Clustered Eye Movement Similarity Matrices

Ayush Kumar
Stony Brook University
New York, USA
aykumar@cs.stonybrook.edu

Neil Timmermans
Eindhoven University of Technology
Eindhoven, The Netherlands
n.timmermans@tue.nl

Michael Burch
Eindhoven University of Technology
Eindhoven, The Netherlands
m.burch@tue.nl

Klaus Mueller
Stony Brook University
New York, USA
mueller@cs.stonybrook.edu

ABSTRACT

Eye movements recorded for many study participants are difficult to interpret, in particular when the task is to identify similar scanning strategies over space, time, and participants. In this paper we describe an approach in which we first compare scanpaths, not only based on Jaccard (JD) and bounding box (BB) similarities, but also on more complex approaches like longest common subsequence (LCS), Frechet distance (FD), dynamic time warping (DTW), and edit distance (ED). The results of these algorithms generate a weighted comparison matrix while each entry encodes the pairwise participant scanpath comparison strength. To better identify participant groups of similar eye movement behavior we reorder this matrix by hierarchical clustering, optimal-leaf ordering, dimensionality reduction, or a spectral approach. The matrix visualization is linked to the original stimulus overplotted with visual attention maps and gaze plots on which typical interactions like temporal, spatial, or participant-based filtering can be applied.

CCS CONCEPTS

• **Human-centered computing → Visualization techniques**;

KEYWORDS

Eye tracking, scanpath comparison, matrix reordering, information visualization, visual analytics

ACM Reference Format:

Ayush Kumar, Neil Timmermans, Michael Burch, and Klaus Mueller. 2019. Clustered Eye Movement Similarity Matrices. In *Eye Tracking and Visualization (ETVIS @ ETRA'19), June 25–28, 2019, Denver , CO, USA.* ACM, New York, NY, USA, Article 4, 9 pages. https://doi.org/10.1145/3317958.3319811

1 INTRODUCTION

Detecting similarities among scanpaths can be of importance, in particular, if the task is to identify strategic difficulties in understanding visual stimuli [Andrienko et al. 2012; Burch et al. 2013]. For example, participants in an eye tracking study might be grouped into different categories of eye movement behavior, while all of those categories are representatives of certain phenomena worth investigating [Yarbus 1967], like the typical and normal eye movement behavior or anomalies and outliers that have to be investigated in more detail.

However, the spatio-temporal patterns are difficult to be grouped or categorized by just one traditional measure [Duchowski 2003; Holmqvist et al. 2011] and hence, we provide a way to improve this grouping strategy. In our approach we first build comparison matrices computed by typical well-known comparison techniques like Jaccard (JD) and bounding box (BB) similarities, but also on more complex approaches like longest common subsequence (LCS), Frechet distance (FD), dynamic time warping (DTW), and edit distance (ED).

These comparison values build a matrix-like scheme, but without further processing the values in terms of reordering the matrices, it is not possible to find the aforementioned scanpath categorizations. In our approach we provide several matrix reordering techniques [Behrisch et al. 2016] that bring structure into the comparison values like hierarchical clustering, optimal-leaf ordering, dimensionality reduction, or a spectral approach.

We represent the comparison values in a color coded and reordered matrix (see Figure 1) that can be scaled down to pixel size in case the number of scanpaths grows into the thousands [Blascheck et al. 2015] [Kumar et al. 2018a]. The data analyst can interact with the visualizations to explore, filter, and navigate in the scanpath data while they are linked with complementary views like visual attention maps, gaze plots, and the original stimuli.

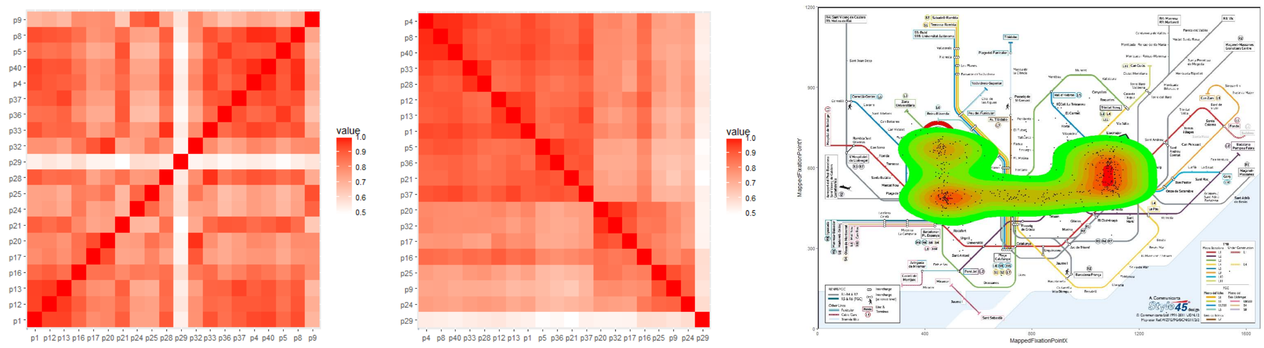
We show the usefulness of the scanpath comparison methods and matrix reorderings by applying it to real-world eye movement data recorded in a route finding experiment in public transport maps [Netzel et al. 2017].

2 RELATED WORK

Scanpaths recorded for many participants produce a data source that is challenging to analyze for common scanning strategies. This is due to the vast amount of data [Blascheck et al. 2015] and the spatio-temporal nature of the data. However, there are already many visualization approaches [Blascheck et al. 2017] and also visual analytics techniques [Andrienko et al. 2012; Burch et al. 2013] that try to combine algorithmic, visual, and interactive concepts, but in

ETVIS @ ETRA'19, June 25–28, 2019, Denver , CO, USA

ACM ISBN 978-1-4503-6729-5/19/06…$15.00
https://doi.org/10.1145/3317958.3319811



(a) (b) (c)

Figure 1: Color coded adjacency matrices showing the strengths of pairwise scanpath comparisons: The initial similarity matrix in (a) is further ordered by a hierarchical clustering approach (b) while the stimulus with the larger cluster of study participants identified in (b) can be seen in (c).

many cases they do not scale to many scanpaths and do not provide an overview about different scanning behavior categories.

