Storyline Visualizations of Eye Tracking of Movie Viewing

J. Timothy Balint* George Mason University

Dustin Arendt† Pacific Northwest National Laboratory

Leslie M. Blaha‡ Pacific Northwest National Laboratory

Abstract

Storyline visualization is a technique that captures the spatio-temporal characteristics of individual entities and simultaneously illustrates emerging group behaviors. We developed a storyline visualization leveraging dynamic time warping to parse and cluster eye tracking sequences. Visualization of the results captures the similarities and differences across a group of observers performing a common task. We applied our storyline approach to gaze patterns of people watching dynamic movie clips. We use these to illustrate variations in the spatio-temporal patterns of observers as captured by different data encoding techniques. We illustrate that storylines further aid in the identification of modal patterns and noteworthy individual differences within a corpus of eye tracking data.

Keywords: eye tracking, dynamic time warping, storyline visualization, layout algorithm, video stimuli

Index Terms: I.3.8 [Computing Methodologies]: COMPUTER GRAPHICS—Applications; G.3 [Mathematics of computing]: Probability and statistics—Time series analysis

1 Introduction

Dynamic stimuli and task environments challenge the limits of current eye tracking analytic techniques. When watching moving scenes or movies, people show broad individual differences in their eye movement responses to visual events and dynamic content. If people are also moving or interacting with visual information (e.g., *in situ* grocery shopping), then the sequences of their own behaviors will further complicate interpretation of the visual events and eye movement patterns. We need visualization tools that enable the identification of individual and group patterns as well as meaningful techniques for aligning variable spatio-temporal characteristics of scanpaths for direct comparison. This will enable researchers to identify observer subgroups utilizing common task strategies and to find events or stimulus characteristics that trigger common or deviating gaze patterns among the observers. We propose that storyline visualization effectively supports these tasks.

It is often claimed, and sometimes tested [10], that people tend to look at the same portion of the screen when watching video clips or sequences. Evaluations of such claims have entailed spatial analysis of fixation positions at selected moments in time with observer data collapsed into a distribution of positions; when a majority of the distribution fall into a common area of interest (AOI), researchers claim people are looking at the same place. Indeed, this assumption is critical to many video compression and enhancement techniques which emphasize the AOI to which a majority of observers attended [10, 27]. There is strong evidence that many common fixation patterns are driven by visual salience and factors that attract exogenous attention [12, 13, 14, 27], which can be leveraged for video processing targeting mean or modal observer behavior.

*e-mail: jbalint2@masonlive.gmu.edu

†e-mail:dustin.arendt@pnnl.gov

‡e-mail:leslie.blaha@pnnl.gov

But do people look at the same places at the same time or in the same sequence? Analytics leveraging spatial aggregation provide only minimal insight into the dynamic sequences of viewing events that constitute the total task behaviors. In video viewing, observers may take different strategies when multiple highly salient objects are in their fields of view. Viewer age, gender, prior experience, and content expectations all shape gaze patterns [10]. Consequently, the one-AOI-fits-all assumption of video compression/enhancement techniques will be less successful for the portion of observers deviating from the mean or modal scanpaths. Behaviors that should be treated as a different cognitive task strategy would be treated as noise and potentially discarded altogether. We hypothesize that storyline visualization, together with the appropriate analytics, is well suited to enable the characterization of spatio-temporal eye gaze patterns. They provide a view of individual observers together with aggregations across observers, aiding the identification of behavioral commonalities in both the space and time dimensions. In this paper, we use dynamic time warping and storyline visualization to accomplish two objectives: (1) demonstrate an analytics pipeline for processing time series data and optimizing storyline layout for eye tracking data, and (2) illustrate the effect that data encoding method has on the resultant visualization.

