Using Text in Visualizations for Micro/Macro Readings

Richard Brath
London South Bank University
brathr@lsbu.ac.uk

Ebad Banissi
London South Bank University
banisse@lsbu.ac.uk

ABSTRACT
Literal text and font attributes can be used in visualizations to create micro/macro encodings in text visualizations. *Alphanumeric marks* can replace geometric marks in standard plots revealing micro-level data and increasing data density. *In-Context Representations* add macro-level information into traditional text lists and flowing text using font attributes to make high-level patterns perceivable. This is a work in progress with novel design contributions regarding generalized use of text and font attributes in visualization.

Author Keywords
Text visualization; font-attributes; proportional length encoding; alphanumeric glyphs.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces: Screen design (e.g., text, graphics, color).

INTRODUCTION
Many approaches to text visualization analyze text content and then present visualizations of summarized text analytics as traditional visual representations (e.g. bar charts and scatterplots); textual representations (e.g. tag clouds or treemaps e.g. newsmap); or text markup (e.g. highlighting text with a background color). Expanding our approaches to represent text in data dense displays creates new possibilities for the visualization front-end to text analytics.

In particular, this paper considers the use of text in visualizations to create micro/macro readings (popularized by Edward Tufte [1]). Encoding opportunities include:

i. **Literal text**. Instead of abstract geometry, alphanumeric characters directly encode data. For macro-readings, characters can be perceived as similar [2] without decoding specific glyphs. For micro-readings, while characters are perceived pre-attentively, people are highly trained to quickly recognize and distinguish characters. These can be single characters (e.g. a leaf in a stem and leaf plot), short mnemonic codes (e.g. country code) or words and phrases.

ii. **Visual Attributes**. Text can further encode additional data via traditional visual attributes that have pre-attentive characteristics, e.g. size and color, or unique text specific *font attributes*, e.g. bold, italic, underline and font family.

These encodings can then be applied to add a micro-level or a macro-level to representations:

a. **Alphanumeric Marks**: Instead of simple geometric marks, (e.g. dots in a scatterplot or bars in a bar chart), text is used instead. The overall macro-level reading of patterns in the plot is maintained, while additional information is encoded at the micro-level of text facilitating direct visual reading of micro-information rather than slower interactive techniques (e.g. pointing the mouse and reading a tooltip).

b. **In-Context Representations**: Flowing textual representations (e.g. text lists or prose) can have additional information encoded via visual attributes. In these representations, the micro-reading of the literal text is maintained while the visualization enhancement provides additional macro-level information for at-a-glance perception.

In both uses, the user can read the visualization at multiple levels without a requirement for interaction; and the lossiness is reduced [3] as information density is increased by layering additional information into a visualization. This paper is a work in progress and provides a contribution regarding the use of text and font attributes to enable multiple levels of reading through a series of illustrated examples.

BACKGROUND
Micro/macro displays show large amounts of high-density information encoded to facilitate the visual reading at various levels ranging from high-level macro patterns such as trends, clusters and outliers; down to low level data points, such as individual observations and local peers.

Tufte's approach is somewhat similar to Shneiderman's *visual information-seeking mantra*: overview first, zoom and filter, then details on demand [4]. In Shneiderman's approach, interactivity is used to reveal the micro-data information whereas Tufte plots it directly. For Tufte, paper is exceedingly high resolution whereas in Shneiderman's case, computers in 1996 are fairly low resolution (72 DPI) with graphics hardware not necessarily suited to managing and manipulating thousands of elements. Modern displays have higher pixel density (e.g. 150-300 DPI) and enhanced graphics capabilities which enables information dense micro/macro displays. Furthermore, non-personal use cases (e.g. video walls, presentations) may have scenarios where alternatives to direct interaction may be desirable.

Typography and “Color”

Leveraging typography in visualization has been discussed in terms of visual attributes [5]. Using type in visualizations utilizes the notion that a well-designed font has an even distribution of text ink across a sequence of letters regardless if the letters are sparse (e.g. i, v) or dense (e.g. m, e). In type design this even distribution of ink is referred to as *color* (not to be confused with hue) e.g. [6,7]. For example, a blurred paragraph from a scan of a typeset book shows the relatively even distribution of intensity in fig. 1.

Figure 1. Blurred text from a typeset book. Words tend to be equally grey, delineated by white space.

The font designer creates even color by adjusting the letterform shape (e.g. the vertical stroke of a c may be thicker than the corresponding stroke of an e) and spacing between letters and specific letter pairs (e.g. kerning, ligatures). As a result, with an even weight across a field of type using the same font, the viewer does not perceive any particular letters standing out. Conversely, concrete poetry adjusts the layout and weights of text to create macro-level shapes, e.g. *Calligrammes* by poet Guillaume Apollinaire [8] or *The Mouse’s Tale* by Lewis Carroll [9].

Micro/macro readings using text in visualizations utilizes the even density of type so that blocks of text can be viewed at a macro-level as if they were marks or bars because they all have the same density (associative [2]); while at a micro-level the individual letters and words can be read. Furthermore, additional micro/macro information can be embedded by adjusting font attributes such as bold, italic, etc.

