Scalable Integration of Analytics and Visualization: Papers from the 2011 AAAI Workshop (WS-11-17)

Scalable Visualization Resizing Framework

Yingcai Wu and Kwan-Liu Ma
VIDI Research Group, University of California, Davis

Abstract

Effective visualization resizing is important for many visualization tasks, where users may have display devices with different sizes and aspect ratios. Our recently designed framework can adapt a visualization to different displays by transforming the resizing problem into a non-linear optimization problem. However, it is not scalable to a large amount of dense information. Undesired cluttered results would be produced if dense information is presented in the target display. We present an extension to our resizing framework with a seamless integration of a sampling-based data abstraction mechanism, such that it is scalable with not only different display sizes, but also different amounts of information.

Introduction

Mobile computing devices such as smartphones and tablets have become more and more powerful. There is, therefore, a growing trend in utilizing mobile devices for visualizing complex data from anywhere and at anytime (Kim et al. 2007; Pattath et al. 2006). However, visualizations have yet to be adopted extensively on mobile devices, because the small screens of these devices limit the use of visualizations (Chittaro 2006). Furthermore, most existing visualizations are specifically created for large displays. They do not consider the different display sizes and screen aspect ratios of different devices that are simultaneously used in one collaborative application. This deficiency motivates us to study the problem of how to resize visualizations in order to fit different device displays in a collaborative application such that embedded useful patterns in the resized visualizations can still be revealed as effectively as before.

In our recent work, we have designed and developed an automatic resizing framework - ViSizer (Wu et al. 2011). The framework was built upon an image warping approach called optimized scale-and-stretch (Wang et al. 2008) that scales important regions uniformly and distorts homogeneous context. With the framework, a visualization can be effectively and automatically scaled to any display size with an arbitrary aspect ratio. In other words, ViSizer ensures that a visualization is scalable with different display sizes and aspect ratios. Although we have demonstrated the effectiveness of ViSizer in our previous work, the use of the framework is still constrained to the density of information in the target display. A visualization could become cluttered during the visualization resizing process with increasing information density of the display. For instance, Fig. 1(b)-(d) show results generated by resizing Fig. 1(a) with the framework. Although the graph structure is well preserved in different displays, the graph nodes and edges become too dense to be distinguishable.

This problem could be solved to some extent by data abstraction techniques (Rafiei and Curial 2005; Ellis and Dix 2007; Schneiderman 1994), which can reduce visual clutter and provide more space for important visual elements. Focus+Context Visualization techniques such as Fisheye (Furnas 1999) and bifocal display (Spence and Apperley 1982) allow users to visualize data on display screens. They distort the visualization space unevenly through magnification functions, allocating more space for salient regions. However, these methods often require additional overhead and ignore different aspect ratios of display screens. They also depend on a user's ability and intuition, or the tendency of the visualization to determine the amount of information to be discarded (Rosenholtz et al. 2005).

In this work, we are extending our resizing framework to address the aforementioned concerns, such that it is scalable with respect to not only different display sizes, but also different amounts of information. Specifically, we are seamlessly integrating a sampling-based data abstraction method into the framework. The main advantage of this approach, compared with other sampling and filtering techniques (Rafiei and Curial 2005; Schneiderman 1994), is that the information density can be automatically adjusted based on the actual display size as well as the quantitative visual clutter magnitude in every local region of the visualization.

Scalable Resizing Framework

Resizing Framework The resizing framework includes a preprocessing component and an optimization component. The preprocessing component first partitions an input visualization with a uniform grid. A significance map, a combination of a degree-of-interest (DOI) map and a visual clutter map, is then created to encode the significance values of every quad in the grid. By combining these two maps, the

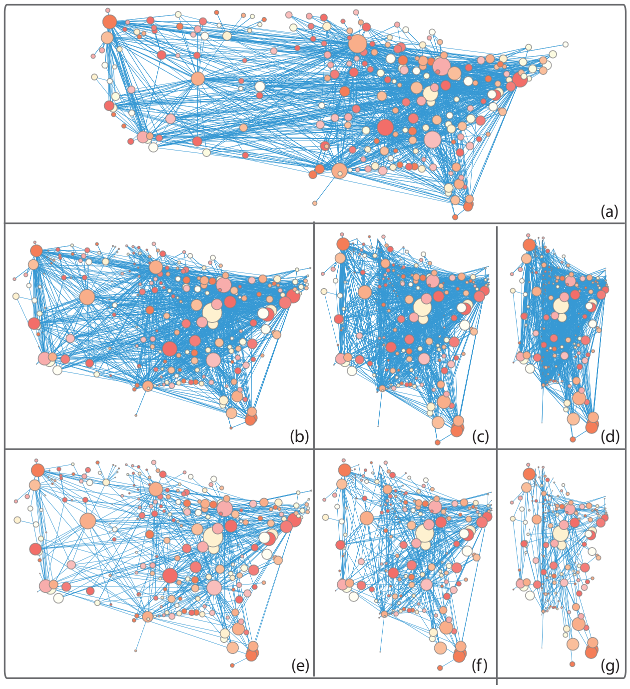

Figure 1: (a) Original network of major US airlines. (b)-(d) Resized results without information density control. (e)-(g) Resized results with information density control.

visualization can be automatically resized to guarantee that important and cluttered regions would largely be preserved while context and uncluttered regions would be deformed. Next, the significance map is used to deform the initial grid and to create a significance-aware grid in which significant regions can be covered by more quads to reduce linearization artifacts.

