Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 180 (2021) 628–638

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Industry 4.0 and Smart Manufacturing

An Empirical Study of Task-Specific Limitations of the Overview+Detail Technique for Interactive Time Series Analysis

Judith Friedl[a,*], Björn Zimmer[a], Lisa Perkhofer[b], Jan Zenisek[a], Peter Hofer[b], Hans-Christian Jetter[c,a]

[a] *University of Applied Sciences Upper Austria, Hagenberg Campus, Softwarepark 11, A-4232 Hagenberg, Austria*
[b] *University of Applied Sciences Upper Austria, Steyr Campus, Wehrgrabengasse 1, A-4400 Steyr, Austria*
[c] *University of Lübeck, Ratzeburger Allee 160, D-23562 Lübeck, Germany*

Abstract

Interactive visualisation and analysis of time series data is a critical part of many data-driven optimisation processes, particularly in Industry 4.0 and Smart Manufacturing. Time series visualisation enables data analysts or domain experts to visually identify problems such as missing values, sensor drift, precision degradation, or faulty data, before or after algorithmic analysis. A common technique to support the visual exploration of large time series is the *overview+detail* (O+D) technique. O+D provides both detail and context information by displaying a detailed view showing the actual data and a thumbnail overview for showing its context. User studies have shown that users analyse and navigate data sets more efficently and effectively with than without O+D, but that this strongly depends on the task and the nature and amount of the data to be displayed.

We present results of a quantitative user study that was performed on Amazon Mechanical Turk with 95 participants to identify scenarios in which O+D could *not* effectively solve the challenge of visualising large time series. By this, we identify potential usability issues of O+D for typical time series analysis tasks and discuss their origins. For each of these usability issues, we also propose alternative interaction and visualisation designs or other strategies to maintain good usability, even for challenging task types and data densities.

Keywords: interactive visualisation; time series; overview+detail; human-centred design; user study

1. Introduction

Industry 4.0 and Smart Manufacturing often rely on data-driven optimisation of business and manufacturing processes, e.g., using sensor or process data for predictive maintenance [36, 37, 48], smarter logistics [8, 44], or better quality assurance [8, 17]. In such scenarios, it is common that time series data, e.g., large data sets containing sensor measurements or performance indicators over time, need to be visualised to enable data analysts or domain experts to interactively review the data before or after algorithmic analysis [48]. For example, interactive time series visualisation

* Corresponding author. Tel.: +43 5 0804 27187 ; fax: +0-000-000-0000.
E-mail address: judith.friedl@fh-hagenberg.at (Judith Friedl).

10.1016/j.procs.2021.01.285

is of critical importance for visually identifying problems such as missing values, sensor drift, precision degradation, or faulty data.

Using interactive visualisation for time series analysis is a typical example of *visual analytics* [28]. By enabling users to interactively visualise data and adjust parameters, visual analytics seamlessly integrate human skills into the semi-automated analysis process to allow users to see, explore, and understand large amounts of information at once [43]. During interactive analysis and reviewing results, visual analytics also enable experts to apply their tacit domain knowledge that is otherwise difficult or even impossible to formalise [28]. In conclusion, visual analytics of time series data has become a critical part of many data-driven processes, particularly in Industry 4.0 and Smart Manufacturing.

A very common technique to support the visual exploration of large time series on displays of limited sizes is the *overview+detail* (O+D) technique [14]. O+D provides both detail and context information by simultaneously displaying two spatially separated rectangular views; one for the context and one for a more detailed perspective at a higher magnification level [35]. In other terms, O+D interfaces are specialised multi-window arrangements, always displaying the entire content of the time series (or document or map, depending on whatever data set is to be shown) in a compressed form in an *overview* and a subset of the content in much greater detail in the *detail view* [23, 24, 35, 46]. Thereby, the detail view is usually a much larger area or covers the entire rest of the screen and shows a close-up of a portion of the content. The overview usually has a visual marker, highlighting the position of the detail view within the overview (Figure 1). This visual marker in the overview helps users to locate the content of the detail view in relation to the overview, so that they do not lose spatial orientation when zooming too far in or out of the data, i.e., getting lost in "desert fog" [27]. Overviews can also be provided in the form of embellished scrollbars [14] that, in contrast to regular scrollbars, additionally portray semantic information, e.g. values or ranges within a time series (Figure 1).

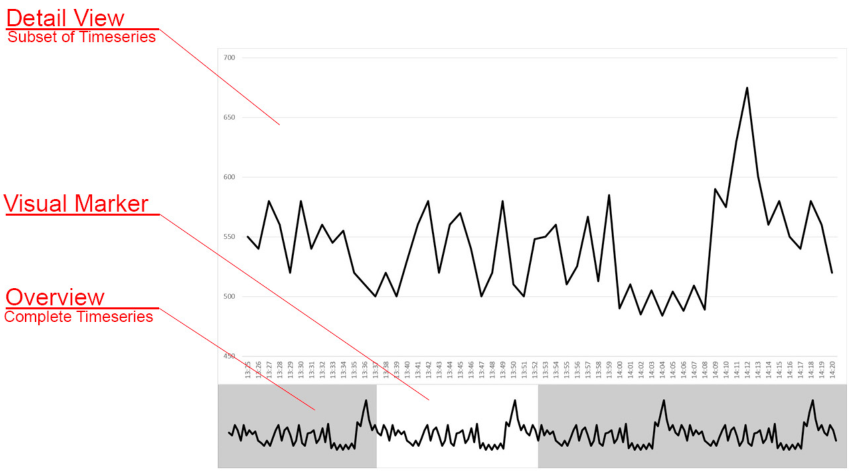

Fig. 1. Schematic view of an O+D chart for time series data. A *detail view* shows a smaller amount of data in great detail at the top. Below that, there is an embellished scrollbar as *overview* that shows a thumbnail view of the entire time series.

