

Progressive Horizon Graphs: Improving Small Multiples Visualization of Time Series

Charles Perin, Frédéric Vernier, Jean-Daniel Fekete

▶ **To cite this version:**

Charles Perin, Frédéric Vernier, Jean-Daniel Fekete. Progressive Horizon Graphs: Improving Small Multiples Visualization of Time Series. IEEE Conference on Information Visualization (INFOVIS), IEEE, Oct 2012, Seattle, Washington, USA, United States. hal-00734497

HAL Id: hal-00734497

https://inria.hal.science/hal-00734497

Submitted on 22 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Progressive Horizon Graphs: Improving Small Multiples Visualization of Time Series

Charles Perin* INRIA and LIMSI-CNRS
Frederic Vernier† LIMSI-CNRS
Jean-Daniel Fekete‡ INRIA



(a) Reduced line charts.

(b) Horizon graphs with standard baseline at the half of the y axis and with two bands.

(c) Progressive horizon graphs with the best baseline and a zoom factor of ten

Figure 1: Three time series visualization techniques: *reduced line charts*, *horizon graphs*, and *progressive horizon graphs* performing the *discriminate* task: "find the time series having the highest value of the three marked points."

ABSTRACT

Many approaches have been proposed for the visualization of time series. The *reduced line charts* (small multiples for time series) and the more recent *horizon graphs* are two of these visualization techniques with benefits for visualizing multiple time series that we propose to unify, using a variant of the pan and zoom interaction on the *y* axis. We compare in a user study *reduced line charts*, *horizon graphs*, and our own contribution—*progressive horizon graphs*—for different tasks and numbers of concurrent time series using datasets with small variations. While recent work has compared *horizon graphs* with others visualization techniques and has made some recommendations on their usability, the real advantages of this technique are not clear. The results of our controlled user study show that *progressive horizon graphs* overcome these two visualization techniques when the number of charts increases.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces (GUI);

1 INTRODUCTION

Time-varying data is predominant in a wide range of domains such as finance (*e. g.* stock prices) and science (*e. g.* climate measurements, medicine). Line charts is one of the most frequently used statistical data graphic and the simplest way of representing time series. However this visual representation, as well as the others, encounter limits for visualizing multiple time series. This article introduces *progressive horizon graphs*, an interactive technique using a variant of pan and zoom for visualizing multiple time series that we designed to scale on the number of concurrent time series one person can monitor and explore efficiently.

2 PROGRESSIVE HORIZON GRAPHS

2.1 Context

Javed et al. classified visualization techniques for multiple time series into two categories [3]. In *shared-space* techniques, time series are overlaid in the same space (e.g., *line graphs*, *braided graphs*, *stacked graphs*). In *split-space* techniques, the space is vertically divided by the number of time series and each time series occupy its own reduced space (*e. g. reduced line charts*, *horizon graphs*). Shared-space techniques having a limit to the number of time series they can handle, we only consider split-space techniques.

Small multiples for time series is a split-space technique which consists of drawing a *reduced line chart* (*RLC*) for each time series by splitting the space into individual line graphs (see Figure 1(a)).

Horizon graphs [1, 4] (*HG*) is a recent split-space technique invented to display large numbers of time series. This technique uses two parameters: the number of bands and the value of the baseline separating vertically the chart in positive and negative values. Data values are represented not only by their vertical height, but also by their color hue and intensity (see Figure 1(b)).

Our work is closely related to two recent studies: Heer et al. [2] evaluated the role of *HG* parameters, focusing on the performance evaluation of the technique. For their *discriminate* task they provide some recommendations, such as the optimal chart height for *HG*. They also show that the number of bands should be less than three and pair. Nevertheless, this recommendation about the number of bands is due to the task, requiring the participants to estimate the value of the time series at a specific point. They limited their study to two simultaneous time series and the number of bands to four. Javed et al. compared *HG* with others visualization techniques for several numbers of concurrent time series (2, 4 and 8 in their main user study and up to 16 in their follow-up) [3]. They limited the *HG* parameters to the recommended ones and did not highlight any considerable advantage of the technique. Moreover, no previous study considered interaction techniques to improve *HG*.

2.2 Technique design

Progressive horizon graphs (*PHG*)(see Figure 1(c)) is an interactive technique designed to control the two parameters of the *HG*: the baseline is controlled through a variant of *panning* and the number of bands through a variant of *zooming*.

2.2.1 Panning: Controlling The Baseline

HG cuts the chart in half, separating positive and negative values. The drawback of a fixed baseline is that the pre-attentive color perception is not always efficient. For instance, if all the values are on one side of the baseline, then only one hue is used (see Figure 1(b)).

In our approach, we call panning the translation of the baseline along the *y* axis. Since the baseline is always at the bottom of the chart, the perceived translation is inside the chart itself and causes no loss of information. Panning is particularly valuable if one is interested in visualizing a time series around a specific value, for instance to observe the human temperature around the normal value of a given patient. With the baseline separating the values below it in blue and above in red, finding the values crossing the specified baseline becomes an easy red/blue pattern finding task, and finding

*e-mail: charles.perin@inria.fr
†e-mail: frederic.vernier@limsi.fr
‡e-mail: jean-daniel.fekete@inria.fr

a maximum value becomes a comparison of level of red interleaved with *y* estimation.

2.2.2 Zooming: Controlling The Number Of Bands

Heer et al. have studied the impact of the number of bands in *HG* [2]. Their results were that time and error increase with the number of bands. However, these results were obtained for values estimation tasks and they aptly noticed that these increases were due to the mental math implied. Nevertheless, if we do not consider this specific kind of task and require as answer only to select a time series, we push away the limits in the number of bands.