Historic Alphanumeric Visualization Examples

In finance, starting in the late 1800’s, charts evolved plotting individual prices in a matrix organized by time (horizontally) and price (vertically) [10,11,12]. These turned into *point and figure charts* (fig. 2 left), using X’s to indicate rising prices, O’s to indicate falling prices, other characters to indicate events (e.g. start of a month), and coloring of characters (e.g. to differentiate columns). *Market profile charts* (fig. 2 right) use characters to indicate the time of day that a commodity trades at a specific price level. A common encoding uses A-X, a-x to indicate half hour intervals starting at midnight with uppercase for trades in the morning, lowercase in the afternoon. Characters are aligned vertically by price and stacked horizontally forming a histogram, enabling a macro-reading (the distribution) and a micro-reading (the individual characters). Variations may use color, background shapes and special characters. .

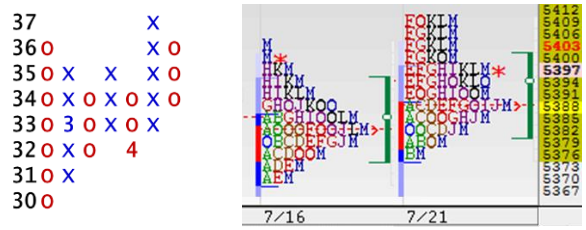

Figure 2. Left: Point and figure chart. Right: Market profile chart (see text for description).

Tukey’s stem and leaf plots [13] similarly stack characters to form distributions: the numbers along the vertical axis correspond to bins of the distribution and characters correspond to next significant digit (fig. 10 left).

Bertin’s *Semiology of Graphics* [2] has very few examples using alphanumeric plots other than labels (e.g. axis labels, cartogram labels, scatterplot labels) although there are a few examples of scatterplots with labels only (no other marks); or numeric tables with added graphical marks.

Tufte [1,14,15] provides seven different examples of alphanumeric charts that include variation of font attributes. E.g. Japanese railway schedules [1] are similar to stem and leaf plots, varying font size, background shape and provide additional tiny glyphs above text. Tufte also calls out a table-graphic with data points set as text in a grid (in a fine italic serif font) with lines of fit (e.g. regression, locus) overlaid and with additional labels (in bold sans serif font) [14].

Within information visualization, typography tends to be relegated to simple labels or representations such as a tag clouds or labeled treemaps. Tag clouds and treemaps vary font size but otherwise do not present different classes of information at different levels of micro/macro reading. In the 150+ infovis examples on *scimaps.org*, there are numerous examples of visualizations with text. In many, text is used as undifferentiated labels, or text size and font are varied to differentiate between title, axis label, legend, notes, etc. Some examples do mix text and other elements to create multiple levels of reading such as Paley’s *Map of Science* [16]. There are also examples that borrow visual cues from cartography, using size, uppercase, italics and tracking to indicate different levels of information [17, 18].

In document visualization, there are techniques to present both the text of interest and surrounding context such as fisheye representations (e.g. [19,20]), text highlighting (e.g. keywords in search such as Adobe Reader) and text scaling to increase visibility of keywords (e.g. [21,22]). Also note that there were early paper-based techniques utilizing font size, color, weight, capitalization, etc [23], now common in most software code editors.

ALPHANUMERIC MARKS

Micro-level data can be added to traditional bar chats and scatterplots by replacing dots and bars with alphanumerics.

Alphanumeric codes instead of Dots

Figure 3 shows a typical bubble plot, in this example, country birth rate (x) vs. death rate (y) and population (dot radius). Macro patterns, e.g. the crescent shape, are visible.

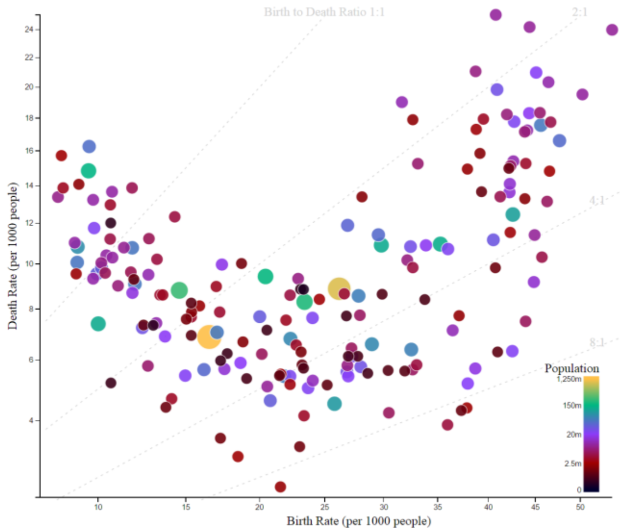

Figure 3. Birth rate vs. death rate by country with size and color proportional to population.

Text supports a high number of unique categoric glyphs. A Latin character set provides 62 unambiguous, uniquely identifiable, sortable/orderable glyphs (A-Z, a-z, 0-9). More unique symbols are available, although they may not be orderable or consistently named (e.g. # may be called hash, number sign, pounds or octothorpe). Small groups of characters may be used to create a large set of unique mnemonic identifiers, such as stock symbols or country codes.