Since its early beginnings in 1980's video games [14], O+D has become a ubiquitous interaction and visualisation technique for zooming and panning in maps (e.g. in Google Earth[1]) and documents (e.g. as "navigator panel" in Photoshop[2]), for scrolling in thumbnail representations of text files (e.g. the "minimap" of the Sublime editor[3]), and also in time series visualisation (e.g. as "range selector" in dygraphs[4], Figure 1 A-D). Several user studies have shown that with O+D users are enabled to analyse and navigate data sets more efficiently and effectively than without O+D. However, this depends on the task as well as the nature and the amount of the data to be displayed [12, 23]. O+D is

[1] https://www.google.com/intl/en/earth/, Accessed June 2020
[2] https://www.adobe.com/at/products/photoshop.html, Accessed June 2020
[3] https://www.sublimetext.com/ Accessed June 2020
[4] http://dygraphs.com/, Accessed June 2020

said to improve usability and task efficiency [39] but there are cases in which it was slightly detrimental (e.g. for maps [23]) or even clearly detrimental (e.g for scatter plots on small mobile screens [13]).

In this paper, we therefore identify task-specific limitations of the O+D visualisation technique when performing typical tasks of time series analysis in the context of Industry 4.0 or Smart Production. To the best of our knowledge, there is yet no detailed study of this kind, particularly not differentiating between the typical tasks and data densities in such scenarios. We conducted a quantitative user study on Amazon Mechanical Turk with 95 participants and discuss in which cases our general hypothesis, i.e. *O+D effectively solves usability issues of visualising time series data*, could *not* be confirmed. By this, we identify potential usability challenges of O+D for Industry 4.0 or Smart Production scenarios and discuss where these problems could have resulted from. For each observed issue, we also propose alternative designs or strategies that could help to solve these usability problems and to maintain effectiveness even for the more challenging task types and data densities.

In the following, we will first summarise related work to provide a theoretical background for our research. We then describe our task and study design. This is followed by a statistical analysis of our results and their discussion, including proposals for redesign and strategies to solve the observed usability problems. We conclude with discussing limitations and future work.

2. Related Work

Time series visualisations are used in various fields, ranging from financial data analysis to the exploration of sensor data from production plants. Different application fields require different tasks and amounts of data. For example, in financial data analysis, it might be of greater interest to identify long-term trends in the data using an overview, whereas analysts of sensor data might be more interested in spotting spikes or recurring patterns in a detail view. It is therefore not possible to easily generalise from the conclusions of previous and related work. Instead, we here revisit some O+D foundations and discuss task-specific nuances.

2.1. Revisiting the Benefits of O+D

O+D techniques have been discussed in human-computer interaction research since the early 1990s, e.g. by Beard and Walker [7] and Plaisant et al. [35]. In 1997, Shneiderman suggested that a two-window O+D visualisation for maps and zoomable user interfaces is useful for zoom factors (i.e. magnification ratios between overview and detail view) of 5.0 to 30.0, after which a cascade of overviews could be better [39]. This contradicts more recent studies for more specific tasks that revealed that an overview might not be necessary at all (or even harmful) and could or must be omitted to save screen real estate [23], especially for mobile devices [12, 13]. All these studies, however, did not focus on one-dimensional navigation in time series but on two-dimensional navigation in maps, scatter plots, or documents.

When using O+D for time series, there are further effects that need to be considered: For example, an overview with low display resolution could lead users to wrong conclusions, as spikes in the data get lost due to the aggregated visualisation [18, 19], an effect that also played a key role in our study.

Furthermore, in time series visualisation with O+D for Industry 4.0 or Smart Production, users also typically need much higher zoom factors than the aforementioned 5.0 to 30.0 from [39] since the data is measured over long periods of time and at high sampling rates. In our study, we therefore use three different data amounts (or data densities DD1-DD3) between overview and detail view. While DD1 and DD2 stay roughly within Shneiderman's range for zoom factors between 5.0 and 30.0 [39] (Figure 2 B-D), DD3 has a zoom factor of 103.8 (Figure 2 A).

Similarly, results from user studies in the 1990s or early 2000s may need to be critically revisited for today's much larger and higher-resolution displays. To account for changes in average screen size and resolution since the 1990s and early 2000s, we recruited 95 participants who worked with their own personal screen setup from 2019 and thus with a more realistic sample.

In addition to changing technology, human-centred constraints may also cause non-trivial changes in user performance. Although technology supports displaying high volumes of data, this does not necessarily mean users are cognitively capable of working with this amount of visual information [6, 16, 41]. Again, it seems sensible to revisit results from 1990s and 2000s given the growth in IT literacy in the user population.

2.2. Advanced Techniques for Time Series Visualisation

There are various advanced techniques for the visualisation and analysis of time series data. VisTree [32] and TimeSearcher [10] facilitate O+D techniques and focus on efficient pattern discovery. Buono et al. extend this approach with multiple views to visualise forecasting in time series [11]. A detailed overview of approaches, techniques, methods, and tasks for time series exploration can be found in [4] and [2]. Others use advanced focus+context techniques to allow for zooming into data directly within an overview [5, 29, 30] and research on perception and visualisation of multiple time series exists [21, 26]. Perin et al. performed an evaluation for pan and zooming techniques in multiple time series [33]. Lam et al. compared the use of low resolution against high resolution overviews during the visualisation of multiple time series [31]. Another task that often occurs is the comparison of values between different charts. Horizon graphs [22] compare values between charts for different data densities. Another approach, that uses different data resolutions in charts was introduced by Isenberg et al. [25]. An evaluation of existing methods for time series visualisation and interaction techniques was performed by Walker et al. [45]. Another study focused on the perception of time series with different interactions and visual encodings for various tasks [1].

