Multi-Perspective Synopsis with Faceted Views of Varying Emphasis

Chris Rooney* Middlesex University
Roger Beecham City University London
Jason Dykes City University London
Cagatay Turkay City University London
Aidan Slingsby City University London
Jo Wood City University London
William Wong Middlesex University

Abstract

Many datasets have multiple perspectives – e.g. time, space and description – and often summaries are required for these multiple perspectives concurrently. A design challenge is to provide such a concurrent summary of perspectives in ways that neither clutter nor visually overload. We report on a framework that helps us do this. We demonstrate its use with an implementation based on a Crime Pattern Analysis case study that produces synoptic summarises of spatial, temporal and descriptive information in crime reports. Our work with crime analysts suggests that the framework offers some potential for Crime Pattern Analysis.

Keywords: Composite Views, Criminal Intelligence Analysis, Small Multiples

1 Introduction

Many datasets have multiple perspectives – e.g. time, space and description – through which they can be considered. Synoptic visual summaries of such data are challenging as they require representations across these multiple perspectives. The design challenge is to provide space-efficient representations that do not clutter visually or confuse cognitively [8, 5]. We develop and report on a framework that helps support such complex representations.

2 Framework and Design

Our *framework* assumes that data are structured as sets of records; each with fields that can be organised into a small number of 'data perspectives'. Our *design* that implements the framework was developed under our Crime Pattern Analysis case study. Each record is a crime and each field relates to either a geographic space, time or descriptive perspective. The following are characteristics of our framework.

- **Faceting.** Conditioning (faceting) a dataset on one of its attributes results in subsets, each of which can be represented by a synoptic overview. Juxtaposing these synoptic overviews using appropriate ordering enables subsets to be compared.
- **Simultaneous multi-perspective synoptic overviews.** Each faceted subset can be summarised through multiple perspectives. We do this simultaneously – in our case, superimposing them.
- **Varying emphasis.** Superimposing multiple perspectives simultaneously will inevitably result in occlusion and clutter. Our solution is to de-emphasise the perspectives not central to the task at hand. Pushing these to the background prevents them from occluding or cluttering the display but retains potentially-important context.

*e-mail:c.rooney@mdx.ac.uk

3 Design

Specific designs that implement this framework may be domain and task dependent. Our design is based on our Crime Pattern Analyst (CPA) case study, in which geographical, temporal and other descriptive characteristics of different sets of crimes need to be compared. Faceting crime records into meaningful subsets, providing synoptic overviews of these, and juxtaposing these, can help analysts compare and contrast the subsets.

From requirements workshops that we ran with regional police analysts in the UK and Belgium, it became clear that analysts need *(a)* overviews that simultaneously summarise multi-perspective aspects of crimes; and *(b)* the ability to split visual summaries into collections and make comparisons across these collections.

Data on crime incidents were collected from the City of Chicago Data Portal (`http://bit.ly/1GmjXCQ`). Each incident has data for each of the three perspectives of geographical space (where it was), time (when it was) and description (other characteristics).

3.1 High abstraction level: summaries of central tendency

Figure 1 shows the abstraction levels we developed. At the most abstract level we use *central tendancy* to give a sense of structure without exposing detail [6]. The *spatial perspective* is represented by standard deviational ellipses [11]. The *temporal perspective* uses polar coordinates to represent cyclic aspects of day and hour with *circular statistics* [3] to represent central tendency and a boxplot to

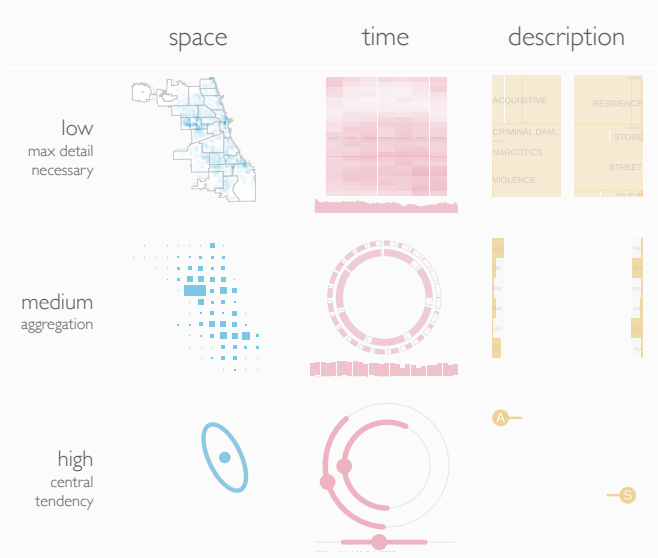

Figure 1: Different designs for each level of abstraction: low (central tendancy), medium and high abstraction (maximum details necessary).