1.1 Related Work

Eye tracking visualizations for dynamic viewing can be classified as temporal, spatial, or spatio-temporal, depending on whether the time dimension is explicitly represented [1]. Temporal plots explicitly represent time as one of the visualization dimensions. Often these collapse the spatial dimension into AOI labels, illustrating dwell time within each AOI and transitions between them. Detailed spatial position information is lost, as in scarf plots [18, 31] and AOI timelines [5]. Spatial visualizations, such as attention maps [4], saccade plots [6], and transition diagrams [3], use a two-dimensional plot or image overlay to capture the two-dimensional position of the eyes. Transitions between AOIs or fixations may be explicitly represented with lines between points, but the temporal information about the scanpath is lost. For our interest in gaze sequences, we seek techniques that capture both spatial and temporal information. Space-time cubes [19, 20] and similar spatio-temporal visualizations attempted to capture this information in 3-D plots. But these suffer from interpretability challenges when rendered on 2-D displays and are prone to clutter and overplotting when individual participant data are combined into a single visualization.

Storyline visualization shows how entities interact over space and time [17, 21, 26, 30, 33]. Storyline visualizations encode time on the horizontal axis and interactions on the vertical axis. When two entities are interacting, their storylines are drawn close together; otherwise, they are drawn apart. Interactions are defined as entities having the same state at the same time. Some visualizations of eye tracking data are reminiscent of storyline visualizations. Grindinger et al. [11] plotted the y-coordinate of the gaze point as a function of time. Löwe et al. [22] used multidimensional scaling to project multiple users' fields of view into one dimension, to be plotted against time. Raiha et al. [28] and Raschke et al. [29] plotted AOIs versus time for multiple users. These visualizations could benefit from storyline layout optimizations to reduce artifacts, such as line crossings, that result from arbitrary line placement. However, similar to storyline visualizations, these techniques are less effective on long time series or suffer overplotting of data recorded at a high sampling rate.

Transformations on raw eye tracking data are used to reduce visual clutter and increase the effectiveness of time-interaction visualizations. Example transformation approaches include aggregating raw eye coordinates over frames with a Gaussian Mixture Model [24], spatial and semantic hierarchical visualizations [2], and Moving Clusters aggregation over time windows [15]. Thompson and Torabi [34] demonstrated the latter technique on static image eye tracking data, but they found the ability to maintain consistent clusters degraded over time. All these techniques are constrained by the number of observers that can be displayed at a given time. We note that many such transformation techniques aggregate the observers, rendering it difficult to examine individual behavior over time.

2 The Storyline Approach

We present a novel application of storyline visualization and dynamic time warping to eye tracking data. Additionally, we developed a new storyline visualization design and layout algorithm to support this specific analysis approach. Rather than focusing on showing raw (x,y) coordinates or AOIs, we show how users' eye gaze patterns change over time. We define a pattern as time sequence of eye movements. We use storyline visualization to convey which users' patterns are very similar within each time window. Our approach can be divided into analytic and visualization stages. In the analytic stage, the interactions between users' patterns are found, using the process described in Section 2.1. In the visualization stage, the layout of the storylines is optimized to improve the usability and aesthetic quality of the drawing. A visual representation of our approach can be seen in Figure 1.

2.1 Eye Gaze Time-series Dynamic Time Warping

Our analytics stage transforms raw eye gaze data into behavior-labeled windows across users. This is accomplished by the following four steps, illustrated in Figure 1.

1. Clean data by retaining fixation and saccade samples, removing eye blinks and recording errors.
2. Divide time series into time windows.
3. Measure warped spatio-temporal distance across all $(observer, window)$ pairs.
4. Cluster the distance matrix to group similar eye gaze patterns.

