Thirteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2019)

Analysis of Long Duration Eye-Tracking Experiments in a Remote Tower Environment

Prithiviraj Muthumanickam
Department of Science and Technology,
Linköping University, Norrköping, Sweden
prithiviraj.muthumanickam@liu.se

Aida Nordman
Department of Science and Technology,
Linköping University, Norrköping, Sweden
aida.nordman@liu.se

Lothar Meyer
Air Navigation Services of Sweden,
LFV, Norrköping, Sweden
lothar.meyer@lfv.se

Supathida Boonsong
Air Navigation Services of Sweden,
LFV, Norrköping, Sweden
supathida.boonsong@lfv.se

Jonas Lundberg
Department of Science and Technology,
Linköping University, Norrköping, Sweden
jonas.lundberg@liu.se

Matthew Cooper
Department of Science and Technology,
Linköping University, Norrköping, Sweden
matthew.cooper@liu.se

Abstract—**Eye-Tracking experiments have proven to be of great assistance in understanding human computer interaction across many fields. Most eye-tracking experiments are non-intrusive and so do not affect the behaviour of the subject. Such experiments usually last for just a few minutes and so the spatio-temporal data generated by the eye-tracker is quite easy to analyze using simple visualization techniques such as heat maps and animation. Eye tracking experiments in air traffic control, or maritime or driving simulators can, however, last for several hours and the analysis of such long duration data becomes much more complex. We have developed an analysis pipeline, where we identify visual spatial areas of attention over a user interface using clustering and hierarchical cluster merging techniques. We have tested this technique on eye tracking datasets generated by air traffic controllers working with Swedish air navigation services, where each eye tracking experiment lasted for $\sim$90 minutes. We found that our method is interactive and effective in identification of interesting patterns of visual attention that would have been very difficult to locate using manual analysis.**

Keywords—**Remote tower, Eye tracking, Spatio-temporal clustering**

I. Introduction

Eye tracking technology has been available for several decades but in the last 20 years has become much more widely used due to the drop in price of the underlying technologies, such as high speed digital cameras and capacious high speed storage in computers. As a result eye-tracking studies are now very common across a wide range of areas from user behaviour analysis to system design and even in film and advertising studies [1]. Most of these studies, however, look at quite short duration scenarios lasting, at most, a few minutes. Even in the case of films, where the duration of the viewing may be hours, the studies can focus on each scene independently and so the duration of each study lasts for only a few seconds. In such a context, the analysis of the eye-tracking data (typically eye gaze position in the scene sampled as x,y coordinates at, often, hundreds of samples per second) is not very challenging. Analysts typically use simple visualization techniques such as heat maps to identify regions of frequent gaze, and linear movements (saccades) to study patterns of eye movement across the display. Combined with interactive windowing of the total duration of the study and limited animation techniques, user behaviour over time can be identified easily.

Recently we have undertaken a number of eye tracking studies, in collaboration with colleagues at Linköping University and Luftfartsverket (Sweden's National Air Traffic Agency), where the eye gaze points for the whole working period of Air Traffic Controllers (ATCOs) are recorded, with each recording typically lasting for around 90 minutes. The studies were primarily aimed at the training process of ATCOs where there is a desire to understand the activities of trainees and to improve the effectiveness of the training process. In such a long scenario of 90 minutes, it becomes much more complex to analyse the eye-tracking data, since the patterns of behaviour are both very complex and can change significantly over a longer period. Variations in the working scenario, such as traffic flows, weather conditions, or exterior lighting conditions, combine to affect the way in which a controller operates. Analysing this data manually can require days or even weeks of painstaking work by researchers to identify the changes in behaviour and to identify the causes of those changes within the scenarios. While the effort is very large, and all expected behaviours can usually be found, there is also an issue in such an analysis that the researchers are never certain that they have found all of the changes since unanticipated changes may be overlooked due to the long duration of the data.

Driven by the need for such long duration eye-tracking studies we have taken methods that we are developing for the analysis of general multivariate temporal data sets and reworked them to address this specific application in order to increase both the speed and, hopefully, the accuracy of the analysis of these data. Our methods are based around a combination of new clustering techniques and some redesigned interactive visualization methods which provide immediate and complete access to even these long duration studies. Initial studies, described later in this paper, have shown that many features can be quickly observed, explored and then compared with the logs of the scenarios to identify the possible causes of the behavioural anomalies. The methods also allow a simple visual comparison of multiple subjects performing the same scenario to identify anomalies peculiar to a single subject, or to compare single subjects between two different scenarios to identify common behaviour between them. Since beginning this work we have received significant interest in the methods from other application areas where long duration eye-tracking studies are used, such as road traffic and aircraft pilot studies of cockpit behaviour and there exist similar difficulties in the analysis of the large and complex behavioural data.