All the aforementioned work focuses on advanced interaction and visualisation techniques that extend the O+D concept. However, our research shows that in real-world practice these techniques are rarely used and simple O+D implementations with an embellished scrollbar that contains a thumbnail representation of the data, remains the gold standard (Figure 1 A). Even for such comparably simple implementations, there is generally little research on O+D benefits or limitations and, in particular, no research that we are aware of for those tasks that are typical in Industry 4.0 and Smart Production scenarios.

3. User Study

Our main assumption for the design of our study was that O+D is an effective and efficient technique for time series analysis, but that it can fail, or at least expose limitations, for specific *Task Types (TT)* and *Data Densities (DD)*. We therefore conducted a quantitative experiment on Amazon Mechanical Turk[5] (MTurk) with 95 participants to measure task *Completion Time*, *Response Accuracy*, self-reported *Mental Effort*, and self-reported user *Satisfaction* for four different analysis tasks and three *DDs*. Concretely, *Completion Time* is the time the users took to complete a task, *Response Accuracy* describes whether the answer was correct. After each task we then asked the participants how they would rate their *Satisfaction* with the visualisation and their *Mental Effort* for that task. Since we were using MTurk, we could not control for screen resolution or size other than to ask participants to use their desktop computer. However, this provides us with ecological validity in terms of typical screen sizes and resolutions in 2019.

We identified the most challenging combinations of *TT* and *DD*, meaning those combinations that resulted in a significantly longer *Completion Time*, lower *Response Accuracy*, higher *Mental Effort*, or lower *Satisfaction*. We could then provide discussion and interpretations of the collected data to suggest possible explanations and to propose alternative designs or strategies.

3.1. Task Design

There is a great variety of different tasks for time series analysis [2, 4]. For our study, we chose four tasks typically used to get an overview of a time series and its initial analysis:

Task 1: Trend Identification: Tasks like this commonly occur in a financial context when analysing market trends [38] or in smart manufacturing, i.e. when analysing power consumption trends over a larger time period [42] or sensor drift. Participants had to report the current direction of a trend in the data (increasing, decreasing, or no trend observable), see Figure 2 (A). Question: “Please examine the overall distribution of the data and evaluate the global trend (decreasing, increasing, not observable).”

Task 2: Discrimination: The second task was a *Discrimination* task for which users had to look up and compare eight data points that were highlighted by red lines in both, the detail view and the overview, and report the highest value, see Figure 2 (B). This *TT* usually occurs in a scenario where an algorithm suggests and highlights candidates for data points in the time series data that could be especially interesting or indicative for a relevant event. A user then needs to

[5] mturk.com, Accessed June 2020

Fig. 2. Four screenshots (A)-(D) of our O+D implementation in the user study. Each screenshot shows a different task: Trend identification (A), Discrimination (B), Extreme Value Identification (C), and Pattern Recognition (D). (A)-(D) also illustrate differences in data density in the *overview*: The *overview* contains 27,500 data points (DD3) in (A), 8,250 (DD2) data points in (B), or 1,375 (DD1) data points in (C&D). The *detail view* always contains 265 data points.

assess the marked points, discriminate their values, and decide whether a specific data point is actually of interest based on its value and context. Similar tasks were analysed in the *Discrimination* tasks of the studies of Heer et al [22] and Javed et al [26]. To simulate a discrimination task, Question: "Please find the marks highlighted in red and report the largest value."

Task 3: Extreme Value Identification: The third task was a *Extreme Value Identification* task meaning that users had to identify the largest value in the entire time series, see Figure 2 (C). This *TT* occurs for example in the analysis of outliers in sensor data from smart grids [40]. Question: "Please find the largest value in the interactive visualization." One maximum peak per data record.

Task 4: Pattern Recognition/Similarity: The fourth and most challeging task was *Pattern Recognition*. Identifying patterns and analysing their context plays a major role in the monitoring of manufacturing processes [42]. In our study, one small section of the data was highlighted in yellow to emphasise a pattern the participants should look for in a larger section highlighted in green. Then they had to report how many times they had encountered the pattern in the green section, see Figure 2 (D). Question: "The following visualization shows a pattern highlighted in yellow. Please report how often the pattern reoccurs inside the green area."

In the following, we describe how these tasks were presented to the participants and how they were asked to complete them.

3.2. Data Sets

Similar to other studies [3, 15, 20, 22, 26] we used synthetic data for our user studies to be able to control the data values and amount of data for the different tasks. Since we focused on different *DDs (DD1 - DD3)* we created a small data set *(DD1)* with 1,375 data points, a medium sized data set *(DD2)* with 8,250 data points, and a large data set *(DD3)* with 27,500 data points for each of the four tasks. However, for more precise analysis of Task 1, we treated *identifying increasing trends* and *decreasing trends* as separate tasks resulting in *Task 1a* (increasing trend) and *Task 1b* (decreasing trend). Thus, we used 15 data sets in total.

$$3 \; data \; densities \times 5 \; tasks = 15 \; data \; sets$$

For generating the consecutive data points of a time series, we developed a synthetic Auto Regressive Moving Average (ARMA) model [9]. Although ARMA models are traditionally used to describe existing time series, we repurposed a synthetically parameterized model to serve as a time series generator. The developed model consists of a constant base value, a random, normally distributed noise term, which mimics signal or measurement noise in our case, and the weighted moving average values of past signal values and noise terms. In order to simulate the targeted data phenomena (see Figure 2 A-D) , we modified the ARMA model accordingly during the generation process at certain points of time.

