Stroscope: Multi-Scale Visualization of Irregularly Measured Time-Series Data

Myoungsu Cho, Bohyoung Kim, Hee-Joon Bae, and Jinwook Seo

Abstract—For irregularly measured time-series data, the measurement frequency or interval is as crucial information as measurements are. A well-known time-series visualization such as the line graph is good at showing an overall temporal pattern of change; however, it is not so effective in revealing the measurement frequency/interval while likely giving illusory confidence in values between measurements. In contrast, the bar graph is more effective in showing the frequency/interval, but less effective in showing an overall pattern than the line graph. We integrate the line graph and bar graph in a unified visualization model, called a ripple graph, to take the benefits of both of them with enhanced graphical integrity. Based on the ripple graph, we implemented an interactive time-series data visualization tool, called Stroscope, which facilitates multi-scale visualizations by providing users with a graphical widget to interactively control the integrated visualization model. We evaluated the visualization model (i.e., the ripple graph) through a controlled user study and Stroscope through long-term case studies with neurologists exploring large blood pressure measurement data of stroke patients. Results from our evaluations demonstrate that the ripple graph outperforms existing time-series visualizations, and that Stroscope has the efficacy and potential as an effective visual analysis tool for (irregularly) measured time-series data.

Index Terms—Irregularly measured time-series data, frequency-aware visualization, uncertainty visualization, long-term case study

1 Introduction

Growing needs of finding important patterns and trends in time-series data in various domains have spurred the development of many interactive visual exploration tools: Line Graph Explorer [17], LiveRAC [20], SignalLens [16], and Data Vases [30] to name a few. As the size and complexity of time-series data increase, visualization researchers now face new challenges and requirements for design of the interactive visual exploration tools. In this paper, we identify some of such challenges and requirements and address them in a new interactive visualization tool called Stroscope.

Most existing time-series data visualization techniques or tools assume that (1) each time-series data is measured regularly over time and (2) all time-series data have the same measurement range. However, there are often the cases that the time-series data do not meet the conditions, e.g., online auction data, regional rainfall/snowfall data, and credit card usage data. For these data, existing visualization tools do not help us much in answering questions regarding frequency-related patterns or trends: For example, (a) which item has the greatest number of bids? (b) in which city does it seldom rain in summer? (c) what is the difference in monthly credit card usage pattern in terms of the amount and frequency?

- *M. Cho and J. Seo are with the Department of Computer Science and Engineering, Seoul National University, Seoul, Korea. E-mail: msdanbi@hcil.snu.ac.kr, jseo@snu.ac.kr.*
- *B. Kim is with the Department of Radiology, Seoul National University Bundang Hospital, Seongnam-si, Gyeonggi-do, Korea. E-mail: bhkim@snubh.org.*
- *H.-J. Bae is with the Department of Neurology, Seoul National University Bundang Hospital, Seongnam, Gyeonggi, Korea. E-mail: braindoc@snu.ac.kr.*

Manuscript received 29 Aug. 2013; revised 13 Nov. 2013; accepted 13 Dec. 2013; date of publication 8 Jan. 2014; date of current version 5 Mar. 2014. Recommended for acceptance by J.-D. Fekete.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2013.2297933

These kinds of data are fairly abundant in the medical domain as well. The examples include, but are not limited to, body temperature, blood sugar level, and blood pressure level in patient records, where the number/interval of measurements and hospitalization time can vary depending on patients' condition. Researchers in the medical domain often have to arduously collect these data to formulate and test hypotheses. Although they may rely on conventional statistical software or Excel for that matter, exploratory analyses for hypothesis formulation are not easy, not to mention that it is neither intuitive to use nor easy to understand the results.

Not much effort has been put into developing visualization models or tools for such data sets in the infovis community. Aris et al. [5] called these data unevenly-spaced time-series data and suggested four representations (i.e., sampled events, aggregated sampled event, event index, and interleaved event index) for the interactive visual exploration of such data. All of them basically regularize the measurement interval, which could lead to loss of information such as measurement frequency or interval which is likely crucial for many tasks.

A well-known time-series data visualization, the line graph, is effective in revealing overall temporal trend of a time-series; however, it is not accurate in showing the measurement frequency or interval. Moreover, in some sense, it harms the graphical integrity [29] because the connected lines lead to false confidence in values between measurements, especially for irregularly measured time-series data. In contrast, the bar graph is effective in showing frequency/interval without interpolating values between measurements, but it is less efficient in showing an overall pattern than the line graph.

In this paper, we propose a unified visualization model, called a ripple graph, that takes the benefits of both of the bar graph and line graph with enhanced graphical integrity for not only the regularly measured but also irregularly measured time-series data. The ripple graph also unveils uncertainty [26] of values between two temporal measurements by varying color intensity depending upon the confidence of the values. In doing so, it can effectively reveal the measurement frequency or interval while still showing the overall temporal pattern of change. We further extend the ripple graph representation into a single unified multi-scale visualization model via an interactive 2D widget to accommodate the advantages of other efficient time-series data visualization techniques while addressing the scalability issue.

Following a participatory design process with neurologists, we designed an interactive visual exploration tool for time-series data, called Stroscope, based on the ripple graph representation and the widget. We conducted a controlled user study to show the efficacy of the ripple graph by comparing it to existing representations for time-series data visualization. We also performed long-term case studies following the multi-dimensional in-depth long-term case studies protocol [28] to show the effectiveness of Stroscope in the real field with real users.

This paper is organized as follows. After describing closely related work, we describe the design rationale along with real world data and a user scenario. We explain the ripple graph and its user interaction, and then introduce the visualization tool—Stroscope. After we present the controlled user study results and the long term case study results, we close this paper with plans for future work and conclusions.

2 Related Work

In this section, we review previous work on representation techniques and interactive visualization tools for time-series data. We deal with temporal event sequence visualization separately. We also review previous work on evaluation of time-series data visualizations.

2.1 (Large) Time-Series Data Visualization

The goals of analyzing time-series are to grasp the evolution of data over time and detect trends and patterns for gaining insights and understanding data [1]. There are many interactive visualization tools to help users achieve the goals. The flexible multi-foci navigation techniques were proposed in KronoMiner [39] and SignalLens [16]. BinX [9] supports different aggregations on time dimension according to the abstraction level defined by a user. ChronoViz [11] is a visualization and analysis tool for time-based data from multiple sources. Visual exploration tools for patient data were proposed in [10], [13] and [24]. These tools are designed for handling a few long time-series, but not for showing an intuitive overview of multiple time-series.

Many representation techniques or interactive tools for large amounts of time-series have been developed. Based on the two-tone pseudo coloring by Saito et al. [25], a more space-efficient visualization technique called a horizon graph [22] was developed using dividing, mirroring, and layering techniques. However, our controlled user study showed that this technique based on filled line chart is not suitable for frequency-related tasks.

Line graph explorer [17] and LiveRAC [20] provide interactive interfaces for exploratory analysis as well as overview for multiple time-series using a Focus+Context technique. Thakur et al. [30] suggested a two-dimensional representation using a symmetric glyph, called a kite diagram, and presented data vases to compactly display multiple time-series. In [7], several representation techniques, each of which is efficient in revealing a different level-of-details were introduced using medical data as examples. All these techniques and visualization tools were designed assuming regularly measured time-series. Thus, it is difficult for them to show the measurement frequency or degree of irregularity in measurement intervals.

Aris et al. [5] suggested four representations for unevenly spaced time-series data. The *sampled events* method and the *aggregated sampled event* method generate an evenly spaced time-series data by sampling at a specific regular interval. The *event index* method distorts the time axis to highlight the number of events. The *interleaved event index* method represents the sequence of events while preserving temporal order of events regardless of their real time interval. However, the first two methods can cause the data loss that can come from sampling and aggregation and the last two methods can distort the time axis by arbitrarily changing the time intervals between two consecutive events. TimeRider [23] also deals with irregularly sampled data and reveals temporal aspects using animation in an animated scatter plot. In this tool, however, it is hard to see an overview because only one time frame can be seen at a time. Our tool supports an intuitive overview while maintaining the graphical integrity for irregularly sampled data.

2.2 Event Sequences Data Visualization

Temporal event sequences in the data such as electronic health records, highway incident logs or web logs, can be thought of as a kind of time-series, where each event does not have a quantitative property but a categorical one with a timestamp. The irregularly measured time-series data handled in this paper covers this event sequence data. There have been visualization tools to help users discover frequent or anomalous patterns in these temporal categorical event sequence data.