In the first step, we represent each fixation or saccade by an (x,y) coordinate or AOI label, with the choice being a parameter chosen by an analyst. We note that this step may leverage AOI labels or event flags generated by eye tracking software associated with specific hardware; it can also use the fixation/AOI label output from a number of fixation identification algorithms [32]. Eye tracking software may also give error event flags (e.g., blinks), which we remove before processing. If AOI labels are used, we include an additional "outlier" label for any point outside the pre-determined, semantically interesting AOIs. This is so all points have a label for our analysis. Depending on the analytic goals, these outliers may also be removed. Data labeling and cleaning greatly reduces the the quantity of data to be processed, and the data is now represented as a sequence of coordinates or labels instead of as an evenly sampled time series. In the second step we choose a meaningful time window length by which we coarsen the data. The selection of the window parameter is data/experiment dependent. It should be chosen by the analyst to capture relevant behaviors at the appropriate scale (or multiple values chosen for meaningful comparisons).

Step three is to measure the distance across all (e_i^s, e_j^t) pairs, for time windows s,t and observers i,j, where e_i^s is the fixation sequence for observer i during time window s. Measuring distance between arbitrary sequences is non-trivial. We rely on *dynamic time warping* (DTW; [25]) to provide meaningful distances, even when participants exhibit variations in the timing of their fixations. DTW creates a transformation between two time signals, and provides a distance cost for the transformation. For AOI sequences, our underlying distance metric is Hamming distance, which provides a generalized edit distance between the sequences of AOI labels. For (x,y) coordinates we use Euclidean distance for the total cost between two 2-D signals. This produces a distance matrix between all observer pairs at each time window over the entire time series.

We input this distance matrix into an off-the-shelf clustering algorithm to find sub-groups. We chose the k-medoids clustering algorithm [16] because it is an intuitive variant of k-means that operates directly on the distance matrix (instead of a feature matrix, which is not available in this case). The clustering algorithm assigns a label to each $(observer, window)$ pair. If two $(observer, window)$ pairs have the same label, they can be assumed to have similar spatio-temporal patterns within this time window. In the storyline visualization, they will be drawn close together.

A key benefit of this approach is that it allows for the visual complexity (i.e., the number of visual elements represented) to depend only on the number of time windows and observers. This approach scales up because the visualization is independent of the temporal frequency that the underlying eye gaze data was captured. Figure 1 illustrates this complete parsing process.

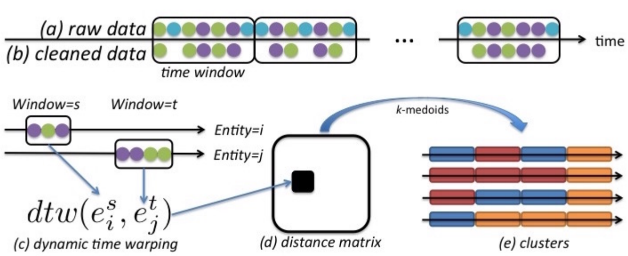


Figure 1: Eye Gaze Time-series Analytics Pipeline. (a) Start with raw gaze data, and (b) omit non-fixation or non-saccade data point (x,y) coordinates or AOIs, and segment time series by time windows. (c) Calculate a warped distance between all pairs of fixation sequences across time windows and entities, and (d) collect these pairwise distances into a square distance matrix. (e) Assign cluster labels to each entity at each time window such that sequences with similar fixation patterns belong to the the same cluster.

2.2 Storyline Visualization Algorithm

We use storyline visualization to show, at a coarse level, how eye gaze patterns change over time across a set of observers viewing identical movie clips. However, determining the vertical placement of each entity over time in a way that produces usable visualizations is a challenging problem, with many different designs and algorithms proposed in the visualization literature. Pilot usability studies conducted within our organization suggest that analysts prefer storyline visualizations where the vertical dimension has a consistent meaning over time. Therefore, our storyline visualization builds on the design proposed by Reda et al. [30] where clusters, which are sequences of similar fixation patterns, are assigned to different vertical layers. With this design constraint, we implemented a custom storyline layout algorithm that re-arranged the cluster layers and storylines to decrease the number of line crossings in the visualization. This is intended to improve the usability and aesthetic quality of the visualization.