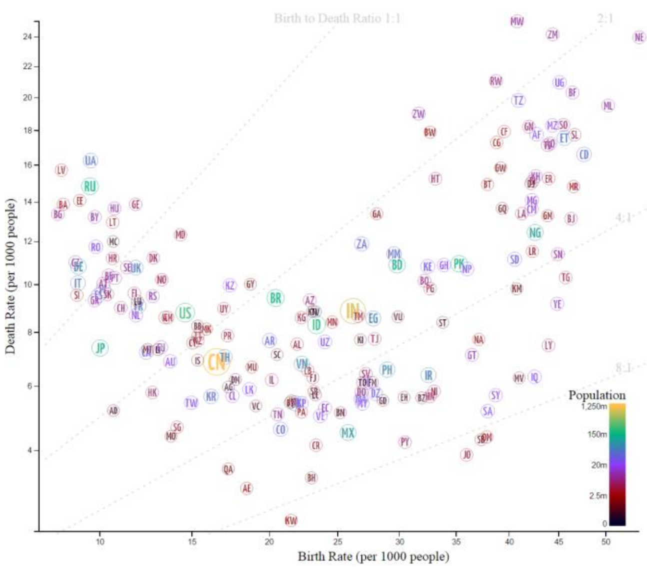

Figure 4. Same data as fig. 3, using country codes.

Fig. 4 shows the same birth rate vs. death rate dataset as fig. 3, in this example using 2-letter country ISO codes as a label-based scatterplot. The primary mark per data point is the 2-letter glyph. A secondary outline circle is also provided to make the marker location unambiguous (i.e. a text label without an added mark requires an additional note to indicate whether the data point is centered on the text, at the bottom of the text, left aligned, etc).

The alphanumeric representation does not require slow interactions to reveal the country names, which in turn may facilitate the ability to discern micro-level relationships in local areas. E.g. in this dataset, toward the upper left, are countries associated with the former east bloc - e.g. LV (Latvia), UA (Ukraine), RU (Russia), BG (Bulgaria), etc. It appears the former east bloc has a shrinking population.

Text vs. Glyphs

Instead of text, micro-level information could be encoded with icons or glyphs. However, ready-made charting software has a limited number of icons, e.g. Excel provides only nine different markers (e.g. square, circle) and D3,js provides six built-in symbols. A renewed interest in glyph design explores multiple attributes in glyphs (e.g. [24,25]).

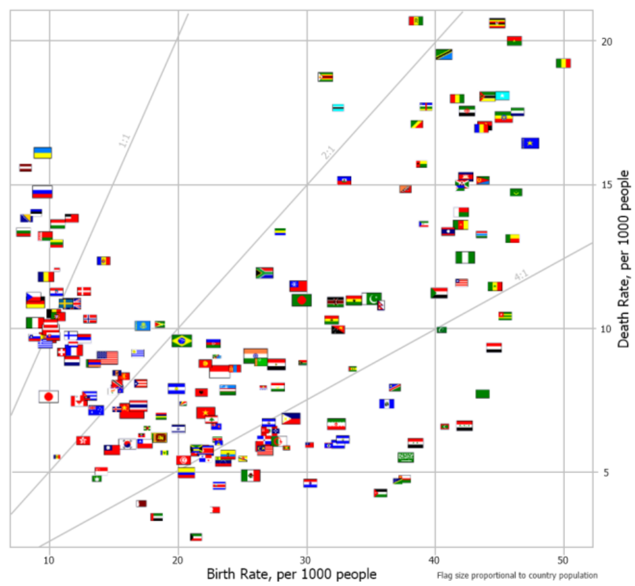

Figure 5. Same data as fig. 3, using country flags.

Figure 5 uses flags instead of dots or alpha-codes. While flags are usable, many people are likely not familiar with more than a dozen flags, thus requiring either a legend or interaction to make the glyphs comprehensible (fig. 6).

Figure 6. Unique glyphs may necessitate a large legend.

Both labels and glyphs can uniquely identify hundreds of items, beyond the fidelity [26] of most other visual channels (e.g. color 10-20). However, the potential advantage of labels over other glyphs (e.g. flags), include:

1. **Mnemonic:** Some text codes may be easy to decipher by their target audience. In the birth rate/death rate scatterplot the top left corner has country codes such as UA, LV, RU, EE, BG which have some similarity to the corresponding country names Ukraine, Latvia, Russia, Estonia, Bulgaria.

2. **Alphasort**: Alphanumeric symbols can be sorted. Given a specific alphanumeric label, e.g. (TV), a legend in sort order aids the viewer to quickly find the target. Conversely given a specific icon of interest (e.g. green, red and green stripe flag), the viewer may need to linearly search through all flags in the legend until the matching entity is found.

3. **Glyph design**: While it may be feasible to design intuitive glyphs, the design task may be difficult when a large number of categorical glyphs are required, e.g. Bertin's map with 59 categories [2, p. 157].

Occlusion and Spatial Separation

In all cases occlusion is an issue. As shown in fig. 7, solid flags (left) obscure the lower flags. With text, the open letterforms provide visibility to lower letters and thin white strokes on the letters helps legibility (fig. 7 center).