For *Task 1a (Increasing Trend)* and *Task 1b (Decreasing Trend)* we used a positive or negative slope to gradually raise or lower the mean value over time.

For *Task 2 (Discrimination)* and *Task 3 (Extreme Value Identification)* we inserted spikes by adding an increment of up to 4.5 to randomly selected values. To control task difficulty and time, each data set had a maximum of eight spikes (one maximum spike, seven distractors). For *Task 2* we then manually selected eight spikes per data set and marked them with a red line.

For *Task 4 (Pattern Recognition)* we created patterns consisting of 25 to 55 consecutive data points with lower values. Each data set for this task contained between four to eight of these patterns (see Figure 2 (D)) and there were no distractor patterns that looked similar to the real one.

3.3. Prototype Implementation

The prototype was implemented as an interactive web visualisation using dygraphs[6] and the meteor framework[7]. It shows 265 data points in the detail view and, depending on *DD1-DD3*, either 1,375, 8,250, or 27,500 data points in the overview.

Since users have to navigate to different areas in the detail view in order to read the exact values, all tasks except for the trend identification require interaction with the time series visualisation. Therefore, the design supports two interactions: (1) Panning: Scrolling through the data sets by left clicking and panning the mouse in the detail view or by left clicking and panning the highlighted area in the overview. (2) Mouseover tooltips: Visualising the exact y-value for a data point at position x when hovering over it with the mouse.

3.4. Participants

We recruited 96 workers on MTurk. We had to reject one worker, since based on the overall time needed and the answers it appeared that he had simply clicked through the study and entered random values. Thus, only the 95 approved workers were used for the analysis and consequently only these 95 workers are reported as participants. Out of the 95 participants 42 were female, 51 were male and two preferred not to disclose their gender. The average age was 39,8 years and ranged from 24 to 69 years. All participants except two were located in the United States. The study took about 35 minutes to complete and each worker was paid ten USD for participation. This is clearly above the minimum wage of 7.25 USD per hour in the United States.

3.5. Measurement

Independent variables: Our first independent variable is *Data Density (DD)* (1,375, 8,250 and 27,500 data points), which represents the zoom factors introduced by [39]. *DD1* with 1,375 data points corresponds to a zoom factor of of 5.2, see Figure 2(C&D), *DD2* with 8,250 data points matches a zoom factor of 31.2, see Figure 2(B), and *DD3* is

[6] http://dygraphs.com/, Accessed June 2020
[7] https://www.meteor.com/, Accessed June 2020

the equivalent of a zoom factor of 103.8, see Figure 2(A). The second independant variable is *Task Type (TT)* (Trend Identification, Discrimination, Extreme Value Identification, Pattern Recognition).
Dependent variable: For our dependent variable we use *Completion Time*, *Response Accuracy*, *Satisfaction* and *Mental Effort*. *Completion Time* was collected by the Online Survey tool[8] in miliseconds (ms), based on how long the users took to continue to the next page. *Response Accuracy* is a binary variable which is either 1 (correct) or 0 (incorrect). *Satisfaction* was determined using a 5-point scale that ranged from 1 (very unsatisfied) to 5 (very satisfied). For *Mental Effort* we used a 7-point scale ranging from 1 (Very low mental effort) to 7 (Very high mental effort).

4. Results & Discussion

A total of 1,425 individual trials were carried out by our 95 participants. Six trials had to be deleted due to missing data. Similar to other research where MTurk was used for evaluating visualisations [34], we eliminated outliers that could result from the absence of a human experimenter during the study by excluding trials that took less than 5% of the maximum task completion time (28 trials) or more than 95% of the maximum task completion time (41 trials). Therefore, the study analysis has been carried out with 1,350 trials.

$$95\ participants \times 5\ tasks \times 3\ repetitions - 75\ excluded\ trials = 1,350\ trials$$

We checked for significant differences in our dependent variables for each *DD* in each TT. For Time and RA we conducted a one-way between groups ANOVA with Post-Hoc Tukey HSD Test for each task. For ME and SAT we ran a Kruskal-Wallis Test as a non-parametric alternative, see Figure 3 for an overview of the results.

4.1. Task 1: Trend Identification

Results: For *Task 1* none of the dependant variables in the tests we conducted showed a significant difference for *DD1-3*. ANOVA *Completion Time*: F(2, 567) = 0.50, p = 0.61; ANOVA *Response Accuracy*: F(2, 567) = 0.55, p = 0.58; Kruskal-Wallis *Satisfaction*: $X^2(2) = 0.00$, p = 1.00; Kruskal-Wallis *Mental Effort*: $X^2(2) = 0.24$, p = 0.89.

Additionally, we analysed the tasks for the positive trend and the tasks for the negative trend separately. However, there was no difference in either of the tests.

Task 1a Positive Trend: ANOVA *Completion Time*: F(2, 282) = 0.20, p = 0.82; ANOVA *Response Accuracy*: F(2, 282) = 0.09, p = 0.92; Kruskal-Wallis *Satisfaction*: $X^2(2) = 0.00$, p = 1.00; Kruskal-Wallis *Mental Effort*: $X^2(2) = 0.22$, p = 0.90.

Task 1b Negative Trend: ANOVA *Completion Time*: F(2, 281) = 0.50, p = 0.61; ANOVA *Response Accuracy*: F(2, 281) = 0.60, p = 0.55; Kruskal-Wallis *Satisfaction*: $X^2(2) = 0.01$, p = 1.00; Kruskal-Wallis *Mental Effort*: $X^2(2) = 1.08$, p = 0.52.