Our storyline layout optimization has two stages: (1) determine a good ordering of the clusters, and (2) determine a good ordering of the storylines. We refer to this ordering as "good" because it decreases the number of crossings from an arbitrary starting state. However, it does not guarantee that the number of crossings is minimum (i.e., cannot be improved). In the first stage, we build a cluster graph $G_c = (V_c, E_c)$ that models the transitions between clusters. The weight of an edge $(u,v) \in E_c$ is the number of times any observer is exhibiting pattern/cluster u at time t and then pattern/cluster v at time $t+1$. Starting from an arbitrary initial permutation π_0 of V_c, we iteratively improve this ordering using the barycenter heuristic [8, 23], which is likely to put edges with larger weights closer together in π_i after the i^{th} iteration.

In the next step, we compute an ordering of the storylines that is dependent on the ordering of clusters found in the previous step. We construct a directed preference graph $G_p = (V_p, E_p)$ that models the preference for one storyline to be above or below another storyline. We consider all pairs of storylines (u,v) belonging to a given cluster c at time t. If u's cluster is above v's cluster at time $t-1$ and $t+1$, then we say that u *prefers* to be above v. A directed edge (u,v) remains in E_p if the total preference of u over v is greater than 0 considering all time windows. Finding the feedback arc set (FAS) of G_p leads directly to an ordering of V_p that best respects the pairwise preferences of the storylines. While finding the FAS is NP-Complete, we rely on a fast approximation algorithm that leads to satisfactory results [7]. We sort storylines within the same time window and cluster according to the FAS order to determine their actual vertical positions.

3 Application

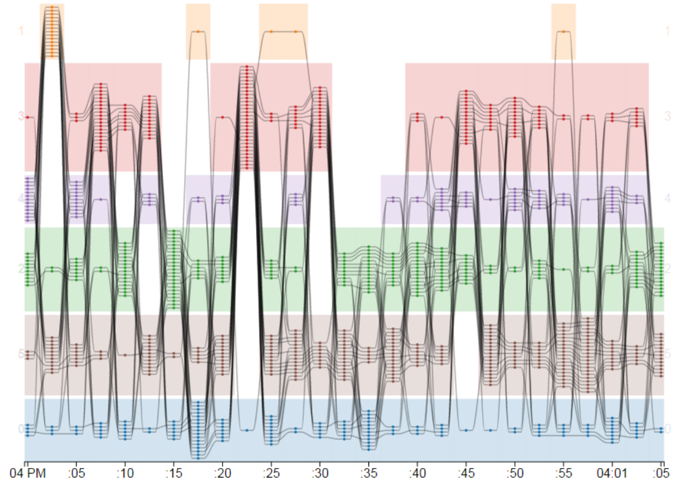
We demonstrate our storyline approach on gaze data recorded during passive viewing of two videos. We used publicly available datasets from the DIEM Project [24]. These data sets contain (x,y) coordinates, pupil dilation, and event flags (fixation, saccade, blink, etc.) recorded at 30 frames per second. The goal of our application is to show examples of the individual and sub-group patterns highlighted by the storyline approach. We also illustrate how algorithmic choices influence the resulting storylines.

In the storyline visualizations, time is on the horizontal dimension, with the points demarcating the sample times from the coarsening process. For this demonstration, the gaze data was coarsened to 2.5 second intervals. The colored groups are the k-medoids clusters. Each curve running between the clusters represents a single observer's data. Curves within a cluster are people exhibiting similar spatio-temporal patterns in that window. More lines clustering together in a time window indicate that something is occurring, externally or cognitively, that triggered a common behavior in all those observers. Splits of the storylines between clusters indicate that subgroups of people are exhibiting different gaze behaviors in that time window. Each curve shifting between clusters indicates a person changing his/her spatio-temporal pattern.