For example, visual attention maps [Bojko 2009; Burch 2016; Spakov and Miniotas 2007] might be useful to detect hot spots in the visual attention, but due to aggregation and overplotting it is pretty difficult to categorize the scanning behavior into groups and categories. The reason for this drawback is that pairwise similarities between the scanpaths are not computed beforehand and the similarity values are also not further ordered and clustered.

Gaze plots [Goldberg and Helfman 2010] on the other hand, show the individual scanpaths overplotted on the visual stimulus, but if many of those are shown, it is pretty difficult or even impossible to find groups of similar scanning behavior. This negative effect comes from the increased visual clutter [Rosenholtz et al. 2005] caused by the line-based representation and the many line overdrawings and crossings.

In our approach we reduce each scanpath to similarity values which are computed in comparison to each other scanpath. Hence, we model those similarities as 2D matrices consisting of real-valued percentage numbers describing the degree of similarity. Those matrices are then reordered, clustered, and color coded based on the given similarity values.

The work by Kumar et al. [Kumar et al. 2016] [Kumar et al. 2018b] is related to our approach. Similarity matrices are computed, but their technique rather focuses on metric-based grouping of eye movements, not on comparing the scanpaths by different options and clustering by several techniques. Most of the existing approaches provide views on stacked scanpaths [Burch et al. 2013; Raschke et al. 2012] while sometimes the stacking order is computed by clustering techniques. The clustering makes use of distance values given in a pairwise scanpath comparison matrix, for example, based on image thumbnail similarities as in the work by Kurzhals et al. [Kurzhals et al. 2016a,b].

3 SCANPATH COMPARISON AND VISUALIZATION

The tool is implemented in Python and consists of several components for exploring eye movements for similar or dissimilar visual scanning patterns. To reach our goal we compare eye movements, order them, visualize those comparison values, and finally, allow interactions in the provided views while they are linked to the original visual stimulus.

3.1 Design Criteria

Based on the aforementioned summary we focus on several requirements for the tool to make it applicable to eye movements in order to derive common scanning strategies, i.e., to identify a good categorization of the eye movement patterns.

- **Stimuli and participants:** The recorded eye movement data can be analyzed for certain stimuli as well as a list of user-selected participants.
- **Scanpath comparison:** The requested scanpath data for the stimulus given by the selected participants can be compared, either completely, or filtered for space and time.
- **Matrix reordering:** The comparison values alone do not provide a structured and clustered view on the scanpath data, hence it must be improved by further matrix reordering techniques.
- **Interactivity:** All provided views and visualizations are interactive while the algorithmic approaches can be exchanged to see the scanpath data from different perspectives.
- **Linked views:** All views and visualizations are linked, meaning a change in one view immediately changes all other views.
- **Easy to use:** The provided algorithms, visualizations, views, and interactions are easy to understand and not much experience is required to get started with the tool.
- **Extendable:** The tool can be extended by extra functionality, e.g., additional comparison algorithms and reordering approaches, but also more visualizations.

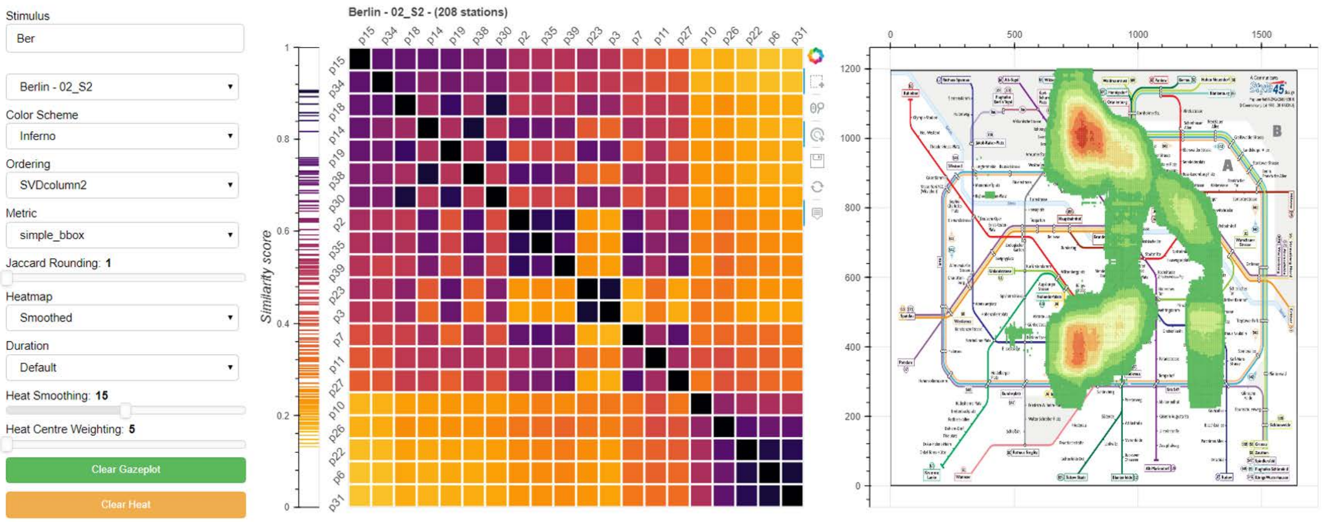


Figure 2: The graphical user interface (GUI) consists of three views while those are interactive and linked: Left: The input panel. Center: The adjacency matrix panel. Right: The stimulus panel with visual attention maps (and gaze plots) if the small cluster at the bottom right is selected.