VISUAL-TimePAcTS [33] shows activities of individuals during a day using a space-time path. ActiviTree [34] visualizes a single event sequence using a circular tree-like representation and Continuum [3] visualizes a hierarchical relationship between temporal events. Lifelines2 provides visual temporal summaries to compare trends across multiple records [35] and an alignment operation that aligns sequences by a specific event [36]. Wongsuphasawat and Shneiderman [37] proposed a temporal categorical similarity measure, called a “Match & Mismatch”, to find similar records. LifeFlow [38] visualizes all possible patterns of event sequences through aggregation while maintaining temporal interval between events. CloudLines [18] represents each event as a circle with the size and opacity varied by importance of the event. These tools represent multiple event sequences by stacking and interpolating them

TABLE 1
Four Clinical Variables (Selected)

Variable	Value	Description
TOAST	1~8	Classification of types according to the causes of stroke
END_Progression	1 or 0	Occurrence of symptom that gets worsening within 24 hours or not
END_Recur	1 or 0	Occurrence of symptom that gets worsening after 24 hours or not
END_Sym_ht	1 or 0	Case in which hemorrhage is observed in image, which is related to the END

There are 29 clinical variables in the 1600 acute ischemic stroke patients' data that we used for the design and development of Stroscope. TOAST (Trial of Oragaranin Acute Stroke Treatment) stands for a classification according to the causes of stroke. END (Early Neurological Deterioration) stands for neurological worsening within hospitalization period from stroke onset.

vertically after aligning each sequence by a time attribute. We took these approaches, except for interpolation, in our tool to show an overview and compare trends across multiple time-series.

2.3 Evaluation

There have been evaluation studies on graphical perception of visual representations for time-series data to recommend appropriate representations for different types of tasks. Javed et al. conducted a controlled user study to evaluate four different visualization techniques (simple line graph, braided graph, small multiples and horizon graph) for local/global tasks in terms of the graphical perception of multiple time-series [15]. This study showed that shared-space techniques are more efficient for local comparison tasks and separate-space techniques are more efficient for dispersed comparison tasks. Aigner et al. showed that an indexing technique—transforming scale of data to a comparable unit—was superior among different representation techniques in comparing two heterogeneous time-series data through a comparative study [2]. A study by Heer et al. [14] compared the performance between line graph and horizon graph and showed that layered bands are more efficient for a small display space. Perin et al. [21] proposed interactive horizon graphs by adding zoom and pan interaction to horizon graph and showed improved performance though an evaluation with non-synthetic data set. These evaluation studies were concerned with the line graph and its variants.

Shneiderman et al. introduced "multi-dimensional in-depth long-term case studies (MILCs)" to evaluate information visualization tools in case studies with real users dealing with real data sets in their workplaces [28]. This evaluation method has been used in many studies [20], [27], and [38]. We also performed long-term case studies following the MILCs.

Arias-Hernandez introduced PairAnalytics [6], which is an approach that a *subject matter expert* (SME) and a *visual analytics expert* (VAE) perform a given task together for real data and problems. In our case studies, we also employed a modified pair analytics method where an experimenter (i.e., SME) demonstrated our tool to participants (i.e., VAE) only when they asked for help as used in [19]. In this way, participants tried using our tool and quickly became familiar. Then we could improve our tool iteratively by removing roadblocks.

3 Problem Analysis

One of the ultimate goals of medicine is to take care of the health and well-being of patients for their whole lifespan. To achieve this goal, it becomes necessary to keep track of individual health records throughout their entire lifespan, which makes it inevitable to deal with irregularly measured time-series data. Body temperature, blood sugar level, blood pressure level, and liver enzyme level are good examples. As an attempt to promote interactive information visualization techniques in the medical domain, a collaborative participatory project was launched with neurologists interested in analyzing relationship between the progression of stroke and the blood pressure change over time. In the following sections, we explain the clinical research problems with the blood pressure data.

3.1 Data Set

A group of neurologists collected time-series data of blood pressure measurements for 1,600 acute ischemic stroke patients at the Seoul National University, Bundang Hospital in Korea. All patients in this data set were hospitalized within 48 hours after the onset of stroke.

There are two sets of data. One includes stroke-related clinical information with 29 clinical variables such as age, gender, and medical history. Four important variables are summarized in Table 1. The other data set includes systolic blood pressure (SBP) and diastolic blood pressure (DBP) values along with the time of measurements. This data set is different from the usual time-series data handled in most conventional time-series data visualization tools. First, it is measured irregularly over time. Second, each patient has a unique hospitalization period, i.e., the total measurement period is different for each patient, ranging from 3 days to 60 days. Third, the first/last measurement time is different for each patient.

3.2 A Scenario—Status Quo

At the beginning of our participatory design process, we observed how neurologists analyzed blood pressure data in their clinical practice. As a result, we came up with a primary persona, Dr. Lee—a neurologist with 20 years of experience, and a representative user scenario that explained the status quo of the data analysis process in the field.

Dr. Lee's goal is to examine if there are differences in blood pressure value and variability between the patients whose symptoms worsen within 24 hours from the onset of stroke and others. He needs to focus on the effect of the END_progression to achieve his goal. He first separated the patients into two groups: patients with END_progression value of 1 (group A) and other patients (group B). He consulted a statistician and decided to control two most

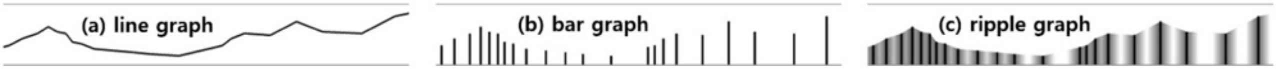

Fig. 1. Concept of ripple graph. Ripple graph (c) takes the benefits of line graph (a) and bar graph (b) with enhanced graphical integrity. It effectively shows the overall temporal trend and the measurement frequency/interval while revealing the uncertainty of values between real measurements.

important compounding variables—gender and age. They generated a new group B′ by choosing the ones in the group B that matched the ones in the group A in terms of gender and age. They used the statistical software R for the matching task, and then they compared the two groups of A and B′ in SPSS in terms of initial value, mean, standard deviation, maximum, and minimum of blood pressure. Although he wanted to see an overview of each group at once, descriptive statistics did not tell him an overall temporal trend in intuitive ways.

Carefully comparing the two groups, he formulated hypotheses. For example, he hypothesized that there would be a meaningful blood pressure change when an END_progression occurs. To verify this hypothesis, he needed to examine how the measurements fluctuate around a clinically meaningful value and how the mean value changes every three- or six-hour period during 24 hours before and after END_progression occurs. These analyses were arduous and time consuming with conventional statistical tools because the numerous iterative filtering of patients and the quantization of time intervals were not efficiently supported in such tools. A bigger problem was that every patient's data in group B′ was required to be aligned along a simulated event to be fairly compared with group A, where a simulated event could be defined as a virtual event occurring at the event time of the matched patient in group A.

3.3 Design Process

We learned that the neurologists have never seen their data in a visualization tool. The fact that they could "see" the data in a more intuitive and informative way and interactively manipulate the data highly motivated them to participate in the design process.

We as information visualization designers collaborated with the real users—three neurologists to understand each other's work. We had met them at the hospital six times over a 6-month period. We alternated between observing users while they performed data analyses with their conventional tools and discussing what they did and why. We also showed them what is possible with interactive visualization tools to educate them about information visualization and to strengthen the partnership.

3.4 Design Rationale

As a result of our observations of and discussions with the real users, we came up with the following design rationale of Stroscope using a new visual representation.

1. *Reveal measurement frequency/interval*. Blood pressure measurement frequency tells analysts a patient's condition, an occurrence of event, or a change in surroundings. We decided to design a new representation based on the bar graph which is useful to grasp the measurement frequency.
2. *Show fluctuation*. Neurologists observe changes in blood pressure around a clinically meaningful reference value possibly different for each patient, which can lead to finding patients with anomalous patterns. We decided to show the fluctuation of measurements around a user-defined clinically meaningful value.
3. *Use focus+context technique*. Systolic blood pressure values are concentrated in a narrow range around 120. It was required to enable users to interactively adjust their specific range of interest and see the detail while maintaining the context.
4. *Keep familiarity*. We integrated the two well-known time-series representations of the bar graph and line graph into a new visual representation to reduce the learning curve.
5. *Provide unified interface*. Different representations are necessary to support different tasks for exploratory data analysis. We designed a unified interface framework into which we can seamlessly integrate multiple representations.
6. *Support multiple time-series*. For visual exploration of multiple time-series, we had to support an intuitive overview and user interactions such as zooming, sorting, filtering, and grouping.
7. *Facilitate comparison task*. One of the main tasks in the analytical exploration of the blood pressure data is to compare between two groups. To reduce bias and the effect of compounding variables in the comparison task, selecting well-matched entities from each group is important. Interactive matching and alignment of patients across two groups are important for accurate comparison.
8. *Integrate visualization and statistical methods*. To facilitate exploratory data analysis, on-demand on-the-spot visualization of statistical summary measures is required. It enables users to perform a quick-and-dirty hypothesis testing on the spot.

4 Ripple Graph: A Multi-Scale Visualization Model for Time-Series Data

We propose a multi-scale time-series data visualization model, called a ripple graph, to represent measurement frequency and uncertainty between measurement points as well as measurements of time-series data. We integrate the line graph and bar graph into the ripple graph (Fig. 1c) to take the benefits of both of them, i.e., the line graph for showing the overall temporal trend (Fig. 1a) and the bar graph for revealing measurement frequency/interval (Fig. 1b). Furthermore, it also takes advantages of space-efficient representation techniques such as the horizon graph [22] and the heatmap-like graph [17] in a multi-scale model.

