IEEE, 1996.