The graphical user interface of our visualization tool can be seen in Figure 2. It consists of three major views which are the input panel, the adjacency matrix panel with extra similarity value distributions (see Figure 3 (a)), and the stimulus panel overdrawn with a visual attention map and a gaze plot (see Figure 3 (b)).

3.2 Scanpath Comparison

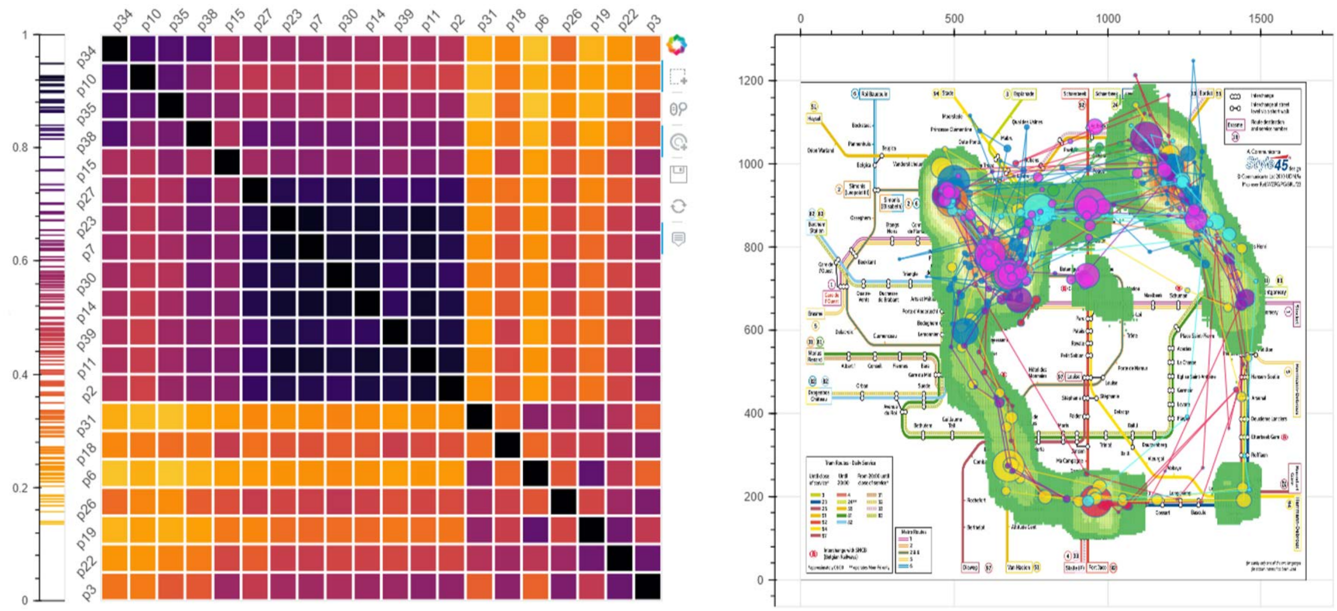
We provide an extendable list of well-known comparison algorithms with Jaccard coefficient, bounding box, longest common subsequence, Frechet distance, dynamic time warping, and edit distance already implemented.

- **Jaccard coefficient (JD):** We define the similarity between two scanpaths by the Jaccard coefficient interpreting each scanpath as a set of fixation points while each can be given a fixation radius to increase the probability of a similarity matching [Levandowsky and Winter 1971]. The Jaccard index is hereby given by the cardinality of the intersection of both sets divided by the cardinality of the union of both sets giving a value between 0 and 1.
- **Bounding box (BB):** The similarity between two scanpaths can also be defined as the size of the overlapping area of the smallest bounding or enclosing boxes for each of the scanpaths divided by the total covered area [Toussaint 1983]. This also gives a value between 0 and 1 while the scanpaths can be filtered for space and time.
- **Longest common subsequence (LCS):** The NP-hard problem of finding the longest common subsequence (LCS) can be used as a similarity measure if the sequence of fixations is taken into account [Maier 1978]. Each fixation sequence is first mapped to a sequence of uniquely labeled AOIs, i.e., a string. This results in a value between 0 and 1 if the length of the LCS divided by the length of the longest sequence is computed.
- **Frechet distance (FD):** The similarity of two curves is given as the length of the shortest leash that can be used to traverse two separate scanpaths [Alt and Godau 1995]. This is a valuable approach if two scanpaths vary a lot over space and time. Computing the maximum of the distances to both paths gives a value between 0 and 1.
- **Dynamic time warping (DTW):** Since scanpath data has a time-dependent nature we also apply a concept from time-series analysis. This is in particular useful, if scanpaths vary a lot in speed (e.g. number of fixations per time unit) [Silva and Batista 2016]. We derive a value between 0 and 1 by computing the difference to the other two sequences.
- **Edit distance (ED):** We first transform two scanpaths into two strings by mapping each fixation to uniquely labeled areas of interest. Then we compute the minimum number of operations to transform one string into the other [Navarro 2001]. The number of required operations divided by the maximum number of operations results in a value between 0 and 1.

3.3 Matrix Reordering

If the comparison values were represented by an unordered matrix visualization, it would be difficult or even impossible to detect group structures in the scanpath data, consequently, we support matrix reordering techniques [Behrisch et al. 2016].

- **Hierarchical clustering:** Taking the comparison values for clustering the matrix hierarchically is a good concept, however, the data groups might not be placed along the diagonal of the matrix. However, we used agglomerative hierarchical clustering to produce an ordered matrix [Eisen et al. 1998].



(a) (b)

Figure 3: Visualizations of the eye movement data for the public transport map of Brussels: (a) An ordered and color coded adjacency matrix with a similarity value distribution. (b) The visual stimulus overdrawn with a visual attention map and a gaze plot if the large cluster in the upper left is selected.