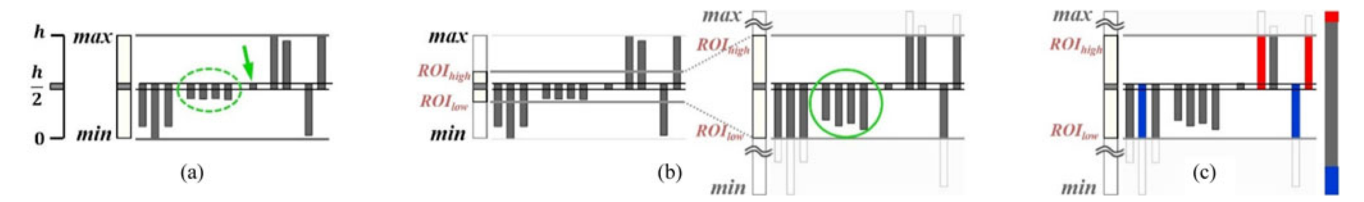

Fig. 2. Dimension zooming and color mapping in ripple graph. The ripple graph represents a time-series with a sequence of 12 values ranging from *min* to *max* for the given height h. (a) General ripple graph. (b) Ripple graph after applying *ROI* parameters: Bars are more distinguishable from each other than bars in (a). (c) Ripple graph after assigning colors to measurement values: A color strip on the right is a user-defined color table.

4.1 Visual Representation

The ripple graph aligns time points horizontally and measurements vertically (Fig. 2a). The horizontal axis runs vertically in the middle of the given space by default. For example, in case of the blood pressure data, the horizontal temporal axis is positioned vertically at the midpoint of the blood pressure value range, and each bar anchored to the horizontal axis is displayed upward for the values over the midpoint value and downward for the values under the midpoint value. In this way, it is easy to see how the blood pressure measurements fluctuate around a specific value which can be any user-defined meaningful value. The default position of the horizontal axis can be changed depending on the problem domain. For example, in case of stock data, the horizontal axis is better to be positioned at the bottom, i.e., at zero.

When one measurement is exactly equal to the user-defined value that the horizontal axis represents, the graph cannot show that measurement point because the height of its corresponding bar is zero, which gives a false impression to users that it was not even measured at the time point. To remedy this problem, we made the horizontal axis a tube-like dual line with the thickness of a small number of pixels. Then, the measurement point can be clearly shown as a blob in the tube as the eighth bar in Fig. 2a (indicated by the green arrow).

4.2 Multi-Scale Modeling

Quantitative comparison in a ripple graph becomes challenging when many graphs have to be shown on the screen. In this case, the height of a ripple graph becomes too small to discriminate each value in the data. In order to resolve this problem, we developed three space-efficient methods in a multi-scale model.

*4.2.1 Dimension Zooming with Range of Interest (*ROI*)*

The first method allows a user to zoom in to a specific range of values that the user is interested in by adjusting the range of the vertical axis. It adjusts the height of a bar accordingly to the range of measurements of interest defined by the user, where the vertical axis spans from the lowest value (ROI_{low}) to the highest value (ROI_{high}) in the range. Then, the bars for any values out of the range have the same height, i.e., reach the top or the bottom of the given space. In this way, the user can zoom in to a specific range of interest to compare values in the range in detail, while maintaining the context, i.e., knowing the existence of the values outside the range. The four downward bars with similar height in Fig. 2a (see the bars within the green dotted circle) can be more clearly distinguishable after setting ROI_{low} and ROI_{high} properly as shown in Fig. 2b (see the bars within the green solid circle).

4.2.2 Color Mapping to Further Distinguish Bars

The second method allows a user to assign colors to measurements, through which each bar is painted in the corresponding color determined by a user-defined color table. This method was similarly used in [7] as "height-coded timelines," but the two methods are different in that our method fills only the bars whereas the height-coded timelines fill the whole space including gaps between bars as well. Fig. 2c shows that the first two bars with the same height are differentiated by color (gray for first bar and blue for second bar) and so are the third and fourth bars to the right end. The user-defined color mapping is shown in the vertical color strip on the right side of Fig. 2c. To give users more flexible control over the visual encoding, the color mapping is independent of ROI.

4.2.3 Moving the Horizontal Axis

The last method allows a user to change the vertical location of the horizontal time axis. To see the bars with the values over a specific value in detail, a user can move the horizontal axis downwards. The horizontal axis can be even located at the bottom of the given space representing the value of ROI_{low}, and then the bars with the values less than ROI_{low} disappear. Fig. 3a shows that the first three bars, which have values less than ROI_{low}, are filtered out. In this way, the vertically movable horizontal axis enables users to filter out some measurements.

When both ROI_{low} and ROI_{high} are set to the minimum in the whole measurements range and the horizontal time axis is located at the bottom (Fig. 3b), all bars have the same height, i.e., the height of the given space for each series. The visualization then becomes a heatmap-like graph [17] where each measurement is represented by a vertical strip with a specific color assigned by users. Since users can reduce the

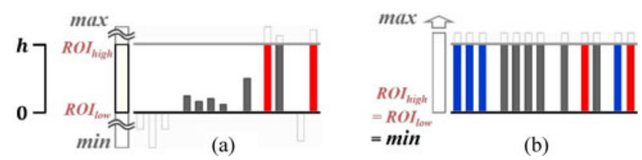

Fig. 3. Moving the horizontal axis. The ripple graph represents the time-series used in Fig. 2. The horizontal axis is located at the bottom. (a) Bars with the values less than ROI_{low} are filtered out. (b) All bars have the same height when ROI_{low} and ROI_{high} are set to the minimum of the measurement value.

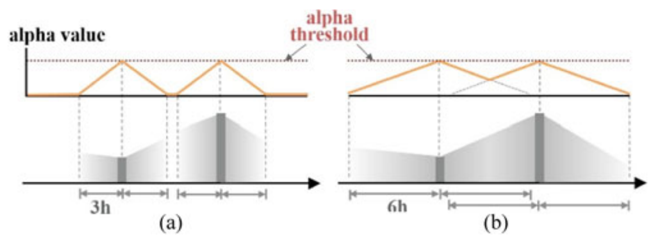

Fig. 4. Ripple graph with confidence Interval of (a) 3-hour and (b) 6-hour. The ripple graph fills the area between bars while changing the alpha value.

given height per series down to 1 pixel, this method provides one of more scalable representations regarding the number of series shown at once.

4.3 Visualizing Degree of Certainty between Measurements

While the ripple graph emphasizes individual measurements, it can also show the overall pattern of change over time. Although the blood pressure changes continuously, it is measured discretely. A linear interpolation is conventionally employed as a simple way to fill the gap between discrete measurements. However, we note that the degree of certainty of an interpolated value between measurements decreases proportionally to the distance from a real measurement point. To the best of our knowledge, there has been no attempt to show the degree of certainty information in time-series data visualizations, which we believe contributes to enhancing the graphical integrity. The ripple graph fills the area between bars with the color of the closest bar while changing the alpha-channel value, where the color becomes more transparent as it gets farther from the closest real measurement point. Similar methods were used for visualizing hierarchical structure in a treemap [12] and density of a cluster in parallel coordinates [31].

Let x be the distance between a real measurement point and a specific point between measurements. The degree of certainty (*DOC*) at the in-between point is determined as follows:

$$DOC(x) = -\frac{x}{c} + 1,$$

where c is a confidence interval, defined by users, which represents the temporal range over which a real measurement holds its confidence. The alpha value at the in-between point is the product of $DOC(x)$ and a maximum alpha threshold (≤ 1). The maximum alpha threshold is empirically set to 0.8. Figs. 4a and 4b show ripple graphs when the confidence interval is 3-hour and 6-hour, respectively. The upper graph in the figure shows how the alpha value changes over time.

The ripple graph has the following advantages with the degree of certainty representation: (1) it enhances the graphical integrity by showing predicted (or interpolated) values along with important context, i.e., the confidence of the predicted values; (2) the variability of measurement values along with measurement frequency is shown more clearly; (3) and it can effectively approximate the real continuous change over time for irregularly measured time-series data.

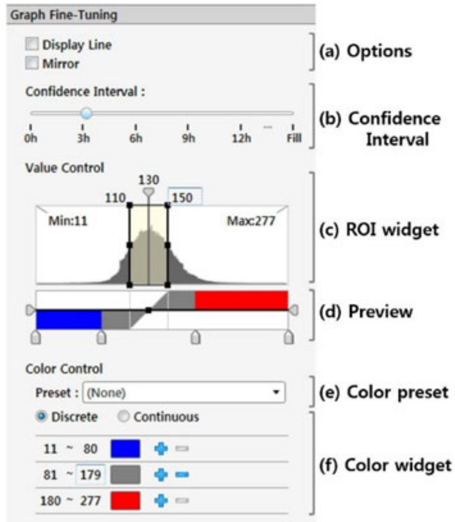

Fig. 5. Control panel for ripple graph manipulation. Controls for (a) connecting bars and flipping bars downward around the horizontal time axis, (b) adjusting confidence interval, (c) adjusting ROI values, (d) showing a preview and changing a position of the time axis, (e) selecting a predefined color palette, and (f) manipulating color-related parameters.