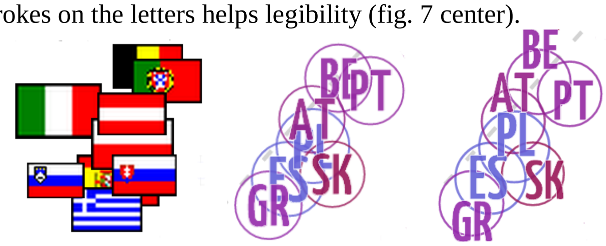

Figure 7. Partial occlusion interferes with identification.

However, there remains perceptual uncertainty with some obscured items. Future work could consider strategies for improving text discrimination when text is overlapping, such as font size, outlines, transparency and color variation.

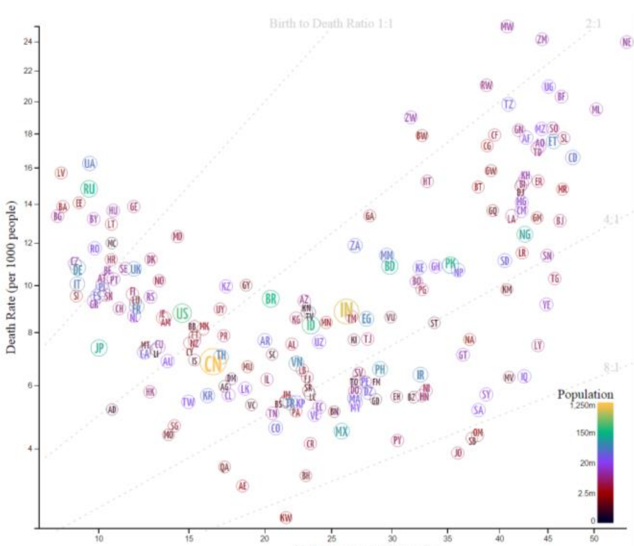

Figure 8. Readability improves without glyph overlap.

Alternatively, clear spatial separation of glyphs will result in better readability. Fig. 7 right shows the same labels algorithmically nudged to reduce overlap while maintaining circle positions to indicate the original data location. The slight deviation of positions does not alter macro-level readings, such as the overall crescent-shaped distribution of the points or individual outliers (e.g. Latvia, LV, top left or Kuwait, KW, bottom) as shown in fig. 8.

Full Labels instead of Mnemonics

Alphanumeric codes maybe effective if available and the intended viewer can readily decode the mnemonic. In some cases, no mnemonic abbreviation is available. For example, fig. 9 shows statistics for U.S. National Parks. Similar to the previous figure, collision detection is used to reduce label overlap improving readability. Narrow fonts are designed for use in tight spaces [6,7] and are used here for labels.

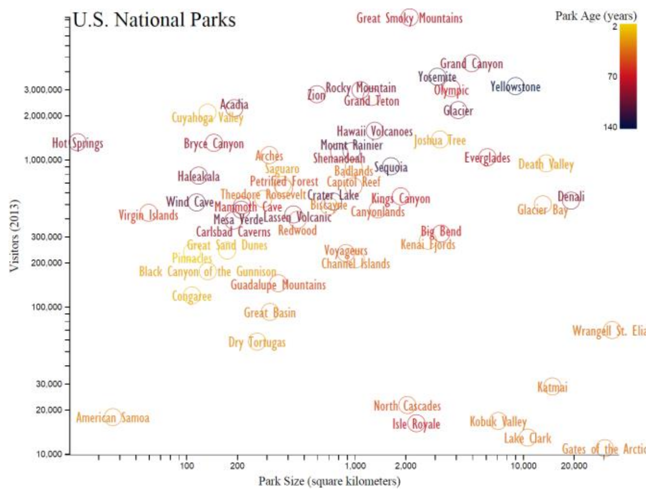

Figure 9. Label scatterplot of US National Parks.

Long labels could potentially hinder macro-level readings such as overly weighting one portion of the plot if long names are concentrated in one area. Furthermore, long labels (e.g. *Black Canyon of the Gunnison*) could have more prominence over short labels (e.g. *Zion*). Cartographers have similar issues labelling maps (e.g. Rome vs. San Francisco) and in practice use a set font weight (e.g. based on city population) regardless of label length.

Large scale macro patterns are visible, such as the large central blob, outlier to the lower left (*American Samoa*) and cluster to the bottom right. Local relationships, such as a pair of very close dots, may become less visible with long labels and the underlying circles may need to be referred to.

FONT ATTRIBUTES AND ALPHANUMERIC MARKS

When using alphanumeric marks, font attributes, such as font weight, font family, underline and italic can be used, in addition to traditional visual attributes such as size of color.

In the previous example (fig. 9) italics indicate location: non-italic for parks in the lower 48 states, italic for Alaska and Hawaii (bottom right) and reverse italic for parks outside the 50 states (e.g. American Samoa, bottom left).

Font Attributes in Sorted Stem and Leaf Plots

Fig. 10 (left) shows a stem and leaf plot [14] indicating volcano heights (stem represents height in thousands of feet, leaves indicate next digit, e.g. 4|6 represents 4600 ft.). By additionally sorting each row, the entire plot can be treated as a sequence and statistically significant values can be encoded with added font attributes. In fig. 10 (right), bold indicates quartiles (and bold italic indicates median); while underline italic indicates mean and further underlines indicates the values for one standard deviation from the mean. From this representation, one can see the median (6|***5*** = 6500 ft.) and mean (7|*0* = 7000 ft.).