- **Optimal-leaf ordering:** Ordering the scanpaths based on the similarities of matrix neighbors results in an ordered matrix. The hierarchical binary tree is used for this ordering strategy. We follow the concept described by Bar-Joseph et al. [Bar-Joseph et al. 2001].
- **Dimensionality reduction:** Multi-dimensional scaling is also applied to matrices [Spence and Graef 1974] for reordering purposes. The scanpaths at the rows and columns are indexed in a way that distances correspond to the dissimilarities as good as possible.
- **Spectral approach:** We use a variant of the rank-two ellipse seriation by Chen [houh Chen 2002] to reorder a scanpath comparison matrix. The goal is to use Eigenvalues and Eigenvectors as a projection into the Eigenspace.

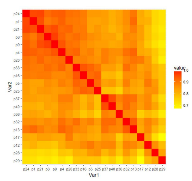
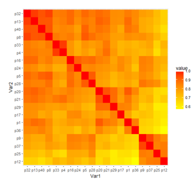
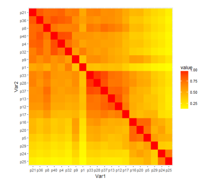
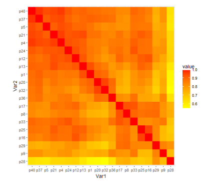
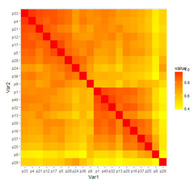
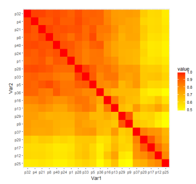
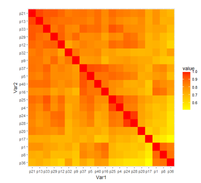
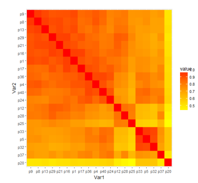
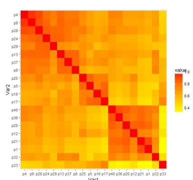
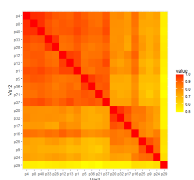
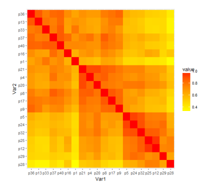
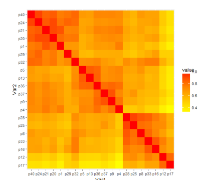
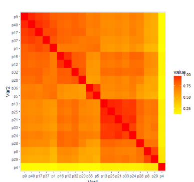
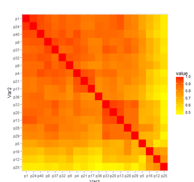
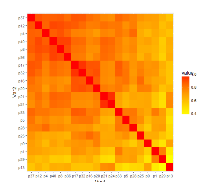
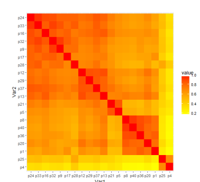
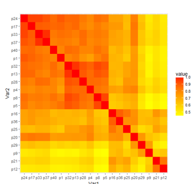
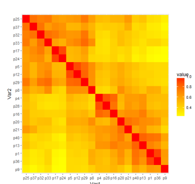
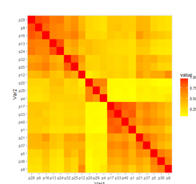
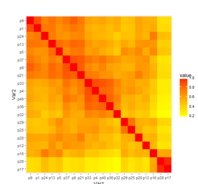
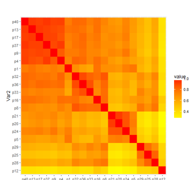
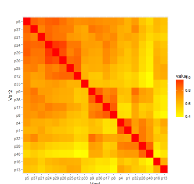
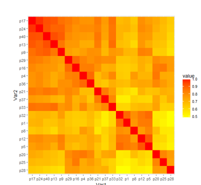
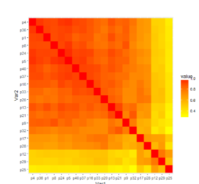
It may be noted that many more matrix reordering techniques might be tested in future. Our tool is implemented in a way that the corresponding algorithms can be added easily since all of them share the same input and output parameters, only the reordering algorithms have to be adjusted.

3.4 Linked Visualization Techniques

We support five major views on the scanpath data, i.e., a matrix visualization, a color coded comparison value distribution bar code, a visual attention map, a gaze plot, as well as the stimulus with overplotted visualizations, shown in Figures 2 and 3.

- **Matrix visualization:** The matrix shows a color coded version of the ordered or unordered comparison values. In this view scanpath groups can be selected which are displayed in a corresponding view on the stimulus. Different reordering algorithms can be applied while the impact of these algorithms on the clustering and ordering can be seen directly.
- **Comparison value distribution bar code:** The number of comparison values can be immense and hence, a histogram-like distribution overview supports the identification of comparison value frequencies. Additional color codings show the value range of the distributions.
- **Visual attention map:** Showing the spatial information of the stimulus together with the visual attention as an aggregated measure over all scanpaths is a useful visual concept in order to provide an overview. The visual attention map can be filtered for densities or spatial regions while the comparison matrices can be adjusted based on the filtered data.
- **Gaze plot:** Exploring the scanpath data in form of line-based representations can provide some insights about the scanning behavior, however, if the number of scanpaths is too large, visual clutter may be a problem in these plots. After filtering, a gaze plot may be a good alternative to the comparison matrix since here the scanpaths are shown completely, and are not aggregated to single real-valued numbers.
- **Stimulus:** To provide contextual information combined with comparison values, visual attention, and scanpaths, we also show the visual stimulus. This can be overplotted by a visual attention map or a gaze plot which is useful to filter the scanpath data based on certain semantic information given by attention hotspots or areas that are not frequently visited.