4.4 User Interface for Ripple Graph Manipulation

We designed an interface to enable users to flexibly adjust the parameters for a ripple graph (Fig. 5). To adjust the range of the vertical axis for a ripple graph, user can drag two draggable vertical edges of the rectangle-shaped range slider on the ROI-widget histogram (Fig. 5c) to set the range from ROI_{low} to ROI_{high} (e.g., 110 to 150 in Fig. 5c). Users can also directly enter the desired value in an edit box on the top of each vertical line. A grey vertical line between two draggable vertical edges indicates the vertical location of the horizontal time axis. When ROI_{low} or ROI_{high} is changed, the grey vertical line moves to the midpoint between ROI_{low} and ROI_{high}.

A preview (Fig. 5d) below the ROI widget shows the height and color of a bar depending on its value ranging from minimum to maximum. A horizontal line on the preview indicates the time axis. To examine the fluctuation of measurements around a specific meaningful value, a user can move the horizontal axis by dragging the line up and down to change the location of the axis between ROI_{low} and ROI_{high}. Then, the grey vertical line within the ROI widget (e.g., 130 in Fig. 5c) also moves accordingly. Any changes of ROI values and the horizontal axis position cause immediate updates in the preview, the ROI widget, and the timeline view.

5 Stroscope

Our 6 month-long participatory design with the neurologists leads us to implementing an interactive visualization tool for time-series data, entitled "Stroscope," where the ripple graph is the main visualization component. In this section, we explain the user interface and interaction models of Stroscope along with related analytical features.

5.1 Layout

Stroscope consists of three main areas (Fig. 6): Control panel, timeline view, and detail view. The control panel in the left area has four tabs: (1) the control tab for sorting,

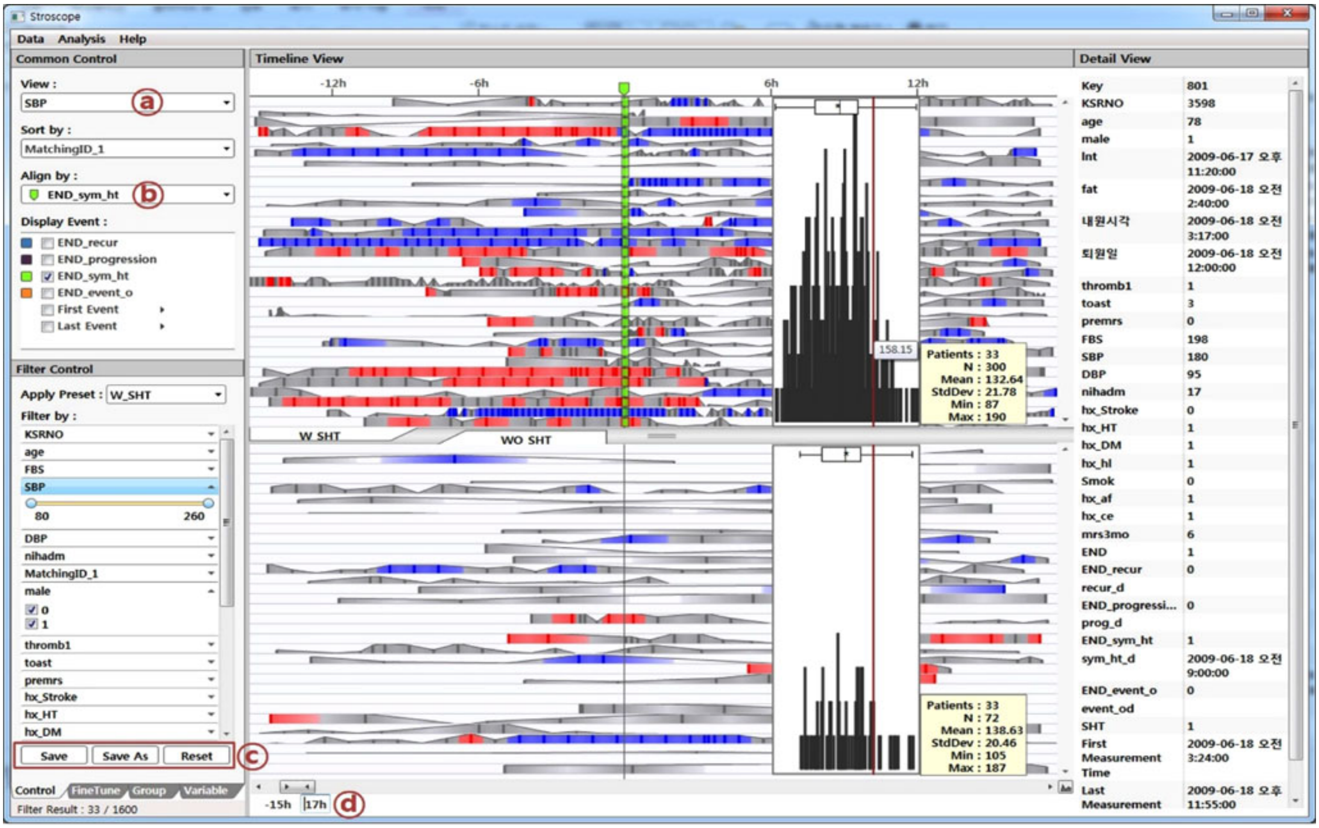

Fig. 6. Overview of Stroscope. Stroscope interface consists of three main areas: control panel (left), timeline view (center), and detail view (right). Stroscope enables users to compare two groups at horizontally splitted views using matching, sorting, and aligning functions. In upper timeline view, records with END_sym_ht were aligned by occurrence time of the event, where a light green rectangle of each row indicates the corresponding event. Users can also select a region of interest by drawing a rubber-band rectangle over the timeline view to check statistical summary measures within the region (gray rectangle). Stroscope provides effective user controls for (a) selecting a measurement variable, (b) selecting a variable for alignment, (c) making a filter present, and (d) specifying a temporal period of interest.

filtering, and aligning time-series, (2) the fine-tuning tab for adjusting various parameters of the ripple graph, (3) the grouping and layout tab for changing the view layout and dividing records into groups, and (4) the variable tab for creating a custom variable out of existing variables.

The timeline view in the middle area visualizes multiple time-series using the ripple graph, where each series is vertically stacked with the same height. All series are aligned by the first measurement time by default. The alignment indicator (i.e., the green marks in the upper timeline view and the vertical line in the lower timeline view in Fig. 6) highlights the alignment time, where the relative time is zero.

The detail view in the right area shows the information of the selected series in the timeline view, and this information is immediately updated upon any selection.

5.2 User Interaction

Stroscope provides effective means to help users efficiently explore a large time-series data. They include common operations that many existing time-series visualization tools support, e.g., ARF framework in [36] for alignment, ranking, and filtering.

5.2.1 Select a Measurement Variable

Time-series data can have more than two contemporaneous variables changing over time, i.e., multivariate time-series. In blood pressure data, there are two such variables: SBP and DBP. Furthermore, Stroscope automatically generates a delta measurement variable that is defined as a sequence of differences between consecutive measurements (ΔSBP and ΔDBP) to help users easily examine variability. Users can select one of these measurement variables using a combo box (Fig. 6a).

5.2.2 Align

An alignment operation enables users to effectively compare temporal patterns before and after a specific event across multiple records. It helps users to predict prognosis and determine a treatment for a patient. Stroscope allows a user to align all records by a clinical variable (such as onset time and hospital arrival time) or by a specific event (such as END_progression). Users can select one of these variables for alignment in a combo box (Fig. 6b).

The alignment indicator at the top of the timeline view is positioned at the top left end by default and filled with the corresponding color to the selected alignment variable. Users can drag the indicator left or right within the entire time range to check the trend before and after the alignment point in detail.

5.2.3 Zoom

Our multi-scale visualization model, the ripple graph allows users to dynamically change the visual representation depending upon the available display space. To maximize space utilization, Stroscope provides a vertical zooming to adjust the height of the ripple graph for each record through wheeling the mouse on the timeline view. A zooming on the horizontal axis is also provided to help users focus on a temporal period of interest (Fig. 6d).

5.2.4 Hierarchical Grouping

To detect differences among groups classified by categorical variables (e.g., gender or age group) is one of the routine analytic tasks for neurologists. Stroscope facilitates this task by enabling hierarchical groupings. The result of a grouping is summarized in an enhanced treeview control (Fig. 13) where each node has both the number of corresponding records and a check box to show/hide the corresponding group. The control panel in Fig. 13 shows the result of a hierarchical grouping where all records are divided into groups by "Cluster" first and then by "END" categorical variable.

5.2.5 Side-by-Side Comparison

One of the main tasks in the analytical exploration of the blood pressure data is to compare between groups. To facilitate such comparison task, a horizontal/vertical juxtaposition of views for the two groups is necessary. A user can define new groups by making filter presets (Fig. 6c) and apply a different preset to each view after splitting the timeline view horizontally or vertically into two. The zoom factor and the position of the alignment indicator are synchronized between two views. Fig. 6 shows two horizontally split views.

5.3 Analytical Features

5.3.1 Statistical Summary Measures on Demand

Providing statistical summary measures on users' demand is important for gaining insights and performing a quick-and-dirty hypothesis testing in an interactive visualization tool. In Stroscope, a user can designate a region of interest by using a rubber-band (rectangle) selection on the timeline view, and check the descriptive statistical summary measures of the region. The summary measures include the number of records, number of measurements, average, standard deviation, minimum value, and maximum value. Stroscope provides the summary measures not only numerically, but also graphically in a histogram and a box plot together within the selection rectangle (gray rectangle on timeline view in Fig. 6).