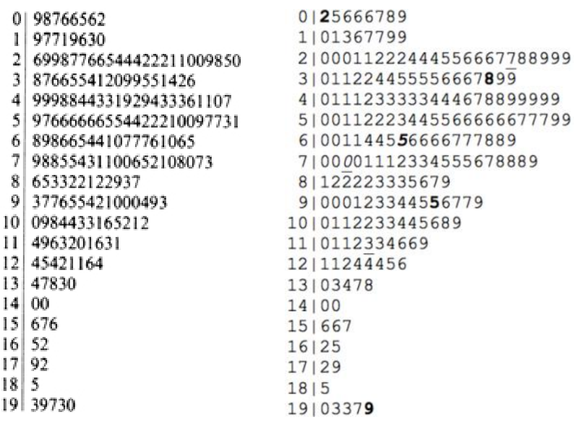

Figure 10. Left: Original Tukey stem and leaf plot [1]. Right: Same plot, sorted and enhanced with quartiles (bold), mean and first standard deviation (underline).

Word Lists, Sentiment Analysis and Descriptor Analysis

Text analytics such as sentiment analysis is popular, with Twitter and news feeds being analyzed to produce sentiment scores and visualizations (e.g. [27,28]). Beyond the raw score, specific words associated with sentiment can be extracted and displayed to provide added context, (e.g., *Tile Apps* depicts Twitter words sized by frequency [29]).

Similar to stem and leaf stacks of numbers, words can be stacked and sequenced for text analytics. Figure 11 shows an analysis of adjectives associated with characters from *Grimms' Fairy Tales* [30]. In this example, adjectives that occur at least 2 times within +/- 3 words of a character are shown in the plot. Each row is a different character and adjectives are ordered left-right with decreasing frequency. Font-weight also indicates adjective frequency, with the lightest weight indicating two occurrences and the heaviest weight indicating ten or more. A horizontal scale is used to indicate the approximate number of words, based on the average word length in the plot to facilitate rapid estimation of the number of words in an adjective list.

At a macro-level, the longest adjective lists in these fairy tales are associated with *birds*, *kings*, *princesses* and *wives*. Font-weight makes more frequently used adjectives visually stand out making a second macro-level reading visible differentiating the more frequent adjectives for each character.

Figure 11. Adjectives associated with fairy tale characters. More frequent adjectives are weighted more heavily.

At a micro-level the individual words can be read. Focusing on characters with frequent adjectives, the most frequent adjectives are as expected: *princesses* tend to be **beautiful** and **young**, whereas *kings* tend to be **old** and **great**, but *girls* are **little** and **poor** while *witches* are **old** (and wicked).

By revealing the details with text, rather than a summarized value represented as a bar, the viewer can form deeper questions. In this example, the character fox has the contradictory adjectives *old* and *young*. Further investigation (e.g. a click or tooltip to show relevant sentences) reveals these adjectives are used to differentiate between specific characters in a tale, such as the old fox vs. the young fox.

Note the horizontal axis and grid lines in fig. 11 are an approximation of the number of words, since words have different lengths. *Hans* is described as **ill** - three short characters - while *birds* are **beautiful** - 8 characters. The axis is based on the average rendered word length, not character count, so that narrow letters and wide letters are accounted for. The degree of error is more noticeable at shorter word lists, e.g. the grid line corresponding to 5 average length words cuts across the word list for *bird* at 3.5 words - an error of 30%. However, the grid line for 10 average length words cuts across *king* at 9.5 words, *princess* at 9 words and *wife* at 10 words, an error of 10%.

Figure 12. Same data as figure 11, oriented at 90 degrees, for an accurate y-axis measuring word count.

If the task is to achieve accurate estimates of word counts, a more accurate representation of the number of words can be created by rotating the plot 90 degrees as shown in fig. 12. The rotated version is less compact as it requires more space between columns such that words do not overlap, but is unambiguous for word counts.

Word Distributions and Opinion Analysis

The same approach can be used for other types of word-based data such as opinion analysis (e.g. [31,32]). For example, movie reviews from many critics may be consolidated into a single score (e.g. Tomatometer at rottentomatoes.com). By summarizing the score into a single value, the dispersion of the scores is not visible and specific critics of interest need to be searched for.