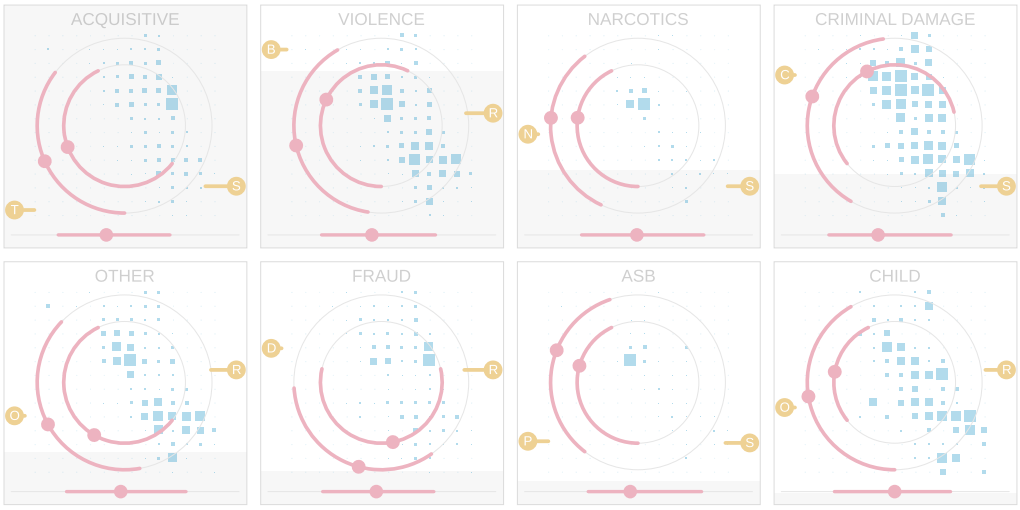

Figure 2: Multi-perspective synoptic overviews of ∼17,000 crime reports, faceted on eight *crime types* in which temporal (red), spatial (blue) and descriptive (brown) perspectives are depicted with *varying levels of emphasis* (low temporal, low descriptive & high spatial emphases).

summarise the temporal distribution as linear time. The *descriptive perspectives* appear at the margins with line length indicating relative entropy [1].

3.2 Medium abstraction level: data aggregation and binning

At the medium level, we move away from statistical summaries and expose more detail by aggregating or binning data [2] and using daily and hourly frequencies [3]. In the *spatial perspective*, we show local densities in a regular grid. For the *temporal perspective*, we use length to show the relative number of crimes occurring by day of week and hour of day and a histogram to summarise frequencies across those days. The *descriptive perspective* remains at the right and left margins, but shows more detail.

3.3 Low abstraction level: maximum details necessary

At the lowest level of abstraction, we attempt to expose as much detail as is desirable or necessary. For the *spatial perspective*, we use kernel-density-estimation (KDE) to create a continuous surface of spatial densities [9]. For the *temporal perspective*, we create a two dimensional 'calendar' showing crime volume by hour of day and day of week. In the *descriptive perspective* we reveal the most detailed sub-type description and embed sub-location descriptions inside the bars forming spine plots [7] which help comparison of proportional differences across parent categories.

4 Evaluation and Further Possibilities

We realised our framework through a prototype (Figure 2) and ran an evaluation with crime analysts and results suggested that although some of the highly schematised views were difficult to interpret immediately, the approach itself was considered to be useful for CPA.

We also have ideas about how these overviews may be used in other contexts, including: as bookmarks (thumbnails that summarise results of a faceted query; e.g. [10]), interactive selection (dynamically summarising interactive selection of records) and as "probes" [4] for summarising records at particular spatial or temporal locations.

5 Conclusion

We present a framework for multi-perspective overview and design graphics for assisting with crime pattern analysis that implements the framework. We are continuing to work with crime analysts to develop our design and hope to report the results.

Acknowledgements

We thank colleagues at West Midlands Police (UK) and Antwerp Local Police (Belgium). This research was funded by the EU under the EC Grant Agreement No. FP7-IP-608142, awarded to Middlesex University and partners.

References

[1] J. Alsakran, X. Huang, Y. Zhao, J. Yang, and K. Fast. Using Entropy-Related Measures in Categorical Data Visualization. In *2014 IEEE Pacific Visualization Symposium (PacificVis)*, pages 81–88, Mar. 2014.

[2] N. Andrienko and G. Andrienko. *Exploratory Analysis of Spatial and Temporal Data*. Springer-Verlag, Berlin, Germany, 2006.

[3] C. Brunsdon and J. Corcoran. Using circular statistics to analyse time patterns in crime incidence. *Computers, Environment and Urban Systems*, 30(3):300–319, 2006.

[4] T. Butkiewicz, W. Dou, Z. Wartell, W. Ribarsky, and R. Chang. Multi-focused geospatial analysis using probes. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1165–1172, 2008.

[5] G. Ellis and A. Dix. A Taxonomy of Clutter Reduction for Information Visualisation. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1216–1223, Nov. 2007.

[6] N. Elmqvist and J. Fekete. Hierarchical Aggregation for Information Visualization: Overview, Techniques, and Design Guidelines. *IEEE Transactions on Visualization and Computer Graphics*, 16(3):439–454, May 2010.

[7] J. Hummel. Linked bar charts: Analysing categorical data graphically. *Computational Statistics*, 11:23–33, 2007.

[8] W. Javed and N. Elmqvist. Exploring the design space of composite visualization. In *Pacific Visualization Symposium (PacificVis), 2012 IEEE*, pages 1–8. IEEE, 2012.

[9] D. O'Sullivan and D. Unwin. *Geographic Information Analysis*. John Wiley & Sons, New Jersey, USA, 2002.

[10] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. H. Nguyen, D. Stephens, B. W. Wong, and Y. Zheng. An extensible framework for provenance in human terrain visual analytics. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2139–2148, 2013.

[11] R. Yuill. The standard deviational ellipse: An updated tool for spatial description. *Geografiska Annaler. Series B, Human Geography*, 53(1):28–39, 2011.