Using the side-by-side comparison feature, users can instantly compare two groups in terms of the statistical summary measures. The linking and brushing technique implemented in Stroscope enables a user to select a rectangular region in one timeline view and see the same region selected in the other timeline view. This feature could encourage users to interactively perform their routine analytic task without cognitive overload.

5.3.2 Accommodating Individual Differences

In the blood pressure data, there exist individual differences among patients in terms of age, weight, medical history, and so on. A value of 150 is generally considered a slightly higher value, but the value can be a critical value for a patient with relatively lower blood pressure. However, for all patients, the value of 150 was represented by the same height and color in the ripple graph. To resolve this issue, Stroscope allows a user to define a reference variable with a different basis value for each patient. For example, let us consider that a user defines the average blood pressure during three days before discharge as a reference variable with "SBPMean" name. Then Stroscope calculates the basis value for every patient according to the definition of the reference variable, and generates a new measurement variable that is defined as the difference between the real blood pressure value and the basis value for each patient. Consequently, a user can accommodate individual differences among patients using the reference variable.

5.3.3 Matching

Users are commonly interested in identifying factors that may contribute to a clinical outcome such as recurrence of END by comparing patients who have the condition (case) with patients who do not have the outcome (control) but are otherwise similar [8]. To help users conduct this kind of case-control study, Stroscope supports the matching of the case to the control. For example, users can match 'patients with the END_recur event' to 'patients without the END_recur event' according to various variables (e.g., gender should be exactly matched and the age should not be different by more than 5). Then, Stroscope generates a new variable with a user-defined name, where one record in the case and one or more matched record(s) in the control have the same value for the variable.

5.3.4 Data-Space Clustering and Image-Space Clustering

Users can gain insights by finding any regularities or anomalies through clustering. We used the I-kMeans algorithm [32], an interactive k-Means clustering method taking advantage of the multi-resolution property of wavelets. We enhanced this algorithm by making it applicable to irregularly measured time-series data because the algorithm assumes that each time-series data is measured regularly.

Stroscope provides two kinds of clustering techniques: data-space clustering and image-space clustering. For the data-space clustering, records that have similar measurement values are grouped together, which always results in the same clusters for the same data set. However, a coarse color mapping could result in a visual inconsistency problem that the resulting visualization does not look well-clustered. For example, a neurologist assigned any measurements of 180 or greater to a red color because those are considered as critical values. However, data-space clustering could separate two measurements in the same red, e.g., 180 and 240 because they were numerically very different in the data space. This result could confuse users due to the inconsistency between users' mental model expressed in color mapping and clustering results based on actual data values.

The neurologist suggested an image-space clustering during the participatory design process in order to resolve this visual inconsistency problem of the data-space clustering. In the image-space clustering, records with a similar color pattern are clustered together, where the clustering results could vary according to the color table defined by the user, but the results make more sense to the user who expresses his intention in his color mapping choice.

5.4 Implementation

Stroscope is implemented in C# with windows presentation foundation (WPF). Stroscope can display more than two different time-series data sets in a multi-monitor environment. It is possible to individualize or synchronize parameters of the ripple graph among different monitors. In a 24″ 1920×1080 resolution monitor, Stroscope can display up to about 950 (1,900 in the side-by-side comparison mode) time-series interactively on screen at once.

6 Evaluation

We conducted a controlled user study to evaluate the ripple graph in terms of how well people could learn and use it in performing graphical perception tasks, compared to other visualization techniques for time-series data. To show the effectiveness of Stroscope in the real field, we also conducted case studies with two neurologists following the evaluation frameworks, multi-dimensional in-depth long-term case studies (MILCs) [28] and PairAnalytics [6]. In this section, we summarize the controlled user study results and the two studies that we conducted with real users in the field.

6.1 Controlled User Study

6.1.1 Participants and Materials

We recruited 14 (five females) volunteers from Seoul National University for the study. The average age of participants was 28.4, ranging from 23 to 35 years of age. Five of them majored in computer science and engineering, and the others were from four different majors, i.e., chemical and biological engineering, pharmacy, economics, and communication studies. All participants were familiar with the line graph and bar graph. The experiment took about 40 minutes and they were rewarded about $5.

We prepared a time-series data set of blood pressure measurements for 1,600 stroke patients. We only used the measurements during the first three days after hospitalization to keep participants from distinguishing records by their different measurement periods.

6.1.2 Visualization Techniques

We compared the following four visualization techniques for time-series data. The first three were among the most representative techniques.

- *Line graph (LG).* We actually used a filled line graph with gray color to ease identification [15].
- *Bar graph (BG).* Each bar with three pixels width was filled with gray color.
- *Interactive horizon graph (IHG).* We implemented the horizon graph with zoom and pan interaction introduced in [21].
- *Ripple graph (RG).* For a fair comparison with other three techniques, participants were only allowed to change the *ROI* values and move the horizontal axis from a control interface for the ripple graph.

We used a split-space technique where each time-series is shown in a row of the same height [15]. We fixed the height for each time-series at 24 pixels to facilitate comparison between our results and previous studies in [14] and [21].

6.1.3 Tasks

Based on the task model suggested by Andrienko and Andrienko [4] and user studies on the graphical perception of multiple time-series [14], [21], we chose the following four types of tasks.

- *Max:* Selecting a time-series with the highest value across all records.
- *Same:* Selecting a time-series which is exactly the same as a given record.
- *Frequency:* Selecting the most frequently measured time-series.
- *Confidence:* Determining the subjective confidence in the value at a given time point. The same number of time points was selected from real measurement points and those inbetween. The value for a point between two adjacent measurement points was linearly interpolated. Because BG represents only real measurement points, it was excluded from the confidence task.

6.1.4 Study Design and Procedure

We ran the study as a within-subjects design, with each participant performing all the tasks using all the visualization techniques. We ran the experiment as a 4 (Visualization technique: LG, BG, IHG, and RG) $\times 4$ (Task type: max, same, frequency, and confidence) $\times 2$ (Number of time-series: 20 and 40) $\times 2$ (trials) deign while counterbalancing the order of visualization techniques. Performance time and correctness of answers were the dependent variables of this study. To avoid a learning effect, we randomly selected a small number (20 or 40) of time-series from the pool of 1,600 time-series for each trial while maintaining comparable complexity across trials.

Before beginning real tasks, we gave participants a tutorial on a visualization technique and showed them how to perform four types of tasks with an example. They also had enough time to try out each technique by themselves. Then, they performed $12(3 \times 2 \times 2)$ tasks for BG or $16(4 \times 2 \times 2)$ tasks for others. They were asked to finish tasks as fast and precisely as they could. We measured the task time and correctness. At the end of a study session, participants filled in a questionnaire for subjective evaluation of each visualization technique.

6.1.5 Hypotheses

We established three hypotheses for this study.

1. For the max and same tasks, IHG and RG will outperform LG and BG in both the task time and correctness.
2. For the frequency task, BG and RG will outperform LG and IHG in both the task time and correctness.

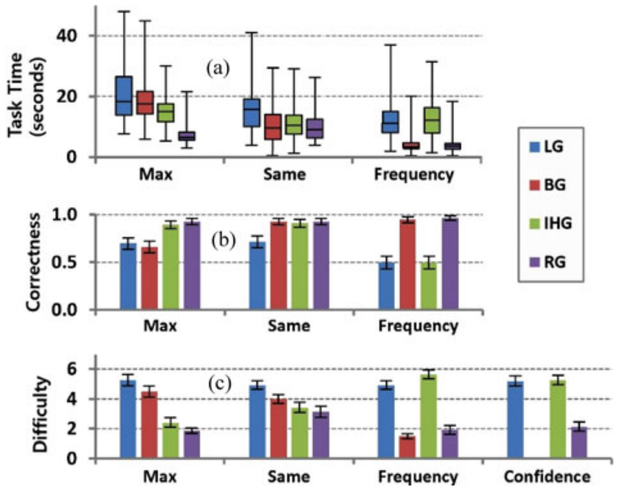

Fig. 7. (a) Task time, (b) Correctness, and (c) Difficulty for each visualization technique by task type.

3. For the confidence task, RG will outperform LG and IHG in distinguishing between real measurement values and interpolated values.

6.1.6 Statistical Analysis and Results

We analyzed the task time and correctness with a 4 (Visualization technique) $\times 3$ (Task type excluding Confidence) $\times 2$ (Number of time-series) repeated-measures analysis of variance (RM-ANOVA). Regarding the task time, we found a significant main effect of visualization technique ($F_{3,312} = 3.53, p < .05$). Figs. 7a and 7b show the mean task time and correctness for each technique by task type, respectively. To analyze differences among the visualization techniques, we also conducted Tukey's HSD post-hoc test for each task type (Table 2).

The results supported our first and second hypotheses except for the same task with BG. Participants completed the same task in a significantly less time and with significantly more correct answers with BG than with LG. The reason might be that BG enabled participants to quickly filter the target time-series by preattentively perceiving the measurement frequency of time-series.