Discussion: Therefore we found no difference between the three different zoom factors in *Task 1*. This result is consistent with our expectations, considering the fact that it was possible to identify the trend by simply looking at the overview. Thus, it is not necessary to achieve optimal usability and mental effort to provide a detail view or a high resolution overview for a simple *Trend Identification Task*. The most important part of the data representation is therefore a simple overview. Without an overview, participants would need to interact with the detail view a lot more, since it would be necessary to use the panning option to scroll through the whole data.

4.2. Task 2: Discrimination

Results: Similar to *Task 1*, none of the tests we conducted for *Task 2* revealed a significant difference. ANOVA *Completion Time*: F(2, 261) = 0.46, p = 0.63; ANOVA *Response Accuracy*: F(2, 261) = 0.14, p = 0.87; Kruskal-Wallis *Satisfaction*: $X^2(2) = 0.02$, p = 0.99; Kruskal-Wallis *Mental Effort*: $X^2(2) = 2.97$, p = 0.23.

Discussion: As in the previous task, there seems to be no effect caused by the different zoom factors in *Task 2*. This result can be explained by the nature of the task. In all trials the users had to compare eight marked data points.The red marks in the overview and detail view are visible in Figure 2 (B). Therefore, there was no difference in workload

[8] https://www.limesurvey.org/de/, Accessed June 2020

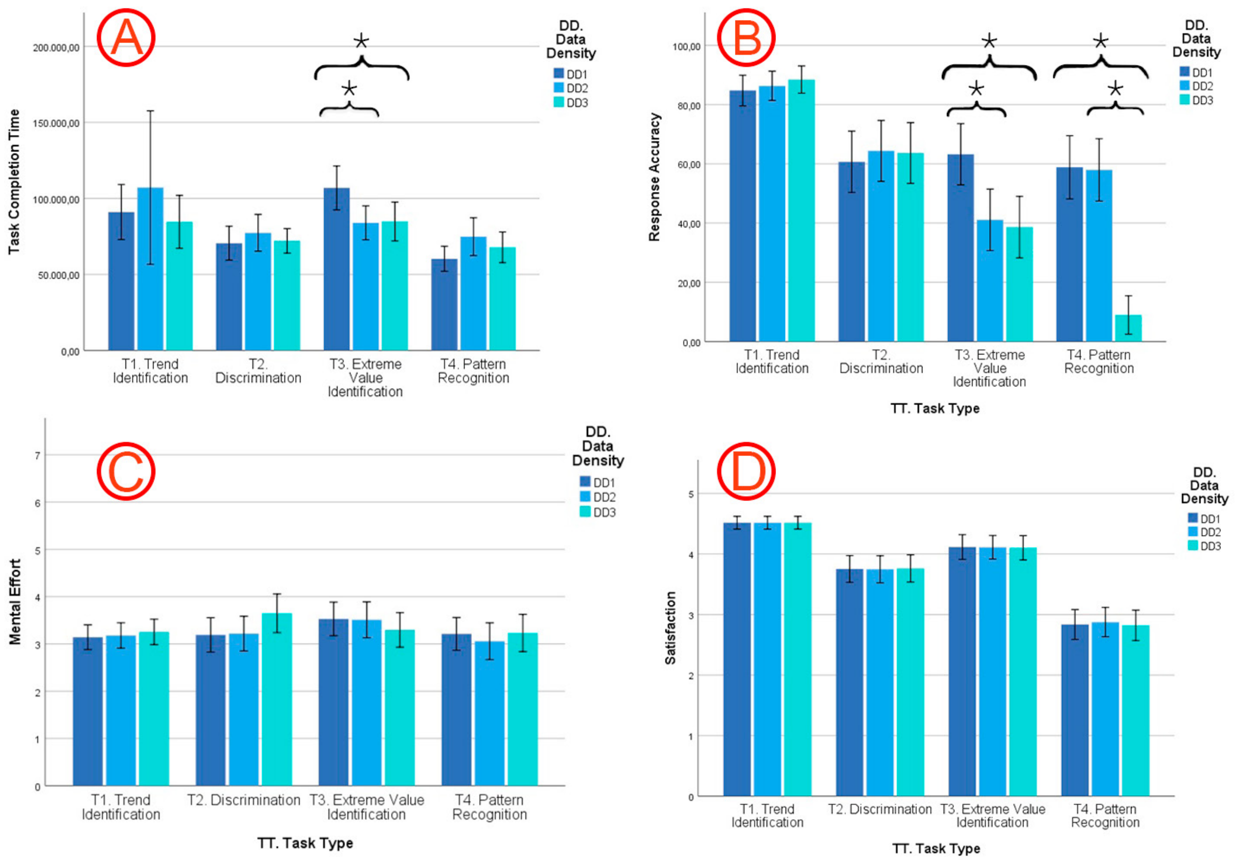

Fig. 3. Overview of the results for each *DD* in each *TT*: Average *Completion Time* in ms (A), Average *Response Accuracy* in % (B), Average *Mental Effort* (C) and Average *Satisfaction* (D). The asterisk and curly brackets marks the statistically significant differences.

for the users. However, it is important to provide an overview for this kind of tasks as it enables the users to navigate directly to the marked data points. Similar to *Task 1* it is not necessary for the overview to have high resolution since the most relevant information in the overview are the marks in the data.

4.3. Task 3: Extreme Value Identification

Results: For *Task 3* we found a significant difference between the different *DDs* regarding *Completion Time* (F(2, 262) = 4.03, p = 0.02) and *Response Accuracy* (F(2, 262) = 6.68, p = 0.00).