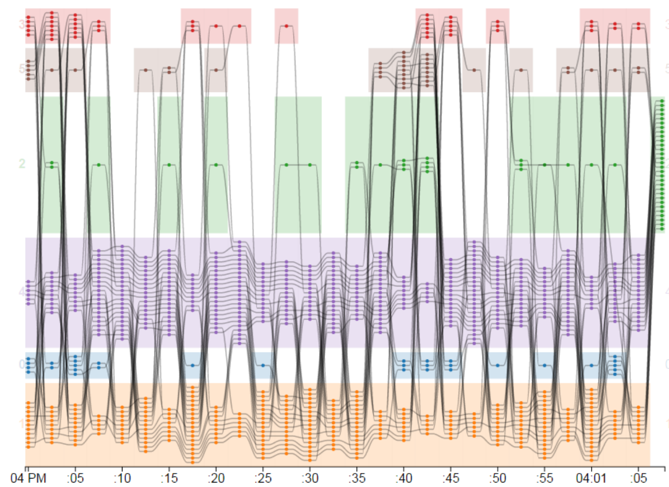
3.1 *The Simpsons* Dataset

To illustrate the storyline differences based on data encoding, we examined viewing behavior of the opening credits from the TV show "The Simpsons." This video sequence uses camera movement to guide observers. Often there is a central focal item in the camera view, which the camera follows. This central object is surrounded by other events and salient moving objects. Camera scene cuts re-focus the observers to new focal objects. Figure 2 illustrates the resulting storylines, with two versions showing how the choice of data type affects cluster assignments. In Figure 2a, the raw (x,y) coordinates are used to inform the clustering. Figure 2b leverages AOI labels. We generate AOI labels using DBSCAN [9] on all (x,y) coordinates, treating each coordinate independently.

The choice to use the raw gaze coordinates or AOI labels determines the distance metric used with DTW, and therefore affects the overall clustering shown. Specifically the use of AOI labels in Figure 2b produced storylines with two dominant clusters across the time series. In comparison, raw (x,y) coordinate storylines in Figure 2a show observers more evenly spread between five clusters. In



(a) Storylines from (x,y)-coordinate Labels



(b) Storylines from AOI Labels

Figure 2: Two storyline visualizations of eye gaze patterns during the "The Simpsons" opening credits. (a) Storylines from raw (x,y) coordinates. (b) Storylines computed from AOI label time series. See text for element descriptions. Note the timestamps in seconds are relative to 0:00 time from the start of the video, arbitrarily labeled here as 4pm.

order to understand where which AOIs are captured in these clusters, we must look at the data, potentially overlaying the AOIs on the video itself. Returning to the data, the dominant AOI clusters almost exclusively contain the outlier label, meaning that observers in these clusters are consistently looking at part of the video that was not previously considered semantically interesting. One implication is that the AOI labeling may be incomplete. As we used a large threshold for the minimum number of points to create an AOI label, we recognize that changing the parameter should affect the AOI labels and identify fewer points as outliers.

Storyline visualization also quickly shows the distribution of observers between different spatio-temporal patterns. Times when almost all participants follow the same behavior appear as a single dominant cluster, as in Figure 2a at 0:22.5 s and Figure 2b at 1:07.5 s. Times when participants are distributed between several patterns appear as more evenly spread across multiple clusters, such as Figure 2b at 0:40.0 s. The current storyline does not display which (x,y) coordinates or AOI labels constitute the clusters, as the DTW process combine spatial and time data instead of individual location points. However, storylines do highlight that these time windows should be targeted for deeper analysis.

3.2 *Fifty Person Questionnaire–Brooklyn* Dataset

We also applied the storyline approach to the "Fifty Person Questionnaire–Brooklyn" video. In contrast to the motion-laden "The Simpsons" video, this video uses a static camera view with either one or two people speaking to the camera. The close-up view often obscures any background activity or objects. The speaker may be left/right/center within the scene. Scene cuts switch between speakers over the course of the film. The storyline result can be seen in Figure 3. In each time window, this dataset displays a strong dominant cluster, which changes over time.