Participants usually spent more time in completing tasks with 40 time-series than with 20 time-series as shown in Fig. 8. However, there were exceptions where participants showed similar performance for both cases. The likely reasons for the exceptions are as follows.

1. For the max task with RG, participants could identify the target time-series at once just by narrowing the range of interest of values down to maximum value.
2. For the same task, participants could identify the target time-series at once by preattentively perceiving a different color of a band with IHG and by grasping the distinctive bar occurrence frequency with BG and RG. In contrast, with LG, they had to use their elementary perceptual skills to find the target.
3. For the frequency task, participants could grasp the frequency at a glance with BG and RG by just perceiving the overall occurrence pattern of bars.

TABLE 2
Results of Post-Hoc Test by Each Task Type

Records	Task	Pairwise comparisons
Task completion time	Max	**RG** < IHG < LG & BG
	Same	**RG** & BG & IHG < LG
	Frequency	**RG** & BG < LG & IHG
Correctness	Max	LG & BG < IHG & **RG**
	Same	LG < BG & IHG & **RG**
	Frequency	LG & IHG < BG & **RG**
Difficulty	Max	**RG** & IHG < LG & BG
	Same	**RG** & IHG < LG
	Frequency	**RG** & BG < LG & IHG
	Confidence	**RG** < LG & IHG

The < sign represents the inequality relation with a statistical difference (p < .05).

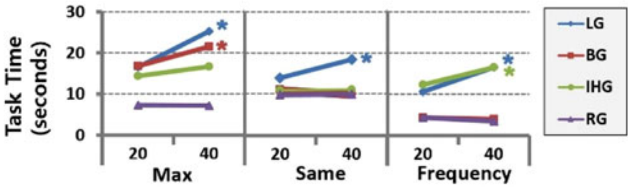

Fig. 8. Task time for each visualization technique by number of time-series (i.e., 20 and 40). * indicates a significant difference in task time between 20 and 40 time-series ($p < 0.05$).

For the confidence task, we analyzed the results after dividing selected time points into three categories: (1) real measurement points with extreme values, (2) real measurement points with non-extreme values, and (3) interpolated points between two adjacent real measurement points (Fig. 9). For real measurement points, the confidences with LG and IHG were lower than those with RG. Especially, for real measurement points with non-extreme values, the confidences of LG and IHG were as low as those for interpolated points. It is likely because connected lines or filled areas in LG and IHG made it difficult to tell if such measurement points are real or not. Participants also answered that they actually felt a difficulty in performing the confidence task with LG and IHG (Fig. 7c). In contrast, participants easily performed the confidence task with more reasonable rating of the confidence values with RG (Figs. 7c and 9).

6.1.7 Subjective Data

We asked each participant to rate how difficult each task was by using a 7 point Likert scale [Rating: 1 = Very easy; 7 = Very difficult] (Fig. 7c). We analyzed the subjective rating results using Friedman test with Bonferroni correction. We found a significant main effect of Visualization technique ($\mathcal{X}^2(3) = 53.2, p < 0.001$). Pairwise comparisons among visualization techniques for each task show that

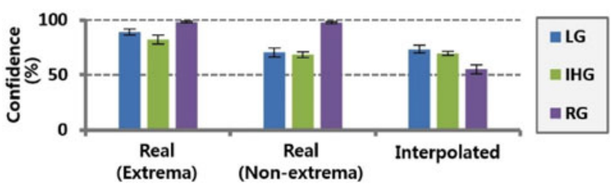

Fig. 9. Average of subjective confidence results for each visualization technique by time point type.

participants had less or same difficulty in performing a task with RG than others across all task types (Table 2).

6.1.8 Discussion

Our controlled user study results supported the three hypotheses, suggesting that the ripple graph was overall the best performing in terms of task time, correctness, and subjective satisfaction. Our results also suggested that the interactive horizon graph was not an appropriate technique for measurement frequency-related analysis of irregularly measured time-series data even though it was known to be good for quantitative comparison of multiple time-series [21]. Although there was no difference in performance for the Frequency task between BG and RG, participants commented that the bar graph caused confusion and eye fatigue when multiple time-series were displayed.

Participants commented that it was helpful to see the degree of certainty between real measurements. They could see the variability clearly in the ripple graph because visualizing the degree of certainty enabled them to perceive a time-series with many individual bars as a gestalt (i.e., unified whole) instead of a series of scattered bars.

6.2 Case Study

6.2.1 Data Set and Procedure

We designed and developed Stroscope involving three neurologists in the design process. After completing an initial prototype, we have conducted two case studies with neurologists for four months in the real field. We had a meeting with two participants (female and male doctors enrolled in a stroke fellowship program) together for about 90 minutes every 2 or 3 weeks for the first two months for exchanging ideas and sharing findings while improving the prototype as well if necessary. We had a 1-hour meeting with each participant every 2 weeks for the rest two months. We used the following procedure for each meeting: (1) Participants gave us feedback on Stroscope. (2) We installed an improved version and explained the improvements. (3) We let participants try Stroscope to confirm whether they understood the new features. In this stage, we employed a modified pair analytics method, where assistance is provided only when participants asked for help. (4) After the meeting, we updated Stroscope following the feedback and maintained contact with the participants by answering their questions via e-mails or phone calls.

When the two participants first tried Stroscope, they were impressed by its visual and interactive nature because they had never used such a visualization tool before. They were also excited that they could find significant patterns in a specific group by comparing different groups using matching, aligning, and clustering.

6.2.2 Participant1 (P1)

P1 was interested in the relationships between the variability in blood pressure and the occurrence of symptomatic hemorrhagic transformation (SHT) of acute ischemic stroke. SHT is one of the important factors that influence the outcome of stroke treatment. Previous studies have shown that the occurrence of SHT relates to high variability in blood pressure. But statistical summary measures did not bring him an intuitive understanding and it was always elusive to examine each record separately in detail using conventional statistical tools.

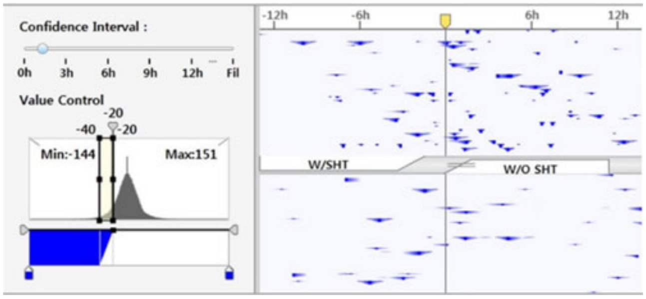

Fig. 10. Stroscope showing only SBP decreasing periods. After selecting △SBP (a difference between consecutive SBP values) as a measurement variable, P1 can observe △SBP values less than –20. The yellow mark indicates when an SHT onset event occurred.

P1 decided to conduct a case-control study with Stroscope. He first defined two groups: cases are the patients who have had a SHT and controls are the patients who have not had a SHT. He used the matching function of Stroscope to match each patient in the cases to the patients in the controls according to the initial SBP ($\pm$ 5 mmHg) and age ($\pm$ 5 years). After he split the timeline view into two (up and down), he assigned the cases to the upper view and the controls to the lower view. He immediately noticed that the blood pressure was measured more frequently and the hospitalization period was longer for the patients in the cases. For a detailed analysis, he aligned patients by the SHT onset time, while aligning patients in the controls by the SHT onset time of the patient matched in the cases. High or low values are observed more frequently near the SHT event in the cases than in the controls.

To analyze the variability of SBP, he first selected ΔSBP as a measurement variable and manipulated the color palette to make positive values red and negative values blue. It was difficult to see the difference of variability between cases and controls by just checking the occurrence of dark blue or dark red regions. So, he dragged the horizontal axis of the ripple graph to the bottom to see only the increasing periods with positive ΔSBP values. In the same manner, he dragged the horizontal axis to the top to see only the decreasing periods with negative ΔSBPs (Fig. 10). Then he could clearly see the difference between cases and controls, e.g., a rapid change in the blood pressure was observed more frequently in the cases near the occurrence of SHT event. During 6 hours before and after the SHT onset event in Fig. 10, blue regions representing the periods with decreasing SBP by more than 20 appear more frequently in the cases (upper view) than in the controls (lower view).

To minimize the influence of individual differences among patients, he decided to examine the deviation of SBP values. He first defined a reference variable named "SBPMean" as the average SBP during 3 days before discharge. And then, he changed the measurement variable to a new reference variable defined by SBP-SBPMean. He easily found that there were many extreme values, especially higher values in dark red in the cases, before the occurrence

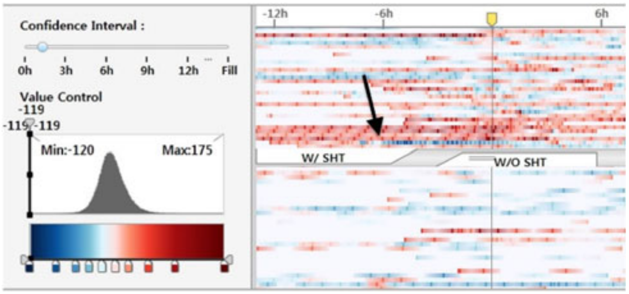

Fig. 11. Stroscope showing a “SBP-SBPMean” variable. A black arrow indicates a patient with an anomalous pattern. Dark blue represents a sudden decrease in SBP (a great negative value of “SBP-SBPMean”). The yellow mark is the indicator of SHT onset event.

of SHT (Fig. 11). In addition, dark red and dark blue colors were observed more frequently in the cases, indicating that the variability of SBP was high.