In the optimization component, we transform the resizing problem into a nonlinear least squares optimization problem through an energy function. The energy function is formulated by considering the significance map, quad deformation, and edge bending. It aims to minimize the deformation of the significant regions and preserving the spatial relations between visual elements. The optimization problem is further transformed to a set of linear equations, which can be iteratively solved to find the optimal solution. When each iteration starts, a new scaling factor will be computed for every quad. This may lead to an undesired distortion to a feature because its covering quads may have varying scaling factors. Thus, the scaling factors are optimized to obtain a new set of smooth scaling factors and thus minimize potential distortion. The iteration repeats until a certain convergence condition is reached, i.e., all vertex movements are very small in the current iteration. Finally, the optimization generates a deformed grid for adjusting the visualization accordingly by interpolation. We have successfully used this framework to resizing different types of visualizations such as graphs, word clouds, line charts, and Treemaps.

Scalability Extension Our resizing framework can scale a visualization to different display sizes and aspect ratios. However, when information in a visualization becomes too dense, it would be challenging and sometimes even impossible for the framework to work properly. This is because the dense information usually leads to a high degree of visual clutter, preventing the framework from finding appropriate regions with less clutter for deforming the visualization. We are extending the framework to improve its scalability by automatically controlling information density such that it can also be scalable with different amounts of information.

A random sampling method is employed to adjust the information density of a visualization, since it is more general and can be used in various types of visualization (Rafiei and Curial 2005). This method determines which items to be removed according to the significance map W. It ensures that salient items will mostly be retained during the random sampling process while less important items will have a higher chance to be removed. The significance map is estimated by a combination of a DOI map DOI and a visual clutter map C, i.e., $W = DOI * C$. We follow Furnas's work (Furnas 1986) and define a DOI function for an item x given the user's focus item y as: $DOI(x|y) = \alpha{\cdot}API(x)+\beta{\cdot}D(x,y)$, where the $API(x)$ represents the general importance of x and the $D(x,y)$ computes the distance from x to y. The clutter map C can be obtained by an efficient quantitative method called Feature Congestion (Rosenholtz et al. 2005). It derives the clutter level by a statistical saliency model based on the observation that unusual items are usually salient. The statistical saliency for a feature vector X can be defined as

$$\Delta = \sqrt{(X-\mu)^T S^{-1}(X-\mu)} \tag{1}$$

where μ and S denote the mean matrix and covariance matrix of the local distribution of the feature vectors.

Given an item i within a grid quad (x, y), we assign the item a probability $p_i = 1-W(x,y)$ for indicating how likely the item will be removed. The DOI and clutter maps are updated at each iteration of the resizing optimization process. We denote the original and target display sizes as S_o and S_t. We also assume that $S_t < S_o$. The amount of information to be removed in the target display could be simply determined by $1 - S_t/S_o$. We randomly filter out the items according to p. The items with lower p are largely maintained while those with higher p have higher chances to be filtered out by the random sampling methods, such that the overall patterns of the visualization could be preserved. Fig. 1 (e)-(g) show our preliminary results created by enabling the dynamic and random sampling mechanism in the resizing optimization process. By comparing Fig. 1 (b)-(d) and (e)-(g), we can see that our new method can be successfully applied to a visualization with a high degree of information density.

Conclusions

In this paper, we have introduced a scalability extension to the resizing framework. The goal of the work is to adapt a visualization to any display size with arbitrary information density. Although the work is still ongoing, the preliminary results look promising. In the future, we will continue to study the scalability issue of the visualization resizing problem and validate our technique with different visualizations such as word clouds and stacked graphs.

Acknowledgments

This research was supported in part by the HP Labs, AT&T Labs Research, and the U.S. National Science Foundation through grants CCF-0808896, CNS-0716691, CCF 0811422, CCF 0938114, and CCF-1025269.

References

Chittaro, L. 2006. Visualizing information on mobile devices. *Computer* 39(3):40–45.

Ellis, G., and Dix, A. 2007. A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics* 13(6):1216–1223.

Furnas, G. 1986. Generalized fisheye views. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 16–23.

Furnas, G. W. 1999. The fisheye view: a new look at structured files. In *Readings in Information Visualization: Using Vision to Think (Interactive Technologies)*. Morgan Kaufmann.

Kim, S. Y.; Jang, Y.; Mellema, A.; Ebert, D. S.; and Collinss, T. 2007. Visual analytics on mobile devices for emergency response. In *IEEE Symposium on Visual Analytics Science and Technology*, 35–42.

Pattath, A.; Bue, B.; Jang, Y.; Ebert, D.; Zhong, X.; Ault, A.; and Coyle, E. 2006. Interactive visualization and analysis of network and sensor data on mobile devices. In *IEEE Symposium On Visual Analytics Science And Technology*, 83–90.

Rafiei, D., and Curial, S. 2005. Effectively visualizing large networks through sampling. In *Proceedings of IEEE Visualization*, 23–28.

Rosenholtz, R.; Li, Y.; Mansfield, J.; and Jin, Z. 2005. Feature congestion: A measure of display clutter. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 761–770.

Schneiderman, B. 1994. Dynamic queries for visual information seeking. *IEEE So* 11(6):70–77.

Spence, R., and Apperley, M. 1982. Data base navigation: An office environment for the professional. *Behaviour and Information Technology* 1(1):43–54.

Wang, Y.-S.; Tai, C.-L.; Sorkine, O.; and Lee, T.-Y. 2008. Optimized scale-and-stretch for image resizing. *ACM Transactions on Graphics* 27(5).

Wu, Y.; Liu, X.; Liu, S.; and Ma, K.-L. 2011. ViSizer: a framework for visualization resizing. In *IEEE Information Visualization (Submitted)*.