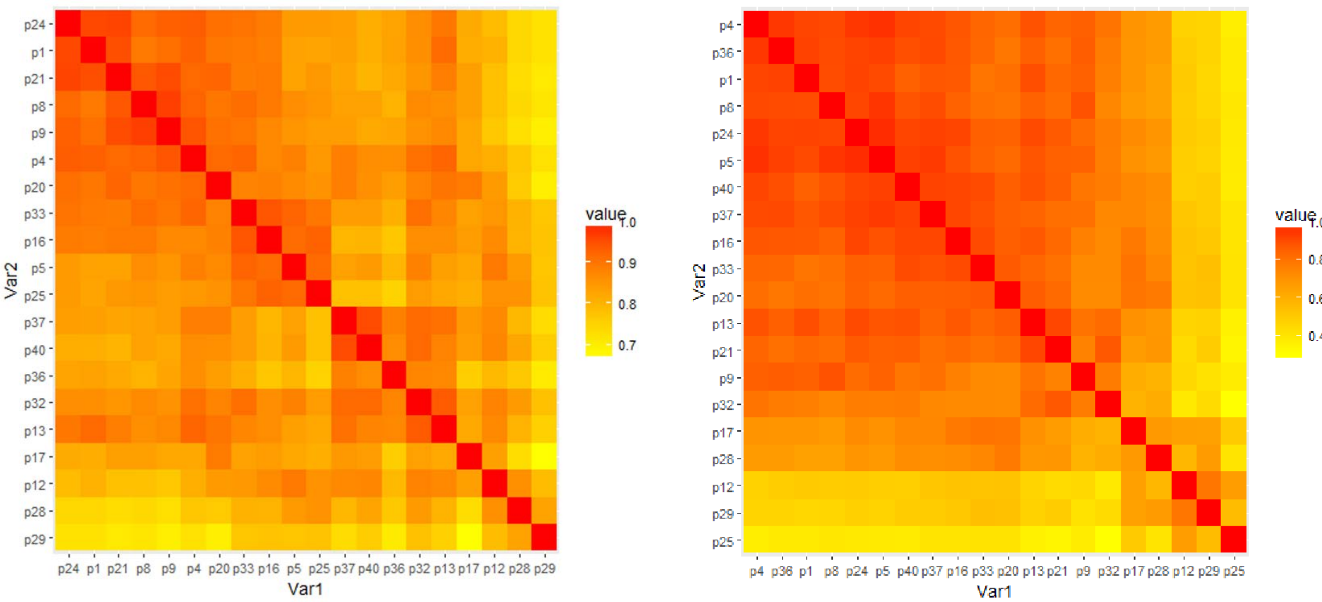
Table 1: Comparison methods for scanpaths and matrix reordering for the pairwise comparison values

	Hierarchical clustering	Optimal-leaf ordering	Dimensionality reduction	Spectral approach
JD				
BB				
LCS				
FD				
DTW				
ED				

3.5 Tool Features and Interactions

The tool provides several features to this end. We are aware of the fact that this is still work-in-progress and plan to add many more of them in the future to make it a useful tool for understanding scanpath patterns in order to improve a visual stimulus. Also several more interactions might be implemented, however, we already support the data analyst by a list of those based on the categorization by Yi et al. [Yi et al. 2007].

- **Scalable adjacency matrix:** We allow to scale down the adjacency matrix to pixel size in case too many scanpaths have to be compared. This approach provides an overview even for large amounts of comparison values.
- **Scanpath selection and splitting:** Scanpaths can even be selected and split into subscanpaths to analyze in-between subsequences. If the scanpaths are not of equal length, the



(a) (b)

Figure 4: Combining scanpath comparison techniques with matrix reordering algorithms: (a) Jaccard coefficient for comparison plus hierarchical clustering for reordering. (b) Edit distance for comparison plus rank-two ellipse seriation as a spectral approach for the reordering.

data analyst might split the scanpaths into those crossing certain areas of interest.

- **Comparison parameters:** Several additional parameters can be changed, for example, the fixation radius for the set-based comparison algorithms like the Jaccard technique. Moreover, the bounding box comparison can be based on strip-like bounding boxes only computing the overlap of the line-based neighborhood regions of scanpaths giving a more accurate comparison.
- **Matrix aggregations:** After a matrix reordering it may be a useful operation to aggregate rows and columns, for example, to explore different groups of scanpaths or to apply the reordering on more aggregated scanpath comparison data.
- **Color scheme selection:** Each visualization can be represented in a certain color scheme. Our visualization tool provides a repertoire of such schemes from which the user can select the most appropriate ones.
- **Hovering:** Hovering over cells in the matrix shows the similarity value represented in these cells and the participant information belonging to these scanpaths.
- **Smoothed visual attention map:** A box reconstruction filter can be applied several times to the visual attention data with the goal to produce a smoother view on the visual attention data. A similar approach can be applied to the gaze plot computing splines instead of line sequences to get a smoother representation although such operations change the original data and have to be taken with care.
- **Code extensions:** Since many scanpath comparison algorithms exist and even many more matrix reordering techniques we support the user by implementing other solutions while the code can be easily added into the script of our tool.