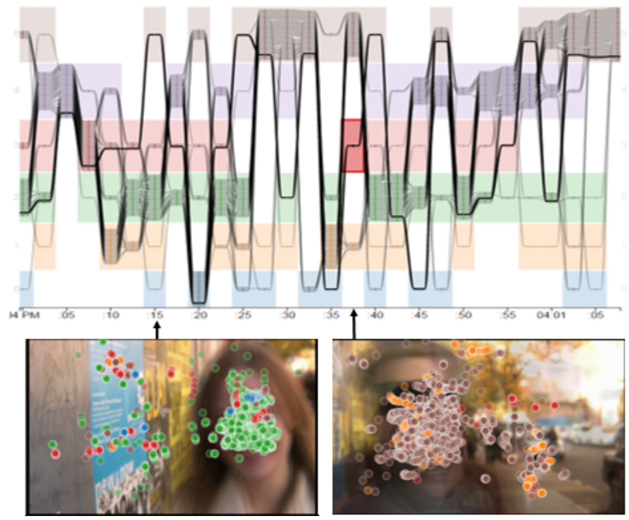
We selected two time windows to dig deeper into the differences between clusters, seen in Figure 3. Both selected time windows have a single dominant cluster, with only a few participants exhibiting different gaze behaviors. One observer (bolded curves in Figure 3a) is in a very sparse grouping for both selected time windows. As storylines only show affinity between participants, we can only know that the DTW process isolated this observer as showing a divergent gaze pattern. To understand the differences, we leverage additional visualizations of key time windows. The average screenshots for each time window, below the storyline, show the (x,y) fixation coordinates for all participants; we can locate the outlier observer's colors (Left: brown; Right: red) to see that s/he is fixating on points outside the main clusters. Figures 3b and 3c show the x position over the selected time windows, with our observer of interest boxed. In Figure 3b, the observer shows a left-ward movement with long fixations; in Figure 3c, the observer shows fixations that deviate from the patterns of the other observers. By combining the storyline with additional visualizations, the storyline facilitates the finding of interesting movement patterns within a set of observers, potentially indicating behavioral outliers, for a given time window.

4 Conclusion

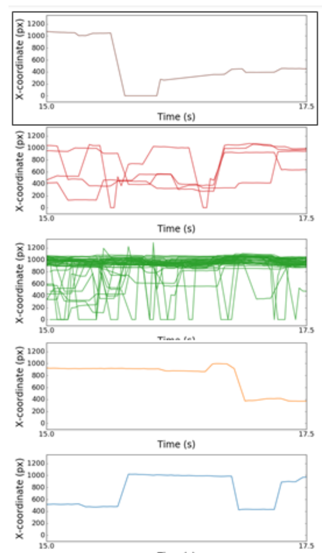
Storyline visualization leveraging DTW is a promising approach enabling the visualization of multiple observers' behaviors when viewing dynamic stimuli. One appeal of this approach is that a consistent visualization technique may be applied to any means of extracting AOI labels, as long as the DTW distance metric is adapted appropriately. As summarized by Salvucci and Golberg [32], fixation identification techniques vary in their use of spatial or temporal data characteristics. According to their taxonomy, each offers trade-offs in implementation difficulty, accuracy, and robustness. Leveraging storylines, all techniques, as well as parametric variations, may be captured in a common visualization technique for direct side-by-side comparison in terms of the spatio-temporal patterns present in the extracted fixation sequences. Novel techniques may further be added to this common visual language and directly compared to standard techniques such that qualitative differences in approaches will be immediately apparent.

Storyline visualizations show whether eye tracking observers are following the same or similar gaze patterns during a given time window. They allow analysts to quickly determine if participants are deviating from the modal behavior patterns. However, the difference in the patterns cannot be determined by storyline visualizations alone. In the application in Figure 3, we leveraged additional plots to examine detailed spatio-temporal characteristics within time windows of interest from the storyline. Storyline visualizations comprise a strong foundation for a powerful tool suite for dynamic viewing eye tracking data.