Specifying number of bands implies sudden transitions between two views of the same *HG* since the scale changes abruptly. The interaction we propose prevents these abrupt changes by introducing a smooth and continuous zooming interaction. While standard zooming techniques consist in focusing on a specific area, losing the context information, our zooming implementation preserves both the visibility of the context and the details around the baseline.

2.2.3 Example

To illustrate the effectiveness of our technique, let's consider the basic task of finding the global maximum over multiple time series. This is accomplished in two steps: first, the baseline is set at its maximum so that all the values are colored blue. Then, the value of the baseline is progressively decreased until red values appear in some charts. If there are several candidates, zooming will grow these areas and the differences in magnitude will be visible.

2.3 Evaluation

The purpose of our experiment was to determine the usefulness of adding interactivity to *HG*. More specifically, we were interested in the limits of *RLC* and *HG* with high numbers of time series with small variations (where the derivatives are on the whole small, avoiding high frequencies) in a small space. To evaluate the impact of our interaction technique, we designed a user study and measured the *time*, the *correctness* and the *error magnitude* ($error_{mag}$) for all combinations of visualization technique *V* and number of concurrent time series *N* with different tasks *T*. We used real stock time series with small variations, because such data have not been well studied; our pilots highlighted that they are appropriate datasets to discriminate the visualization techniques; and such datasets are common in a wide range of domains (*e. g.* finance, network logs).

2.3.1 Experiment Factors

Below are detailed our three experimental factors *V*, *T* and *N*.

***V**: RLC*, *HG* and *PHG*.

T: With respect to previous studies, we evaluated three tasks: *Max* consists in comparing multiple time series values at a shared marked point and determining in which one the highest value is. *Disc* is similar to *Max*, but with each time series having its own particular marked point (Figure 1). *Same* consists in picking the time series being exactly the same than a separated reference on. Note that we did not measure the $error_{mag}$ for *Same*.

N: We considered 2 and 8 concurrent time series (*N2* and *N8*) to compare our results with previous ones but also 32 time series *(N32)* and went deeper in the study of split-space techniques scaling.

2.3.2 Results

We applied a log transform to the measures of time and the trials followed the normal distribution. We analysed using ANOVA and the Bonferroni adjustments for pair-wise means comparison.

We briefly detail now the most important results we obtained.

For low numbers of time series (*N2*), participants were slower using *PHG* than using *HG* and *RLC*. This result is due to the users interactions, making them waste time but bringing no benefit.

For medium numbers of time series (*N8*), the interesting result is that *PHG* had higher *correctness* (17% and 20.5% better for *Max* and *Disc*, respectively), and lower $error_{mag}$ than *RLC*. Nevertheless, we did not find any significant result between *PHG* and *HG*.

For high numbers of time series (*N32*), we found the following: for *Max*, *PHG* and *HG* had significantly more *correctness* than *RLC*. Mean *correctness* for PHG was 3% higher than for *HG* and 48% higher than for *RLC*. $error_{mag}$ was also significantly lower for *PHG* and *HG* than for *RLC*. For *Same*, *PHG* had significantly more *correctness* than *RLC* (31% higher). For *Disc*, *PHG* had significantly more *correctness* than *HG* (17% higher) and *RLC* (41% higher) and we obtained strict relationships between all three techniques: each time, the *correctness* as well as the $error_{mag}$ were better for *PHG*, then for *HG*, and finally for *RLC*. *N32* is the only number of time series (and the highest) we tested involving clear differences between the three techniques. These differences had not been highlighted in previous studies [3] and are explained by the features of our data, *i. e.* time series with low variations.

Based on the results, we suggest the following design guidelines: *RLC* are acceptable for low values of *N*. *HG* are acceptable for low to medium values of *N*. For high numbers of *N*, we recommend the use of interactive techniques such as the *PHG* we present. Nevertheless, because *PHG* embeds both *RLC* and *HG*, our technique can be used for low, medium, and high values of *N*.

3 Conclusion and Future Work

We have presented *progressive horizon graphs*, an efficient interactive technique relying on pre-attentive features which unifies two split-space visualization techniques for multiple time series: *RLC* and *HG*. In a user study, we found that the limits of *PHG* are larger than for the others thanks to interactive control of the parameters (position of the baseline and number of bands). We also found that *RLC* scale less well than *HG* and that *PHG* scale to at least 32 concurrent time series. We highlighted not only a significant effect of technique on *correctness*, but an important one with strong differences (*PHG* had 17% more correctness than *HG* and 41% more than *RLC*) whereas previous studies did not find it and this is due to our dataset's properties as well as the higher number of time series.

Future work will entail considering more than 32 time series using more specialised hardware such as wall-sized screens. We also identified that automatic parametrization of *HG* rarely led to acceptable visualizations for time series with small variations. We proposed a pan and zoom variant to adjust these parameters but other interactive techniques such as brushing and zooming or rectangular selection can be linked with an automatic process to determine the best values for the baseline and the number of bands.

Finally, the possibility of switching between two visualization techniques in a smooth way allows the use of the advantages of either and the unification of other time series visualizations such as *braided graphs* and *stacked graphs* offers promising perspectives.

References

[1] S. Few. Time on the horizon. available online at `http://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf`, Jun/Jul 2008.

[2] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In *CHI '09*, pages 1303–1312, 2009.

[3] W. Javed, B. McDonnel, and N. Elmqvist. Graphical perception of multiple time series. *TVCG '10*, 16(6):927–934, 2010.

[4] H. Reijner. The development of the horizon graph. available online at `http://www.stonesc.com/Vis08_Workshop/DVD/Reijner_submission.pdf`, 2008.