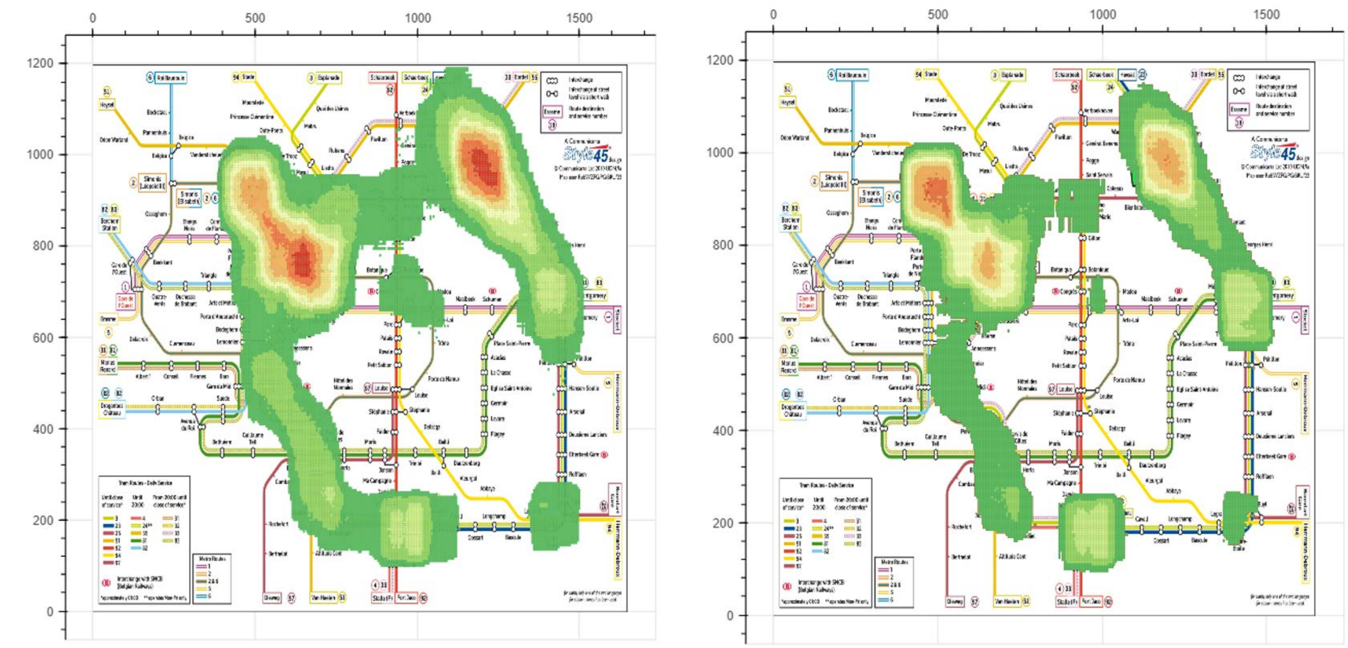
4 APPLICATION EXAMPLE

We applied the clustered eye movement similarity matrices to real-world eye tracking data recorded in a route finding experiment showing people public transport maps while highlighting start and destination stations [Netzel et al. 2017]. The study participants had to find a route and tell the names of the interchange stations while at the same time the eye movements were recorded by a Tobii T60 XL eye tracking device. Although the eye tracking software already provides visualizations in form of visual attention maps and gaze plots, it is pretty hard to identify common visual scanning strategies. Moreover, in typical eye tracking data analysis software not many algorithms for comparison and reordering strategies are integrated.

For illustrative purposes we focus on the public transport map of Brussels in Belgium, but we could have picked any of the 24 metro maps provided by the publicly available dataset from the eye tracking experiment. Each metro map was visually inspected by 20 study participants while looking for suitable routes in the metro map from the highlighted start to the destination station.

4.1 Comparison Results

Table 1 summarizes the 24 color coded matrix results of the application of the 6 comparison techniques and 4 reordering algorithms for the map of Brussels. The first impression that we can get is that the color coded matrices all look a bit differently compared to



(a) (b)

Figure 5: The visual attention map for the public transport map of Brussels in Belgium: (a) without the fixation duration information and (b) with the fixation duration information.

each other. This indicates that it is worth experimenting interactively with different approaches to find a suitable cluster and group structure among the scanpaths that can be further explored in the corresponding visual stimuli in form of a public transport map.

Interactively checking various parameters algorithmically as well as visually while observing their impact on the results in a visual form is a powerful concept in visual analytics with the goal to build, refine, confirm, or reject hypotheses [Keim 2012]. This parameter variation is useful to build a model for the scanning behavior of different participant groups and can help to identify design flaws or problems in the visual stimulus. Only looking at one individual parameter setting oftentimes does not show the full potential of an analysis process and hence, many insights still remain hidden in the data.

Browsing through the individual matrices in an enlarged form while exploring the involved scanpaths and their comparison values can reflect the grouping structure in a more detailed view (see Figure 4). For example, in (a) we can see a scanpath comparison based on the Jaccard coefficient [Levandowsky and Winter 1971] reordered by a hierarchical clustering [Eisen et al. 1998] while in (b) the same scanpath dataset is compared by computing the edit distance [Navarro 2001] and the reordering algorithm is based on the rank-two ellipse seriation [houh Chen 2002].

The approaches in Figure 4 (a) seem to give a better grouping result. Many more subclusters are visible along the diagonal compared to the approach visible in Figure 4 (b). But, for example, in (a) the hierarchical clustering may give too many clusters and it may be more difficult to identify common scanning behavior. For this the data analyst might either choose a different reordering strategy or might switch to another comparison strategy, for example, an edit distance approach. However, if the edit distance is combined with a spectral reordering algorithm, we end up in the situation shown in Figure 4 (b). The formerly fine substructures are now merged together into one large cluster and three outlier scanpaths by the participants numbered p12, p25, and p29. This process indicates how important it can be to experiment with several algorithmic approaches and parameters in order to find out the best parameter setting to find insights in the data, or to confirm or to reject hypotheses as it is a typical analysis strategy in visual analytics.