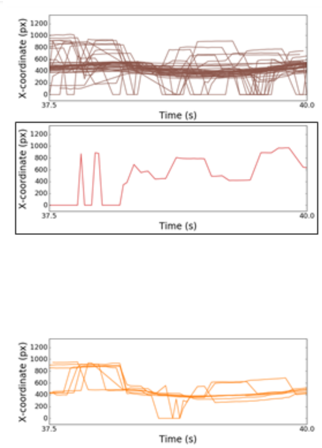
Beyond passively viewing movies, the analysis of eye tracking data in dynamic tasks offers additional challenges with which storylines may be able to assist. Dynamic tasks may include random (sequences of) events to which an observer must respond. People may adopt different strategies, which may be classified according to their performance storylines. Shifts in strategy, or changes due to training/experience, may be tracked as individuals shift their storyline membership. Further, storylines offer a visualization approach to identifying subsequences of dynamic events that prompt similar behaviors among observers.



(a) Storylines from (x,y) coordinate labels



(b) x position, 0:15 s time window



(c) x position, 0:37.5 s time window

Figure 3: (a) Storyline visualizations processed using the (x,y) coordinates time series for "Fifty Person Questionnaire–Brooklyn" dataset. Below the storyline is the averaged images for two selected time windows (indicated with arrows). Fixations are overlaid on the images in colors that correspond with the cluster membership for each point. (b-c) The *horizontal eye movements* against *time* of all participants for each 2.5 s time window, centered at the indicated time stamp. The color of each participant corresponds to the storyline cluster in the given window, and participants of the same cluster are shown in the same plot.

Acknowledgements

The research described in this document was sponsored the U.S. Department of Energy (DOE) through the Analysis in Motion Initiative at Pacific Northwest National Laboratory. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

References

[1] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. State-of-the-art of visualization for eye tracking data. In R. Borgo, R. R. Maciejewski, and I. Viola, editors, *Proceedings of Eurographics Conference on Visualization (EuroVis)*, 2014.

[2] T. Blascheck, K. Kurzhals, M. Raschke, S. Strohmaier, D. Weiskopf, and T. Ertl. AOI hierarchies for visual exploration of fixation sequences. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, pages 111–118. ACM, 2016.

[3] T. Blascheck, R. M., and T. Ertl. Circular heat map transition diagram. In *Proceedings of the 2013 Conference on Eye Tracking South Africa (ETSA'13)*, pages 58–61, 2013.

[4] A. A. Bojko. Informative or misleading? Heatmaps deconstructed. In *International Conference on Human-Computer Interaction*, pages 30–39. Springer, 2009.

[5] M. Burch, A. Kull, and D. Weiskopf. AOI rivers for visualizing dynamic eye gaze frequencies. *Computer Graphics Forum*, 32:281–290, 2013.

[6] M. Burch, H. Schmauder, M. Raschke, and W. D. Saccade plots. In *Proceedings of the 2014 Symposium on Eye Tracking Research & Applications (ETRA?14)*, pages 307–310, 2014.

[7] P. Eades, X. Lin, and W. F. Smyth. A fast and effective heuristic for the feedback arc set problem. *Information Processing Letters*, 47(6):319–323, 1993.

[8] P. Eades and N. C. Wormald. Edge crossings in drawings of bipartite graphs. *Algorithmica*, 11(4):379–403, 1994.

[9] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. volume 96, pages 226–231, 1996.

[10] R. B. Goldstein, R. L. Woods, and E. Peli. Where people look when watching movies: Do all viewers look at the same place? *Computers in Biology and Medicine*, 37(7):957–964, 2007.