Post-Hoc Comparison using a Tukey HSD Test revealed that for the *Completion Time DD1* (M=106,857.01; SD=67,677.13) was significantly different from *DD2* (M=83962.33, SD=53116.79) and *DD3* (M=84,816.93; SD=60,181.81).

For *Response Accuracy* it similarly showed that *DD1* (M=0.63, SD=0.49) was significantly different from *DD2* (M=0.41, SD=0.50) & *DD3* (M=0.39, SD=0.49).

The Kruskal-Wallis Test found no significant difference between the groups for *Satisfaction* $X^2(2) = 0.03$, p = 0.99) and *Mental Effort* ($X^2(2) = 0.97$, p = 0.62).

Therefore, it took the users longer to complete the task when the zoom factor was small. However, the Response Accuracy was significantly higher for the smallest zoom factor. The medium and the large zoom factor lead to a faster task completion while lowering the percentage of correct answers significantly.

Discussion: One possible explanation for this phenomenon is that with a lower zoom factor like in *DD1* the users can identify each peak in the data by simply looking at the overview. The task itself is then similar to *Task 2*. The user simply navigates to every peak they see in the overview. However, with the higher zoom factors, like in *DD2* and *DD3*, more data points have to be displayed per pixel in the overview. Thus, with common aggregation functions, like the one used in this study, the overview will render a mean of the values that have to be displayed on each pixel. This results in some peaks not being visible at all within the overview. This issue of low resolution has also been reported in previous research [7, 31]. The users then need to navigate through all data points in the detail view. To do so, users often pan the overview and use it like a scroll bar. However, with higher *DD*, like *DD2* & *DD3*, this automatically leads to a lower control display ratio. Thus, a small movement of the pointing device in the overview leads to faster scrolling in the detail view, covering a larger amount of data points than a similar movement in *DD1*. While this interaction technique enables the user to quickly go through the data and usually results in lower completion time, it may also lead to the user overlooking some of the peaks. Thus, the time needed for completing the task is lower but there is a higher percentage of incorrect answers.

Design Implications: This problem could be solved by applying a different method that considers data point importance when rendering the overview.Such an approach is described in [19] using perceptually important points. This would enable the user to identify the peaks in the overview even with large zoom factors.

4.4. Task 4: Pattern Recognition

Results: Like for *Task 3* we found a significant difference between the *DDs* for *Task 4* when it comes to *Response Accuracy* (F(2, 248) = 33.63, p = 0.00). However, there was no significant difference in the *Completion Time* (F(2, 248) = 1.97, p = 0.14).

A Post-Hoc Tukey HSD Test revealed that *DD3* (M=0.09, SD=0.29) is significantly different from *DD1* (M=0.59, SD=0.50) & *DD2* (M=0.58, SD=0.50).

The Kruskal-Wallis Test found no significant difference between the groups for *Satisfaction* $X^2(2) = 0.08$, p = 0.96) and *Mental Effort* ($X^2(2) = 1.07$, p = 0.59).

In summary, there is no significant difference in either *Completion Time*, *Mental Effort* or *Satisfaction*. However, *Response Accuracy* is significantly lower in the condition with the largest zoom factor.

Discussion: This may be due to the fact, that with the lower zoom factors of *DD1* & *DD2* the patterns can still be identified by looking at the overview. However, for *DD3*, patterns are no longer clearly visible in the overview. The user therefore has to scroll through the data. As mentioned above, the control display ratio is much lower in *DD3*. Thus, users can easily overlook a relevant pattern.

Design Implications: To overcome this challenge for high *DD*, higher factors for control display ratio may be helpful.Users would then take longer to scroll through the whole detail view, however, it may be easier to find patterns. A different approach would be a fish eye lense in the overview, similar to distortion lenses like described in [29]. This interaction technique would enable the user to quickly run through the overview using the lense to identify similar patterns. In that case a detail view might not be necessary at all. Another option would be to implement an interaction technique where users can sketch or highlight a pattern they would like to look for and an algorithm runs through the data highlighting those patterns, similar to [11, 47]. Like in *Task 2* the user would then simply navigate to the marked areas and decide if the pattern is relevant for their analysis, based on the context of the data points surrounding the pattern.

5. Conclusion

In this paper, we reported the results of an experimental user study on the limitations of overview+detail in time series analysis. Visual analysis of time series data enables domain experts in fields like Industry 4.0 and Smart Manufacturing to identify problems while drawing on their implicit domain knowledge. We looked at the differences in *Completion Time*, *Response Accuracy*, *Mental Effort* and *Satisfaction* which are caused by different zoom factors in four different tasks. Our results showed no differences for all dependent variables for the *Trend Identification* and *Discrimination* task, but revealed a difference in *Completion Time* and *Response Accuracy* in the *Extreme Value Identification* Task and a difference in *Response Accuracy* in the *Pattern Identification Task*. Since we did not study the influence of screen size and resolution on the four dependant variables, this could be addressed by future work using a controlled in person experiment. Furthermore, we discussed interaction techniques as well as visualisation techniques (such as reducing the overview dimensionality by using perceptually important points or using distortion lenses within the overview) to mitigate the challenges posed by high zoom factors within these tasks.

Acknowledgements

This project is partially financed by research subsidies granted by the government of Upper Austria for the project X-Pro, and the USIVIS project of the Austrian Research Promotion Agency (FFG).