4.2 Contextual Information

To obtain contextual information from the displayed visual stimulus we can select any kind of cluster in the color coded matrix and show the corresponding scanpaths in an overplotted form in the corresponding public transport map (here Brussels). Figure 5 (a) gives an impression about the general scanpath trend in the Brussels public transport map if we have selected the large cluster identified by computing the edit distance and by ordering the matrix by a rank-two ellipse seriation. We can see that most of the people focus on a similar path building a circular-based shape (with small visual attention gaps). Figure 5 (b) shows the effect of changing visual parameters like including the fixation duration in the visual attention maps. We may even show the gaze plot on top of the visual attention map (see Figure 3 (b)) but in this case this would result in visual clutter.

5 DISCUSSION AND LIMITATIONS

We described a combination of comparison and reordering algorithms supported by interactive visualization with the goal to explore scanpath data recorded in eye tracking experiments. Although we have implemented several of those candidates in combination as a powerful concept we are aware of the fact that some approaches perform much better than others while also different optimal or not-that-optimal results are generated.

5.1 Algorithmic Challenges

We have to deal with algorithmic scalability problems, for example, the number of scanpaths and the length of the scanpaths can become a serious problem for both the comparison algorithms but also for the reordering approaches. However, the lengths of the scanpaths only have an impact on the comparison algorithms but not on the reordering algorithms since for the reordering we only require a matrix-like scheme of real-valued numbers.

We experimented with 6 comparison and 4 reordering algorithms, all of them coming with different runtime complexity classes. However, our experiments showed that the techniques still work interactively for our metro map examples, but we are aware of the fact that the larger the number of scanpaths becomes, the slower the algorithms will work. This can be of particular interest for comparison algorithms based on the longest common subsequence, the Frechet distance, dynamic time warping, and the edit distance, in particular, if additional parameters are adapted like the string alphabet for the edit distance.

The clustering results (if all parameters of comparison algorithms are fixed) typically look differently for each matrix clustering/reordering algorithm. To explore where the clusters are placed in each of the matrices, the data analyst should apply interaction techniques to identify the linkings. However, if the same clustering/reordering is applied to the same comparison data, we will get the same ordered matrix with the clusters at the same positions, a required design criterion.

5.2 Visual Challenges

Visual scalability can become a problem if too many scanpaths have to be compared and be represented in a matrix. But, however, a matrix representation has the benefit that it can be scaled down to pixel size, easily showing thousands of values in an overview. Such an overview is a good concept to guide a data exploration process while still preserving the mental map and the contextual information. On the negative side it is challenging to select individual pixels from the matrix without more advanced interaction concepts.

The extra views in form of visual attention maps show an aggregated representation of the scanpath data and hence, do not show the visual scanning behavior for pairwise scanpath behavior. Only the hot spots of visual attention are visible. For the gaze plots, on the other hand, we soon run into trouble with visual clutter effects caused by many overplotted scanpaths. For this reason we argue that a comparison-based matrix visualization is useful for an overview, to navigate and to filter in the scanpath data. A combination of the algorithmic concepts with the repertoire of interactive visualizations generates a good solution to these scalability issues.

In general, two or more results of comparison algorithms can be combined. They can either be shown by linked visualizations, each showing one matrix of comparison values, or the results of different comparison algorithms might be combined and then these combined comparison values are grouped/clustered. However, aggregating several comparison results leads to a loss of information, hence, we argue for showing the comparison values in separate views and interactively link them together.

5.3 Perceptual Challenges

The color coded and ordered matrices reflect scanpath clusters, assumed that the color scheme is well chosen. Since we do not exactly know which properties a dataset under exploration will have and if the user suffers from color deficiency issues, we provide a repertoire of color schemes. We experimented with several of those as can be seen in the corresponding figures.

A similar problem occurs if several color codings are used in the same view, for example, if the visual stimulus is shown, overplotted with a visual attention map or a gaze plot, or even both. To provide a solution to this problem, the user can adapt each color coding apart from the one given in the visual stimulus.

6 CONCLUSION

In this paper we described a scanpath comparison approach that is not built for just one comparison technique, but benefits from a larger repertoire of candidates like Jaccard and bounding box similarities, but also more complex ones like longest common subsequence, Frechet distance, dynamic time warping, and edit distance. The generated comparison values result in a matrix-like scheme that is hard to interpret if no structure is computed. For this reason, we provide matrix reordering techniques that are hierarchical clustering, optimal-leaf ordering, dimensionality reduction, and a spectral approach. The data analyst can interactively experiment with these algorithms while the visual output is shown in a color coded matrix representation. Typical interactions like filtering and cluster selection are supported to analyze the scanpath data on different properties while the views are linked, for example, with the visual stimulus overplotted with visual attention maps or gaze plots showing the scanpath data in focus. We experimented with real-world data from a formerly conducted eye tracking study investigating route finding tasks in public transport maps. For future work we plan to integrate more comparison methods and matrix reordering techniques while we also plan to test our approach with dynamic stimuli like videos, animations, or interactive user interfaces. Moreover, we would also like to conduct a user experiment with experts to investigate the usefulness of our approach.

ACKNOWLEDGMENTS

This research was partially supported by NSF grant IIS 1527200 and MSIT, Korea, under the ICT Consilience Creative program (IITP-2019-H8601-15-1011) supervised by the IITP. We would like to thank Robin Woods from Communicarta Ltd. for providing the public transport maps.

REFERENCES

Helmut Alt and Michael Godau. 1995. Computing the Frechet Distance Between Two Polygonal Curves. *International Journal of Computational Geometry and Applications* 5, 1-2 (1995), 75–91.

Gennady Andrienko, Natalia Andrienko, Michael Burch, and Daniel Weiskopf. 2012. Visual Analytics Methodology for Eye Movement Studies. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2889–2898.

Ziv Bar-Joseph, David K. Gifford, and Tommi S. Jaakkola. 2001. Fast Optimal Leaf Ordering for Hierarchical Clustering. *Bioinformatics* 17 (2001), 22–29.