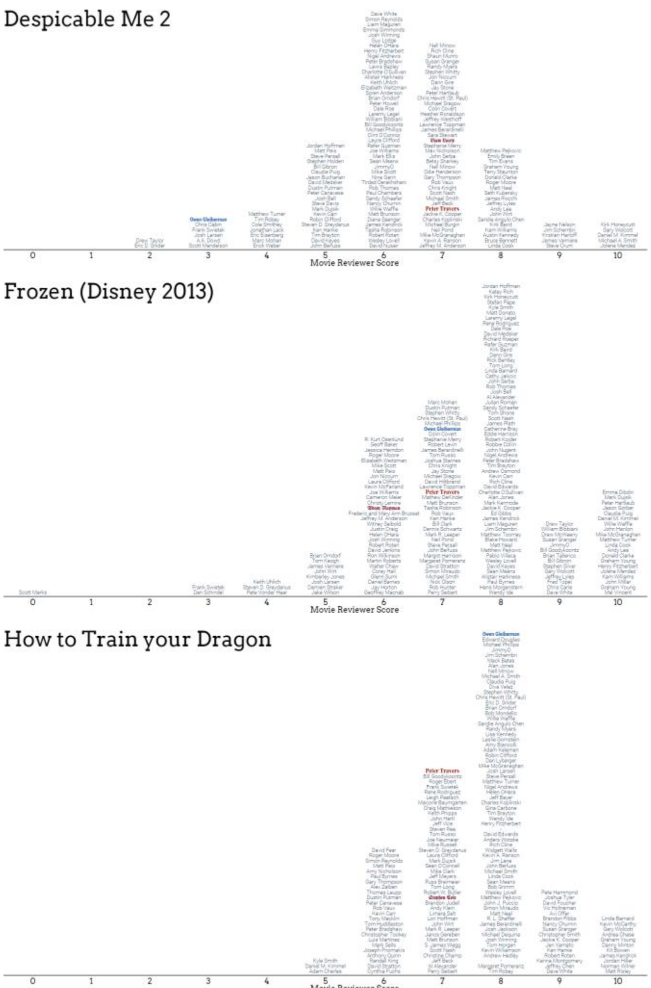

Figure 13. Distribution of movie review scores as stacked names of critics with specific critics indicated via font.

Fig. 13 shows a visual opinion analysis as a distribution of movie critic scores for some popular animated movies. The macro-level view of the distribution is clearly perceivable (e.g. *Despicable Me 2* achieved lower reviews on average than the other two movies); as well as the difference in dispersion (e.g. *Frozen* generated a wider range of critic scores than did *How to Train your Dragon*).

In this example, specific critics are identified with different color and font families in a heavier weight than the surrounding fonts. For example, the *Rolling Stones'* critic Peter Travers is differentiated with a **red high-stress serif font** similar to the magazine's logo, while the critic of the *Globe and Mail* is identified with a **blackletter font** (often used in newspaper titles), as shown in the close-up (fig. 14). One may be able to identify micro-patterns too, e.g., Peter Travers consistently gave all three movies a score of 7.

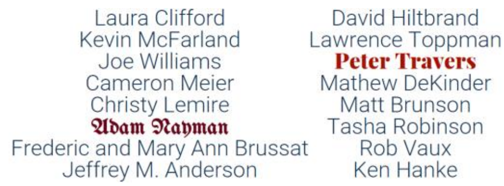

Figure 14. Specific reviewers are identified with unique font families and colors.

Alphanumeric Marks Applied to Other Chart Types

Besides bar charts and scatterplots, this alphanumeric marks approach can also be applied to other familiar representations. For example, instead of choropleth maps and distortion-cartograms, label-based cartograms with font-attributes can be used; or instead of simple node-link graphs, label-based graphs with font attributes can be used [33].

IN CONTEXT VISUALIZATION

As opposed to starting with traditional charts and adding micro-level information; another approach is to start with traditional micro-level text-based representations and layer in macro-level information via font attributes:

Proportional Encoding and Opinion Analysis

Consider movie review data. A reader may be interested in sampling different reviews while understanding the larger context of the overall distribution. Movie review website Rotten Tomatoes provides a key quote for each reviewer, set out on a web page for the user to read a few quotes. However, with 100+ movie reviews without any additional macro-level ordering, the viewer is simply reviewing random quotes. Fig. 15 shows eleven quotes evenly sampled from a list of quotes sorted by score. The *proportion of bold* along the quote indicates the review score, i.e. movie with a score of 4 out of 10 is bold across 40% of its length. In this novel approach, font formatting is not restricted to full words, rather the font formatting is applied to the length of text corresponding to a value. While bold is shown here, any other visual attribute, (e.g. color, font italic, capitals, etc) could be used.

Macro-patterns can be seen in this representation. For example, the overall amount of bold is indicative the overall rating: e.g. *Despicable Me 2* has less bold than *How to Train your Dragon*. The slope formed by the edge of the bold provides an indication of dispersion, e.g. *Despicable Me 2* has a shallower slope than *How to Train your Dragon* indicating higher dispersion for the former.