While observing the patients in the cases, he found one patient with an anomalous pattern: the patient exhibited a dramatic decrease of the “SBP-SBPMean” value about seven hours before the occurrence of SHT (see the black arrow in Fig. 11). After zooming in to the patient further, he found that the “SBP-SBPMean” value decreased approximately from 25 to –50 (see the black arrow in Fig. 12). He wondered why the SBP value suddenly decreased. He checked the medical history of the patient in the EMR (electronic medical record) system. He found that the patient received a treatment called mechanical thrombectomy for an occlusion in the sphenoid segment of the middle cerebral artery (a part of blood vessel in the brain). He reconfirmed the sudden decline in the SBP value though two MRI scans before and after the treatment to remove the offending thrombus.

6.2.3 *Participant2 (P2)*

In P2’s clinical research, she often found that patients’ blood pressure increased or decreased rapidly when they got worse. But there are previous studies that have shown conflicting results because most of these studies were cross-sectional which compared only statistical summary measures between groups without taking into account the temporal aspect of blood pressure change. P2 wanted to go beyond the statistical summary measures by visually exploring individual blood pressure values and their changes over time using Stroscope.

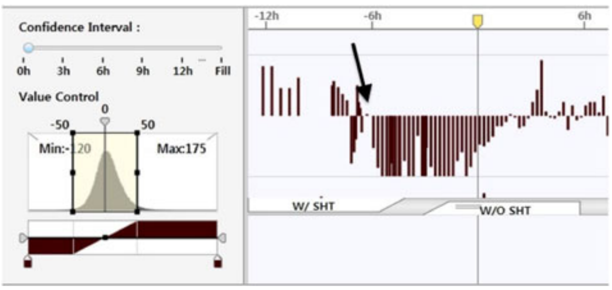

Fig. 12. Stroscope showing the detail of the patient with anomalous pattern in Fig. 11. P1 identified that the “SBP-SBPMean” value decreased approximately from 25 to –50 about 7 hours before the SHT onset. The yellow mark is the indicator of SHT onset event.

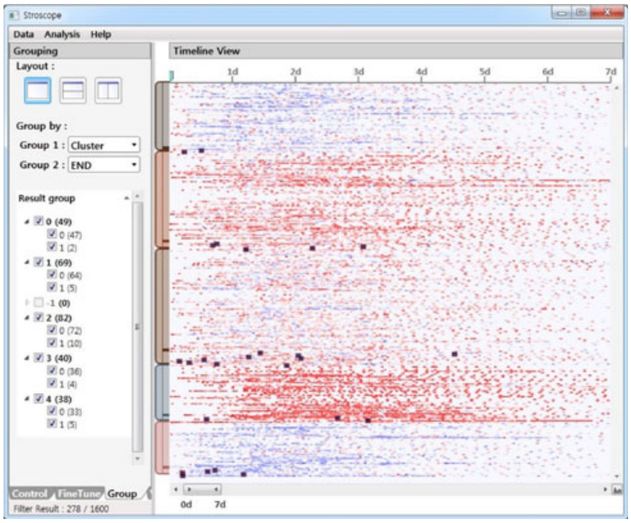

Fig. 13. Stroscope showing hierarchical grouping result. P2 aligned patients by onset time and grouped patients according to *Cluster* and *END* variables after performing a clustering function. Black rectangular spots in the timeline view represent the END_progression event. In the two clusters with low blood pressure (the first and last clusters), the event tends to occur within about 30 hours from onset.

As soon as she loaded her data set and sorted patients by TOAST (a classification stroke according to the causes of stroke), she started to make her discoveries in her data set: (1) Overall, SBP value was decreasing during a day after the first measurement; (2) patients in ‘TOAST 2’ and ‘TOAST 6’ groups were hospitalized for a shorter period of time; and (3) SBP values of patients in ‘TOAST 3’ and ‘TOAST 4’ groups tended to be relatively low.

Using the matching and alignment functions in Stroscope, P2 also figured out that the blood pressure of patients with lacunar infarcts in ‘TOAST 2’ was increasing or decreasing rapidly before and after an END_progression event. Then, she partitioned patients with lacunar infarcts into five clusters by performing a clustering using our enhanced I-kMeans algorithm (Section 5.3): two clusters with high blood pressure, two clusters with low blood pressure, and a cluster with slightly high blood pressure. She aligned records by the onset time to check whether there exists any difference among the clusters in the elapsed time from onset to the END_progression event. She observed that the END_progression event occurred within about 30 hours from onset in the clusters with low blood pressure (Fig. 13). But, the number of patients with the event was not enough to conclude that her observation was meaningful. So she decided to explore in the same manner after collecting more time-series for patients with lacunar infarcts.

6.2.4 *Discussion*

Two long-term case studies helped us test the efficacy and utility of Stroscope. Although Stroscope was the two participants’ first visualization tool for time-series data, they became rapidly proficient in using it. We allowed them to ask us for help whenever they were faced with any problems. But, they used Stroscope skillfully without any help after the first one month of the case study.

They used sorting, aligning, and matching functions for comparison of two groups, which was one of the main tasks in the analytical exploration of the blood pressure data. They could easily find differences in measurement frequency and variability as well as measurements, especially before and after a specific event. These findings confirmed what they already knew and also yielded the results contradicting previous studies.

They changed the visual representation from a familiar graph such as the line graph or bar graph to their own ripple graph depending upon their goals and the available display space. To see an intuitive overview of multiple records, they often made all bars have the same height of 1 pixel and then adjusted the color palette and confidence interval (Fig. 13). For participant 1, to see only the increasing/decreasing periods, he adjusted the position of horizontal axis and ROI values. He was satisfied with that he created his own graph to show the peak only. He could also observe one record in detail by adjusting ROI values after increasing the height of the ripple graph (Fig. 12). He commented on our visualization model and interactive widget as follows: "It is very nice that I can progressively narrow down to a range of values of my interest after understanding the context."

Although our multi-scale visualization model enables users to choose the best representation for a given display space, the scalability issue still remains. It can only scale up to a point where each time-series takes a pixel height. It is possible to scale up further by employing aggregation or data reduction techniques, but then we may lead to a more aggregated overview, thus inducing information loss. Such information loss is in general unacceptable in the medical field since it could complicate or mislead medical decision-making.

Another limitation is that we conducted two case studies in only one domain, i.e., medical domain with a blood pressure data. Further case studies are required to show that Stroscope based on the ripple graph is not a domain-specific tool. Thus, more case studies in other domains can be meaningful future work.

7 Conclusion and Future Work

In this paper, we presented a multi-scale visualization model, i.e., a ripple graph for irregularly measured time-series data, concerned with measurement frequency and confidence in values between measurements. To investigate the efficacy and potential of the ripple graph, we implemented an interactive visualization tool, Stroscope in which we provided an interactive widget to enable intuitive control of the ripple graph and several analytical functions. We then evaluated the ripple graph and Stroscope by conducting a controlled user study and two long-term case studies with neurologists. Results showed a promising possibility that our ripple graph is generally applicable visualization model for time-series. Case study participants could efficiently exploit the visualization model and the analytical functions of Stroscope throughout their exploratory analysis processes.

While the ripple graph focused on quantitative values only at discrete time points, more work is needed to generalize Stroscope to deal with a time axis not only as time points but also as time intervals (e.g., a representation like Gantt charts). We are also planning to adopt a different method to accommodate individual differences among patients in the analysis of temporal change rate. For example, conditional variance used in stock data analysis can be used as an alternative variable. Furthermore, to make our multi-resolution clustering technique more generalizable, it is necessary to adopt a different wavelet transform, e.g., lifting scheme, which makes it possible to do a discrete wavelet transform without regularizing irregularly measured time-series.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean Government (MSIP) (No. 2011-0030813, 2010-0002266 and NRF-2013R1A1A3006706). The ICT at Seoul National University provided research facilities for this study. Jinwook Seo is the corresponding author.

References

[1] W. Aigner, S. Miksch, W. Mueller, H. Schumann, and C. Tominski, "Visualizing Time-Oriented Data: A Systematic View," *Computers and Graphics*, vol. 31, no. 3, pp. 401-409, 2007.

[2] W. Aigner, C. Kainz, R. Ma, and S. Miksch, "Bertin was Right: An Empirical Evaluation of Indexing to Compare Multivariate Time-Series Data Using Line Plots," *Computer Graphics Forum*, vol. 30, no. 1, pp. 215-228, 2011.

[3] P. Andre, M. Wilson, A. Russell, D.A. Smith, A. Owens, and M.C. Schraefel, "Continuum: Designing Timelines for Hierarchies, Relationships and Scale," *Proc. ACM Symp. User Interface Software and Technology*, 2007.

[4] N. Andrienko and G. Andrienko, *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. first ed., Springer, 2006.

[5] A. Aris, B. Shneiderman, C. Plaisant, G. Shmueli, and W. Jank, "Representing Unevenly-Spaced Time Series Data for Visualization and Interactive Exploration," *Proc. IFIP TC13 Int'l Conf. Human-Computer Interaction*, pp. 835-846, 2005.