[11] T. Grindinger, A. T. Duchowski, and M. Sawyer. Group-wise similarity and classification of aggregate scanpaths. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pages 101–104. ACM, 2010.

[12] L. Itti. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing*, 13(10):1304–1318, 2004.

[13] L. Itti and P. Baldi. A principled approach to detecting surprising events in video. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 631–637. IEEE, 2005.

[14] L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40:1489–1506, 2000.

[15] P. Kalnis, N. Mamoulis, and S. Bakiras. On Discovering Moving Clusters in Spatio-Temporal Data. In *International Symposium on Spatial and Temporal Databases*, pages 364–381. Springer, 2005.

[16] L. Kaufman and P. Rousseeuw. *Clustering by means of medoids*. Statistical Data Analysis Based on the L1 Norm. North-Holland, 1987.

[17] N. W. Kim, S. K. Card, and J. Heer. Tracing genealogical data with timenets. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 241–248. ACM, 2010.

[18] K. Kurzhals, F. Heimerl, and D. Weiskopf. Iseecube: Visual analysis of gaze data for video. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 43–50. ACM, 2014.

[19] K. Kurzhals and D. Weiskopf. Space-time visual analytics of eye-tracking data for dynamic stimuli. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2129–2138, 2013.

[20] X. Li, A. Çöltekin, and M.-J. Kraak. Visual exploration of eye movement data using the space-time-cube. In *Geographic Information Science*, pages 295–309. Springer, 2010.

[21] S. Liu, Y. Wu, E. Wei, M. Liu, and Y. Liu. Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2436–2445, 2013.

[22] T. Löwe, M. Stengel, E.-C. Förster, S. Grogorick, and M. Magnor. Visualization and Analysis of Head Movement and Gaze Data for Immersive Video in Head-mounted Displays. In *Proceedings of the First Workshop on Visualizing Eye Tracking Data (ETVIS)*, 2015.

[23] E. Mäkinen and H. Siirtola. The barycenter heuristic and the reorderable matrix. *Informatica (Slovenia)*, 29(3):357–364, 2005.

[24] P. K. Mital, T. J. Smith, R. L. Hill, and J. M. Henderson. Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, 3(1):5–24, 2011.

[25] M. Müller. Dynamic Time Warping. In *Information Retrieval for Music and Motion*, pages 69–84. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[26] M. Ogawa and K.-L. Ma. Software evolution storylines. In *Proceedings of the 5th International Symposium on Software Visualization*, pages 35–42. ACM, 2010.

[27] W. M. Osberger and A. M. Rohaly. Automatic detection of regions of interest in complex video sequences. In *Photonics West 2001-Electronic Imaging*, pages 361–372. International Society for Optics and Photonics, 2001.

[28] K.-J. Räihä, A. Aula, P. Majaranta, H. Rantala, and K. Koivunen. Static visualization of temporal eye-tracking data. In *IFIP Conference on Human-Computer Interaction*, pages 946–949. Springer, 2005.

[29] M. Raschke, X. Chen, and T. Ertl. Parallel scan-path visualization. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 165–168. ACM, 2012.

[30] K. Reda, C. Tantipathananandh, A. Johnson, J. Leigh, and T. Berger-Wolf. Visualizing the evolution of community structures in dynamic social networks. In *Computer Graphics Forum*, volume 30, pages 1061–1070, 2011.

[31] D. C. Richardson and R. Dale. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6):1045–1060, 2005.

[32] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, pages 71–78. ACM, 2000.

[33] Y. Tanahashi, C.-H. Hsueh, and K.-L. Ma. An efficient framework for generating storyline visualizations from streaming data. *IEEE Transactions on Visualization and Computer Graphics*, 21(6):730–742, 2015.

[34] J. Thompson and T. Torabi. Visualising Moving Clusters using Cluster Flow Diagrams. In *Proceedings of the 38th Australasian Computer Science Conference (ACSC 2015)*, volume 27, pages 37–45, 2015.