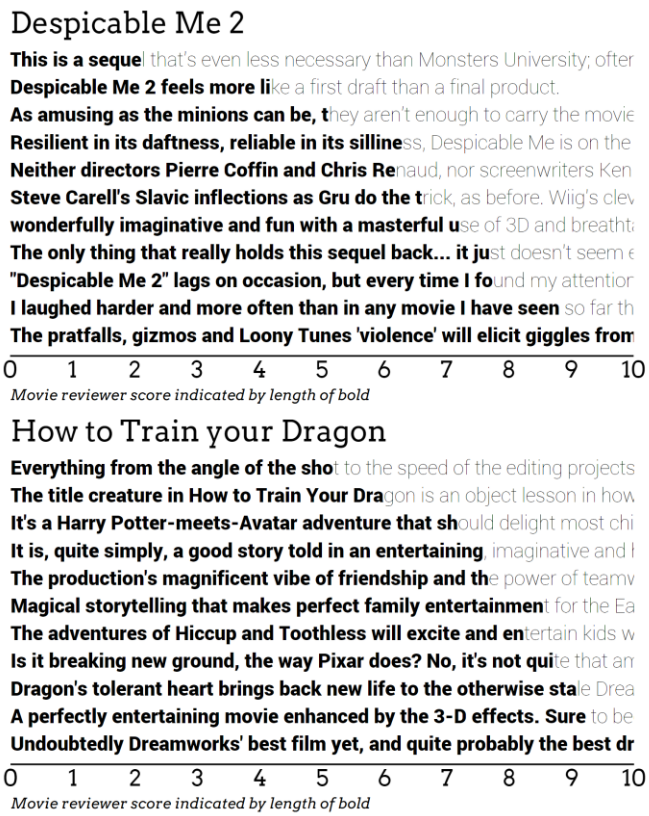

Figure 15. Movie review quotes with review score encoded via length of bold.

Skim Formatting

Text skimming is a reading technique of rapid eye movement across a large body of text to get the main ideas. At a high level, this strategy focuses on the structure of the text, such as the title and first sentence of each paragraph (e.g. www.aacc.edu/tutoring/file/skimming.pdf). At a lower level, the strategy requires the reader to dip into the text looking for words such as proper nouns, unusual words, enumerations, quantifying adjectives and typographic cues.

Fig. 16 shows the opening paragraphs of *The Wizard of Oz*, with font weight inversely applied to English language word frequency (to make uncommon words stand-out). Italics are applied to parts of speech such as prepositions, articles and pronouns to create a Gestalt figure/ground separation (e.g. [34]) pushing those words into the visual background by further differentiating them from the desired parts of speech (e.g. nouns, verbs and adjectives) presented in a plain heavy-weight font.

CONCLUSION

The unique contribution of this paper is to outline a wide range of areas where text and font-attribute techniques can enhance text visualization to achieve micro/macro readings. In particular, by adding micro information via alphanumeric marks or by adding macro information by adjusting font attributes in-context in flowing text representations.

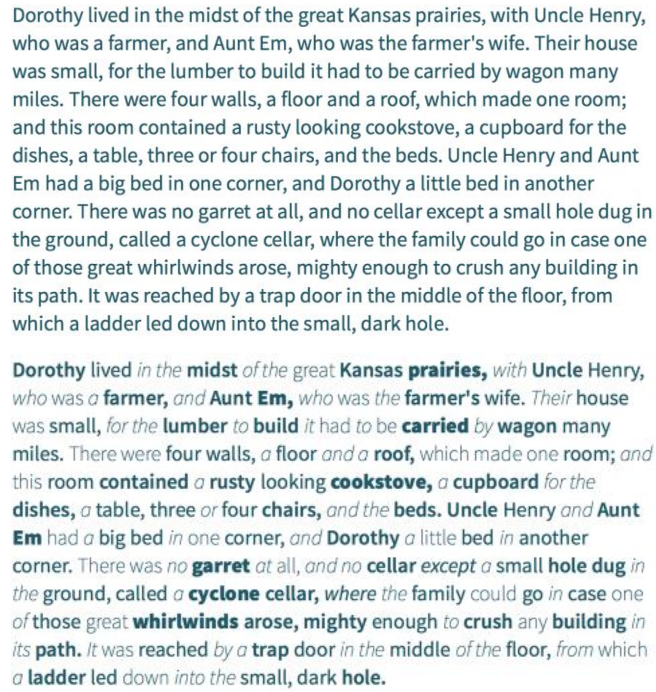

Figure 16. First paragraph of *The Wizard of Oz* without formatting (top) and skim formatting (below) with least frequent words heavily weighted.

Potential micro/macro text visualization applications include character traits, sentiment analysis, visual opinion analysis and skim formatting. The potential benefits include:

1) In all cases, multiple levels of patterns are immediately visible, such as the individual countries and crescent shaped pattern (fig. 4) or individual reviewers and dispersions (fig. 13), without reliance on interaction such as filtering or tooltips which are slower than shifting visual attention. A related benefit is the potential to see serendipitous patterns otherwise not visible, such as the movie reviewer who scored all movies the same (fig. 13) or the declining population in the former east bloc (fig. 8).

2) Information visualization is fundamentally lossy [3] and micro/macro techniques have the potential to reduce lossiness and increase data density. In the U.S. park scatterplot (fig. 9), two additional dimensions are layered in via text and italics, beyond the base scatterplot x, y, color and size.

3) Micro/macro text-based techniques can layer in additional data with potentially preattentive visual attributes in general and font attributes in particular (which are unique to text-based visualization techniques) such as counts and scores (fig. 11, 16), highlights (fig. 14) and statistics (fig. 10). These formats can span across the letters or words to create novel macro-patterns such as proportion of bold along a length of text (fig. 15).

There are many areas for potential future work. Some care is required in the use of label-based techniques, e.g. dealing with variable word length and overlapping labels: both are areas where future work could be done.

Micro/macro is related to proxemics: interfaces that adapt based on the distance of the user to the display. Similarly, maps can also work at more levels than just binary macro/micro levels (e.g. cities, states, countries within in single map). This work is largely binary micro/macro and could be extended to multiple levels of scale and proxemics.

In addition, future work should include usability testing to validate and refine the implied heuristics outlined herein.