References

[1] Adnan, M., Just, M., Baillie, L., 2016. Investigating time series visualisations to improve the user experience, in: Proc. CHI '16, ACM, New York, NY, USA. pp. 5444–5455. doi:10.1145/2858036.2858300.
[2] Aigner, W., Miksch, S., Schumann, H., Tominski, C., 2011. Visualization of Time-Oriented Data. 1st ed., Springer Publishing Company, Incorporated, London.
[3] Albers, D., Correll, M., Gleicher, M., 2014. Task-driven evaluation of aggregation in time series visualization, in: Proc. CHI '14, ACM, New York, NY, USA. pp. 551–560. doi:10.1145/2556288.2557200.
[4] Andrienko, N., Andrienko, G., 2005. Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer-Verlag, Berlin, Heidelberg.
[5] Balakrishnan, R., Chevalier, F., Zhao, J., Pietriga, E., 2011. Exploratory analysis of time-series with chronolenses. IEEE TVCG 17, 2422–2431. doi:10.1109/TVCG.2011.195.
[6] Bawden, D., Robinson, L., 2009. The dark side of information: overload, anxiety and other paradoxes and pathologies. Journal of Information Science 35, 180–191. doi:10.1177/0165551508095781.
[7] Beard, D.V., II, J.Q.W., 1990. Navigational techniques to improve the display of large two-dimensional spaces. Behaviour & Information Technology 9, 451–466. doi:10.1080/01449299008924259.
[8] Beham, A., Raggl, S., Hauder, V.A., Karder, J., Wagner, S., Affenzeller, M., 2020. Performance, quality, and control in steel logistics 4.0. Procedia Manufacturing 42, 429 – 433. doi:https://doi.org/10.1016/j.promfg.2020.02.053. international Conference on Industry 4.0 and Smart Manufacturing (ISM 2019).
[9] Box, G.E.P., Meirion, J.G., 1970. Time series analysis: forecasting and control. 2nd ed., Holden-Day, San Francisco, CA, USA.
[10] Buono, P., Aris, A., Plaisant, C., Khella, A., Shneiderman, B., 2005. Interactive pattern search in time series, in: Visualization and Data Analysis 2005, Society of Photo-Optical Instrumentation Engineers (SPIE). pp. 5669 – 5669 – 12. doi:10.1117/12.587537.
[11] Buono, P., Plaisant, C., Simeone, A., Aris, A., Shneiderman, B., Shmueli, G., Jank, W., 2007. Similarity-based forecasting with simultaneous previews: A river plot interface for time series forecasting, in: 2007 11th International Conference Information Visualization (IV '07), IEEE, Zurich, Switzerland. pp. 191–196. doi:10.1109/IV.2007.101.
[12] Burigat, S., Chittaro, L., 2013. On the effectiveness of overview+detail visualization on mobile devices. Personal Ubiquitous Comput. 17, 371–385. doi:10.1007/s00779-011-0500-3.
[13] Büring, T., Gerken, J., Reiterer, H., 2006. Usability of overview-supported zooming on small screens with regard to individual differences in spatial ability, in: Proc. AVI '06, Association for Computing Machinery, New York, NY, USA. p. 233–240. doi:10.1145/1133265.1133310.
[14] Cockburn, A., Karlson, A., Bederson, B.B., 2009. A review of overview+detail, zooming, and focus+context interfaces. ACM Comput. Surv. 41. doi:10.1145/1456650.1456652.
[15] Correll, M., Albers, D., Franconeri, S., Gleicher, M., 2012. Comparing averages in time series data, in: Proc. CHI '12, ACM, New York, NY, USA. pp. 1095–1104. doi:10.1145/2207676.2208556.
[16] Falschlunger, L., Lehner, O., Treiblmaier, H., 2016. Infovis: The impact of information overload on decision making outcome in high complexity settings, in: Proc. SIG HCI '16, pp. 1–5.
[17] Frustaci, F., Perri, S., Cocorullo, G., Corsonello, P., 2020. An embedded machine vision system for an in-line quality check of assembly processes. Procedia Manufacturing 42, 211 – 218. doi:https://doi.org/10.1016/j.promfg.2020.02.072. international Conference on Industry 4.0 and Smart Manufacturing (ISM 2019).
[18] Fu, T.c., 2011. A review on time series data mining. Engineering Applications of Artificial Intelligence 24, 164–181. doi:10.1016/j.engappai.2010.09.007.
[19] Fu, T.c., Lai Chung, F., Kwok, K.y., Ng, C.m., 2008. Stock time series visualization based on data point importance. Engineering Applications of Artificial Intelligence 21, 1217–1232. doi:10.1016/j.engappai.2008.01.005.
[20] Fuchs, J., Fischer, F., Mansmann, F., Bertini, E., Isenberg, P., 2013. Evaluation of alternative glyph designs for time series data in a small multiple setting, in: Proc. CHI '13, ACM, New York, NY, USA. pp. 3237–3246.
[21] Hao, M.C., Marwah, M., Janetzko, H., Dayal, U., Keim, D.A., Patnaik, D., Ramakrishnan, N., Sharma, R.K., 2012. Visual exploration of frequent patterns in multivariate time series. Information Visualization 11, 71–83. doi:10.1177/1473871611430769.
[22] Heer, J., Kong, N., Agrawala, M., 2009. Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations, in: Proc. CHI '09, ACM, New York, NY, USA. pp. 1303–1312. doi:10.1145/1518701.1518897.
[23] Hornbæk, K., Bederson, B.B., Plaisant, C., 2002. Navigation patterns and usability of zoomable user interfaces with and without an overview. ACM Trans. Comput.-Hum. Interact. 9, 362–389. doi:10.1145/586081.586086.

[24] Hornbæk, K., Frøkjær, E., 2001. Reading of electronic documents: The usability of linear, fisheye, and overview+detail interfaces, in: Proc. CHI '01, ACM, New York, NY, USA. pp. 293–300. doi:10.1145/365024.365118.