Michael Behrisch, Benjamin Bach, Nathalie Henry Riche, Tobias Schreck, and Jean-Daniel Fekete. 2016. Matrix Reordering Methods for Table and Network Visualization. *Computer Graphics Forum* 35, 3 (2016), 693–716.

Tanja Blascheck, Michael Burch, Michael Raschke, and Daniel Weiskopf. 2015. Challenges and Perspectives in Big Eye-Movement Data Visual Analytics. In *Proceedings of the 1st International Symposium on Big Data Visual Analytics*. 17–24.

Tanja Blascheck, Kuno Kurzhals, Michael Raschke, Michael Burch, Daniel Weiskopf, and Thomas Ertl. 2017. Visualization of Eye Tracking Data: A Taxonomy and Survey. *Computer Graphics Forum* 36, 8 (2017), 260–284.

Agnieszka Bojko. 2009. Informative or Misleading? Heatmaps Deconstructed. In *Human-Computer Interaction – INTERACT*. Springer, 30–39.

Michael Burch. 2016. Time-Preserving Visual Attention Maps. In *Proceedings of Conference on Intelligent Decision Technologies*. 273–283.

Michael Burch, Gennady L. Andrienko, Natalia V. Andrienko, Markus Höferlin, Michael Raschke, and Daniel Weiskopf. 2013. Visual Task Solution Strategies in Tree Diagrams. In *Proceedings of IEEE Pacific Visualization Symposium*. 169–176.

Andrew T. Duchowski. 2003. *Eye Tracking Methodology - Theory and Practice*. Springer.

Michael B. Eisen, Paul Spellman, Patrick O. Brown, and David Botstein. 1998. Cluster Analysis and Display of Genome Wide Expression Patterns. In *Proceedings of the National Academy of Sciences*, Vol. 95. 14863–14868.

Joseph H. Goldberg and Jonathan I. Helfman. 2010. Visual Scanpath Representation. In *Proceedings of the Symposium on Eye-Tracking Research and Applications (ETRA)*. 203–210.

Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost van de Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press.

Chun houh Chen. 2002. Generalized Association Plots: Information Visualization Via Iteratively Generated Correlation Matrices. *Statistica Sinica* 12, 1 (2002), 7–29.

Daniel A. Keim. 2012. Solving Problems with Visual Analytics: Challenges and Applications. In *Proceedings of Machine Learning and Knowledge Discovery in Databases - European Conference*. 5–6.

Ayush Kumar, Michael Burch, Irma van den Brand, Laurens Castelijns, Frank Ritchi, Freek Rooks, Hero de Smeth, Neil Timmermans, and Klaus Mueller. 2018a. Eye Tracking for Exploring Visual Communication Differences. (2018).

Ayush Kumar, Rudolf Netzel, Michael Burch, Daniel Weiskopf, and Klaus Mueller. 2018b. Visual Multi-Metric Grouping of Eye Tracking Data. *Journal on Eye Movement Research* 10, 5 (2018), 1–17.

Ayush Kumar, Rudolf Netzel, Michael Burch, Daniel Weiskopf, and Klaus Mueller2. 2016. Multi-similarity matrices of eye movement data. In *2016 IEEE Second Workshop on Eye Tracking and Visualization (ETVIS)*. 26–30. https://doi.org/10.1109/ETVIS.2016.7851161

Kuno Kurzhals, Marcel Hlawatsch, Michael Burch, and Daniel Weiskopf. 2016a. Fixation-Image Charts. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, ETRA*. 11–18.

Kuno Kurzhals, Marcel Hlawatsch, Florian Heimerl, Michael Burch, Thomas Ertl, and Daniel Weiskopf. 2016b. Gaze Stripes: Image-Based Visualization of Eye Tracking Data. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 1005–1014.

Michael Levandowsky and David Winter. 1971. Distance between sets. *Nature* 234, 5 (1971), 34–35.

David Maier. 1978. The Complexity of Some Problems on Subsequences and Supersequences. *J. ACM* 25, 2 (1978), 322–336.

Gonzalo Navarro. 2001. A guided tour to approximate string matching. In *ACM Computing Surveys*, Vol. 33. 31–88.

Rudolf Netzel, Bettina Ohlhausen, Kuno Kurzhals, Robin Woods, Michael Burch, and Daniel Weiskopf. 2017. User Performance and Reading Strategies for Metro Maps: An Eye Tracking Study. *Spatial Cognition & Computation* 17, 1–2 (2017), 39–64.

Michael Raschke, Xuemei Chen, and Thomas Ertl. 2012. Parallel Scan-path Visualization. In *Proceedings of the 2012 Symposium on Eye-Tracking Research and Applications, ETRA*. 165–168.

Ruth Rosenholtz, Yuanzhen Li, Jonathan Mansfield, and Zhenlan Jin. 2005. Feature Congestion: A Measure of Display Clutter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 761–770.

Diego Furtado Silva and Gustavo E. A. P. A. Batista. 2016. Speeding Up All-Pairwise Dynamic Time Warping Matrix Calculation. In *Proceedings of the 2016 SIAM International Conference on Data Mining*. 837–845.

Oleg Spakov and Darius Miniotas. 2007. Visualization of Eye Gaze Data Using Heat Maps. *Electronics and Electrical Engineering* 2, 74 (2007), 55–58.

Ian Spence and Jed Graef. 1974. The Determination of the Underlying Dimensionality of an Empirically Obtained Matrix of Proximities. *Multivariate Behavioral Research* 9, 3 (1974), 331–341.

Godfried Toussaint. 1983. Solving Geometric Problems with the Rotating Calipers. (1983).

Alfred L. Yarbus. 1967. *Eye Movements and Vision (Translated from Russian by Basil Haigh. Original Russian edition published in Moscow in 1965.)*. New York: Plenum Press.

Ji Soo Yi, Youn ah Kang, John T. Stasko, and Julie A. Jacko. 2007. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1224–1231.