REFERENCES

1. Tufte, E. *Envisioning Information*, Graphics Press, Cheshire, CT. 1990.
2. Bertin J. *Semiologie Graphique*. Gauthier-Villars, Paris, 1967.
3. Chen, Min, and Luciano Floridi. An analysis of information visualization. In *Synthese* 190.16 (2013): 3421-3438.
4. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proceedings of the *IEEE Symposium on Visual Languages*, pages 336-343, Washington. IEEE Computer Society Press, 1996.
5. Brath, R. and Banissi, E. The design space of typeface. Visualization and Computer Graphics, IEEE Transactions on (VisWeek 2014). IEEE Computer Society Press, 2014.
6. Lupton E.: *Thinking with Type*. Yale, 2004.
7. Ambrose G. & Harris P. *The Fundamentals of Typography 2nd Edition,* AVA Publishing SA, 2011.
8. Apollinaire, G. *Calligrammes*. Mercure de France, Paris, 1918
9. Carroll, L. *Alice's Adventures in Wonderland.* MacMillan and Co. London. 1866. p. 37.
10. Hoyle. *The Game in Wall Street*. Ogilvie Publishing, New York, NY, 1898.
11. Wyckoff R.: *Studies in Tape Reading*. Noble Offset Printers, New York, NY, 1910.
12. Livermore J.: *How to Trade in Stocks*. Quinn & Boden Company, Rahway, NJ, 1940.
13. Tukey J.: *Some graphic and semigraphic displays. Statistical Papers in Honor of George W. Snedecor*, 1966.
14. Tufte, E. *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT. 1983.
15. Tufte, E. *Beautiful Evidence*, Graphics Press, Cheshire, CT. 1996.
16. Paley B. *Map of Science*, 2006. URL: http://wbpaley.com/brad/mapOfScience/index.html.
17. Ellingham H. *A chart illustrating some of the relations between the branches of natural science and technology*, 1948.
18. Skupin A. The world of geography: Visualizing a knowledge domain with cartographic means. *Proceedings of the National Academy of Sciences of the United States of America 101*, Suppl 1, 5274–5278. 2004.
19. Mackinlay, J. D., Robertson, G., & Card, S. K. The perspective wall: Detail and context smoothly integrated. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 173-176). ACM 1991.
20. Apperley, M. D., Tzavaras, I., & Spence, R. A bifocal display technique for data presentation. In Proceedings of Eurographics 1982.
21. Stoffel, A., Strobelt, H., Deussen, O., & Keim, D. A. Document thumbnails with variable text scaling. In *Computer Graphics Forum*. Blackwell Publishing. June 2012.
22. Buchanan, G., & Owen, T. Improving skim reading for document triage. In *Proceedings of the second international symposium on information interaction in context*. ACM 2008.
23. Baecker, R., & Marcus, A. On enhancing the interface to the source code of computer programs. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM 1983.
24. Borgo R., Kehrer J., Chung D. H., Maguire E., Laramee R. S., Hauser H., Ward M., Chen M.: Glyph-based visualization: Foundations, design guidelines, techniques and applications. In *Eurographics State of the Art Reports, EG STARs*, Eurographics Association. 2013.
25. Brath R.: *Multiple Shape Attributes in Information Visualization: Guidance from Prior Art and Experiments*. IEEE Information Visualization. 2010.
26. Brath, R., & Banissi, E. Evaluating lossiness and fidelity in information visualization. In *IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics*. 2015.
27. Kaye, J. J., Lillie, A., Jagdish, D., Walkup, J., Parada, R., & Mori, K. Nokia internet pulse: a long term deployment and iteration of a twitter visualization. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 829-844). ACM. 2012.
28. Bloomberg Introduces Sentiment Analysis Tools. http://www.tradersdna.com/news/bloomberg-introduces-twitter-sentiment-analysis-tools/. Retrieved 12/29/2014.
29. Cheng, D., Schretlen, P., Kronenfeld, N., Bozowsky, N., & Wright, W. Tile based visual analytics for Twitter big data exploratory analysis. In *Big Data, 2013 IEEE International Conference on* (pp. 2-4). IEEE. 2013.
30. Grimm, J., Grimm, W., Taylor, E. & Edwardes, M. *Grimms' Fairy Tales*. www.gutenberg.org/cache/epub/2591/pg2591.txt, extracted 12/27/2014.
31. Oelke, D., Hao, M., Rohrdantz, C., Keim, D. A., Dayal, U., Haug, L., & Janetzko, H. Visual opinion analysis of customer feedback data. In *Visual Analytics Science and Technology, 2009. VAST 2009.* IEEE 2009.
32. Wu, Y., Wei, F., Liu, S., Au, N., Cui, W., Zhou, H., & Qu, H. OpinionSeer: interactive visualization of hotel customer feedback. In *Visualization and Computer Graphics, IEEE Transactions on*, 16(6). 2010.
33. Brath, R. & Banissi, E. Using Font Attributes in Knowledge Maps and Information Retrieval. In *Proceedings of Knowledge Management and Information Retrieval (KMIR) 2014 Workshop at Digital Libraries,* 2014.
34. Figure and Ground. www.afn.org/~gestalt/fignd.htm, accessed 01/09/2014.