[6] R. Arias-Hernandez, L.T. Kaastra, T.M. Green, and B. Fisher, "Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics," *Proc. 44th Hawaii Int'l Conf. System Sciences*, 2011.

[7] R. Bade, S. Schlechtweg, and S. Miksch, "Connecting Time-Oriented Data and Information to a Coherent Interactive Visualization," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '04)*, pp. 105-112, 2004.

[8] E.J. Bergstralh and J.L. Kosanke, "Computerized Matching of Cases to Controls," Technical Report 56, Mayo Clinic, Section of Biostatistics, 1995.

[9] L. Berry and T. Munzner, "BinX: Dynamic Exploration of Time Series Datasets across Aggregation Levels," *Proc. IEEE Symp. Information Visualization*, pp. 215-216, 2004.

[10] D. Brodbeck, R. Gasser, M. Degen, S. Reichlin, and J. Luthiger, "Enabling Large-Scale Telemedical Disease Management through Interactive Visualization," *European Notes in Medical Informatics*, vol. 1, no. 1, pp. 1172-1177, 2005.

[11] ChronoViz, http://chronoviz.com/, 2014.

[12] Y.H. Fua, M.O. Ward, and E.A. Rundensteiner, "Hierarchical Parallel Coordinates for Exploration of Large Datasets," *Proc. Conf. Visualization*, pp. 43-50, 1999.

[13] T. Gschwandtner, W. Aigner, K. Kaiser, S. Miksch, and A. Seyfang, "CareCruiser: Exploring and Visualizing Plans, Events, and Effects Interactively," *Proc. IEEE Pacific Visualization Symp.*, 2011.

[14] 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," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '09)*, pp. 1303-1312, 2009.

[15] W. Javed, B. McDonnel, and N. Elmqvist, "Graphical Perception of Multiple Time Series," *IEEE Trans. Visualization and Computer Graphics*, vol. 16, no. 6, pp. 927-934, Nov./Dec. 2010.

[16] R. Kincaid, "SignalLens: Focus+Context Applied to Electronic Time Series," *IEEE Trans. Visualization and Computer Graphics*, vol. 16, no. 6, pp. 900-907, Nov./Dec. 2010.

[17] R. Kincaid and H. Lam, "Line Graph Explorer: Scalable Display of Line Graphs Using Focus+Context," *Proc. Working Conf. Advanced Visual Interface (AVI '06)*, pp. 404-411, 2006.
[18] M. Krstajic, E. Bertini, and D. Keim, "CloudLines: Compact Display of Event Episodes in Multiple Time-Series," *IEEE Trans. Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2432-2439, Dec. 2011.
[19] B.C. Kwon, B. Fisher, and J.S. Yi, "Visual Analytic Roadblocks for Novice Investigators," *Proc. IEEE Symp. Visual Analytics Science and Technology (VAST '11)*, 2011.
[20] P. McLachlan, T. Munzner, E. Koutsofios, and S. North, "LiveRAC: Interactive Visual Exploration of System Management Time-Series Data," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '08)*, pp. 1483-1492, 2008.
[21] C. Perin, F. Vernier, and J. Fekete, "Interactive Horizon Graphs: Improving the Compact Visualization of Multiple Time Series," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '13)*, pp. 3217-3226, 2013.
[22] H. Reijner, "The Development of the Horizon Graph," *Proc. Workshop on from Theory to Practice: Design, Vision and Visualization*, 2008.
[23] A. Rind, W. Aigner, S. Miksch, S. Wiltner, M. Pohl, F. Drexler, B. Neubauer, and N. Suchy, "Visually Exploring Multivariate Trends in Patient Cohorts Using Animated Scatter Plots," *Proc. Int'l Conf. Ergonomics and Health Aspects of Work with Computers (EHAWC'11)*, pp. 139-148, 2011.
[24] A. Rind, W. Aigner, S. Miksch, S. Wiltner, M. Pohl, T. Turic, and F. Drexler, "Visual Exploration of Time-oriented Patient Data for Chronic Diseases: Design Study and Evaluation," *Proc. Information Quality in e-Health*, pp. 301-320, 2011.
[25] T. Saito, H.N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, and T. Kaseda, "Two-Tone Pseudo Coloring: Compact Visualization for One Dimensional Data," *Proc. IEEE Symp. Information Visualization*, pp. 173-180, 2005.
[26] J. Sanyal, S. Zhang, G. Bhattacharya, P. Amburn, and R. J. Moorhead, "A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1209-1218, Nov./Dec. 2009.
[27] M. Sedlmair, P. Isenberg, D. Baur, M. Mauerer, C. Pigorsch, and A. Butz, "Cardiogram: Visual Analytics for Automotive Engineers," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '11)*, pp. 1727-1736, 2011.
[28] B. Shneiderman and C. Plaisant, "Strategies for Evaluating Information Visualization Tools: Multi-Dimensional In-Depth Long-Term Case Studies," *Proc. AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV '06)*, pp. 1-7, 2006.
[29] E.R. Tufte, *The Visual Display of Quantitative Information*. Graphics Press, 2001.
[30] S. Thakur and T. Rhyne, "Data Vases : 2D and 3D Plots for Visualization Multiple Time Series," *Proc. Int'l Symp. Visual Computing (ISVC '09)*, pp. 929-938, 2009.
[31] J.J. van Wijk and H. van de Wetering, "Cushion Treemaps: Visualization of Hierarchical Information," *Proc. IEEE Symp. Information Visualization*, pp. 73-78, 1999.
[32] M. Vlachos, J. Lin, E. Keogh, and D. Gunopulos, "A Wavelet-Based Anytime Algorithm for K-Means Clustering of Time Series," *Proc. Workshop on Clustering High Dimensionality Data and Its Applications*, pp. 23-30, 2003.
[33] K. Vrotsou, K. Ellegard, and M. Cooper, "Everyday Life Discoveries: Mining and Visualizing Activity Patterns in Social Science Diary Data," *Proc. 11th Int'l Conf. Information Visualization (IV '07)*, pp. 130-138, 2007.
[34] K. Vrotsou, J. Johansson, and M. Cooper, "ActiviTree: Interactive Visual Exploration of Sequences in Event-Based Data Using Graph Similarity," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 6, pp. 945-952, Nov./Dec. 2009.
[35] T.D. Wang, C. Plaisant, B. Shneiderman, N. Spring, D. Roseman, G. Marchand, V. Mukherjee, and M. Smith, "Temporal Summaries: Supporting Temporal Categorical Searching, Aggregation and Comparison," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1049-1056, Nov./Dec. 2009.
[36] T.D. Wang, C. Plaisant, A. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman, "Aligning Temporal Data by Sentinel Events: Discovering Patterns in Electronic Health Records," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '08)*, pp. 457-466, 2008.
[37] K. Wongsuphasawat and B. Shneiderman, "Finding Comparable Temporal Categorical Records: A Similarity Measure with an Interactive Visualization," *Proc. IEEE Symp. Visual Analytics Science and Technology (VAST '09)*, 2009.
[38] K. Wongsuphasawat, J.A. Guerra Gómez, C. Plaisant, T.D. Wang, M. Taieb-Maimon, and B. Shneiderman, "LifeFlow: Visualizing an Overview of Event Sequences," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '11)*, pp. 1747-1756, 2011.
[39] J. Zhao, F. Chevalier, and R. Balakrishnan, "KronoMiner: Using Multi-Foci Navigation for the Visual Exploration of Time-Series Data," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '11)*, pp. 1737-1746, 2011.

Myoungsu Cho received the BS degree in computer science and engineering from Kyungpook National University, Daegu, Korea, and the MS degree in computer science and engineering from Seoul National University, Seoul, Korea in 2002 and 2004, respectively. She is currently working toward the PhD degree in the Department of Computer Science and Engineering, Seoul National University, Seoul, Korea. Her current research interests include HCI and information visualization.

Bohyoung Kim received the BS and MS degrees in computer science from Seoul National University, Seoul, Korea, and the PhD degree in computer science and engineering from the same university in 2001. She is currently an assistant professor in the Department of Radiology, Seoul National University Bundang Hospital, Korea. Her research interests include computer graphics, volume visualization, medical imaging, and information visualization.

Hee-Joon Bae received the MD and MS degrees in Seoul National University College of Medicine from Seoul, Korea, and the PhD degree from the Department of Preventive Medicine from Korea University College of Medicine, Seoul, in 1985, 2000, and 2003, respectively. He is currently an associate professor in the Department of Neurology at Seoul National University College of Medicine and the director of the Stoke Center at Seoul National University, Bundang Hospital. His research interests include acute management of stroke, prevention of stroke, and vascular cognitive impairment.

Jinwook Seo received the BS and MS degrees in computer science from Seoul National University, Seoul, Korea, and the PhD degree in computer science from the University of Maryland, College Park, MD, in 1995, 1997, and 2005, respectively. He is currently an associate professor in the Department of Computer Science and Engineering, Seoul National University, where he is also the director of the Human-Computer Interaction Laboratory. His research interests include HCI, information visualization, and biomedical informatics.