[25] Isenberg, P., Bezerianos, A., Dragicevic, P., ekete, J.D., 2011. A study on dual-scale data charts. IEEE TVCG 17, 2469–2478.

[26] Javed, W., McDonnel, B., Elmqvist, N., 2010. Graphical perception of multiple time series. IEEE TVCG 16, 927–934. doi:10.1109/TVCG.2010.162.

[27] Jul, S., Furnas, G.W., 1998. Critical zones in desert fog: Aids to multiscale navigation, in: Proc. UIST '98, Association for Computing Machinery, New York, NY, USA. p. 97–106. doi:10.1145/288392.288578.

[28] Keim, D., Andrienko, G., Fekete, J.D., Görg, C., Kohlhammer, J., Melançon, G., 2008. Visual Analytics: Definition, Process, and Challenges. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 154–175. doi:10.1007/978-3-540-70956-5_7.

[29] Kincaid, R., 2010. SignalLens: Focus plus Context Applied to Electronic Time Series. IEEE TVCG 16, 900–907. doi:10.1109/tvcg.2010.193.

[30] Kincaid, R., Lam, H., 2006. Line graph explorer: Scalable display of line graphs using focus+context, in: Proc. AVI '06, ACM, New York, NY, USA. pp. 404–411. doi:10.1145/1133265.1133348.

[31] Lam, H., Munzner, T., Kincaid, R., 2007. Overview use in multiple visual information resolution interfaces. IEEE TVCG 13, 1278–1285. doi:10.1109/TVCG.2007.70583.

[32] Lin, J., Keogh, E., Lonardi, S., 2005. Visualizing and discovering non-trivial patterns in large time series databases. Information Visualization 4, 61–82. doi:10.1057/palgrave.ivs.9500089.

[33] Perin, C., Vernier, F., Fekete, J.D., 2013. Interactive horizon graphs: improving the compact visualization of multiple time series, in: Proc. CHI '13, ACM, New York, NY, USA. pp. 3217–3226. doi:10.1145/2470654.2466441.

[34] Perkhofer, L., Walchshofer, C., Hofer, P., 2020. Does design matter when visualizing big data? an empirical study to investigate the effect of visualization type and interaction use. Journal of Management Control , 1–41.

[35] Plaisant, C., Carr, D., Shneiderman, B., 1995. Image browsers: Taxonomy, guidelines, and informal specifications. IEEE Software 12, 21–32.

[36] Poor, P., Basl, J., Ženíšek, D., 2020. Assessing the predictive maintenance readiness of enterprises in west bohemian region. Procedia Manufacturing 42, 422 – 428. doi:https://doi.org/10.1016/j.promfg.2020.02.098. international Conference on Industry 4.0 and Smart Manufacturing (ISM 2019).

[37] Quatrini, E., Costantino, F., Pocci, C., Tronci, M., 2020. Predictive model for the degradation state of a hydraulic system with dimensionality reduction. Procedia Manufacturing 42, 516 – 523. doi:https://doi.org/10.1016/j.promfg.2020.02.039. international Conference on Industry 4.0 and Smart Manufacturing (ISM 2019).

[38] Schreck, T., Tekušová, T., Kohlhammer, J., Fellner, D., 2007. Trajectory-based visual analysis of large financial time series data. SIGKDD Explor. Newsl. 9, 30–37. doi:10.1145/1345448.1345454.

[39] Shneiderman, B., 1997. Designing the User Interface: Strategies for Effective Human-Computer Interaction. 3rd ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[40] Sun, L., Zhou, K., Zhang, X., Yang, S., 2018. Outlier data treatment methods toward smart grid applications. IEEE Access 6, 39849–39859.

[41] Sweller, J., 2010. Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational Psychology Review 22, 123–138. doi:10.1007/s10648-010-9128-5.

[42] Tao, F., Qi, Q., Liu, A., Kusiak, A., 2018. Data-driven smart manufacturing. Journal of Manufacturing Systems 48, 157–169. doi:https://doi.org/10.1016/j.jmsy.2018.01.006.

[43] Thomas, J.J., Cook, K.A., 2005. Illuminating the Path: The Research and Development Agenda for Visual Analytics. National Visualization and Analytics Center. URL: https://www.hsdl.org/?view&did=485291.

[44] Vieira, A.A., Dias, L., Santos, M.Y., Pereira, G.A., Oliveira, J., 2020. Supply chain risk management: an interactive simulation model in a big data context. Procedia Manufacturing 42, 140 – 145. doi:https://doi.org/10.1016/j.promfg.2020.02.035. international Conference on Industry 4.0 and Smart Manufacturing (ISM 2019).

[45] Walker, J., Borgo, R., Jones, M.W., 2016. Timenotes: A study on effective chart visualization and interaction techniques for time-series data. IEEE TVCG 22, 549–558. doi:10.1109/TVCG.2015.2467751.

[46] Wang Baldonado, M.Q., Woodruff, A., Kuchinsky, A., 2000. Guidelines for using multiple views in information visualization, in: Proc. AVI '00, ACM, New York, NY, USA. pp. 110–119. doi:10.1145/345513.345271.

[47] Wattenberg, M., 2001. Sketching a graph to query a time-series database, in: CHI EA '01, Association for Computing Machinery, New York, NY, USA. p. 381–382. doi:10.1145/634067.634292.

[48] Zimmer, B., Zenisek, J., Jetter, H.C., 2018. Towards uncertainty visualization in smart production environments, in: Proc. VINCI '18, Association for Computing Machinery, New York, NY, USA. p. 116–117. doi:10.1145/3231622.3231640.