This paper builds on two previous publications: a concept paper presented at BELIV2016 [2], and a technology paper which was presented at the IEEE VAST conference in Berlin, 2018 [3]. Consequently we emphasize the results of the studies in this work and the effectiveness of the methods with specific reference to the application area of Air Traffic Control, while referring the reader to the IEEE VAST publication for a more thorough description of the technical details.

II. Related Work

As mentioned above there is little related work in this area since long duration eye-tracking experiments of this kind are relatively rare, but there is some related work with which ours can be compared. The straightforward approach to identify Areas of Attention (AoAs) is to divide the scene into predefined grids, or use manually defined AoAs and assign labels to those eye points that fall under them [4]. Pre-defining AoAs without taking into account the underlying characteristics of the eye tracking can fail to capture the features that are inherently unique to every eye tracking dataset. In order to overcome this problem, clustering methods such as Mean-shift [5], Gaussian mixture models [6] have been used to perform clustering on the entire data set, but the local AoAs in a scene get saturated into very few clusters (AoAs) in case of very long duration eye tracking experiments. Some of the methods use image processing algorithms [7], [8] to look for features or objects in the display to which eye gaze can be attached to allow for the identification of AoAs which move over time following an optical flow in the display. One such method is that of the ISeeCube tools described in [9] which are used to analyze gaze in a changing display such as that in a video where objects can be expected to move and so AoAs move with them. Our approach is to identify spatial AoAs as they evolve over time rather than relying on object movement to guide them. A state of the art study on analysis and visualization of spatio-temporal eye tracking data is thoroughly discussed in [10]–[13].

Other approaches, such as those described in [14] and [15], work in multi-target environments like that seen in ATC and use the objects to analyse the gaze movements over time. This is related to our work but presumes that only objects are the carriers of attention, overlooking the regions of the display which may contain no moving objects can also be AoAs at different times during a long and evolving scenario. Our method, being untied from object movement, relies on no such features to convey the changing attention of the user over time.

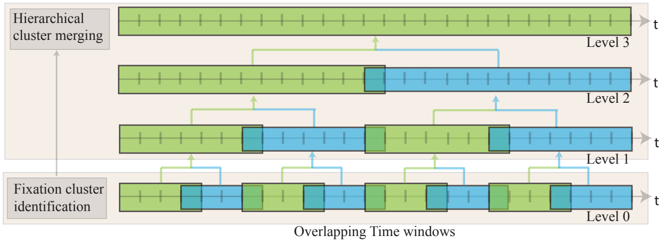


Figure 1: Eye tracking data is divided into overlapping time windows and fixation clusters are identified at level 0. Adjacent pair of time windows from level 0 are merged using a hierarchical merging process until only one time window remains. At each level of the hierarchy, the parallel computation is performed across the time windows to achieve efficient computation.

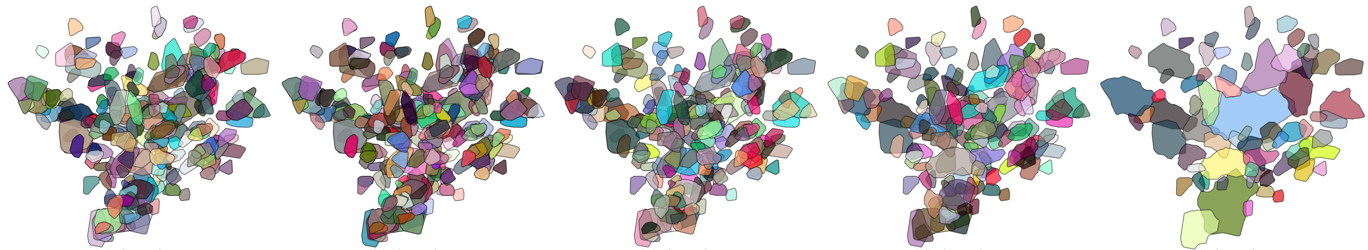
Methods such as McClung et. al [16] search for circular, linear and mixed type scanning patterns in eye gaze data collected from ATCOs. Statistical analysis of eye gaze data such as distribution of fixation count, fixation duration and saccadic velocity were used in Wang et. al [17] to study eye gaze patterns of ATCOs in order to understand the scanning patterns used to acquire situational awareness during their work. These approaches can provide useful short-duration knowledge but is not as representative of the evolutionary change which occurs in the basic attention of the subject during an eye-tracking study as our method provides. While the methods described may have significant benefit for some problems in this area, our approach addresses a more basic need: to understand the changes in user attention over time and how the AoAs appear, disappear and recur over time without being tied to the presence of moving objects.

III. Methodology

The data produced in our studies is in the form of 2D gaze coordinates across the display and acquired at 60 samples per second. For the 90 minute scenario involved some $\sim 320K$ samples are acquired for a single user. Errors due to occlusion, head movement out of range and other measurement issues mean that about 10% of the samples can be discarded thereby leaving behind $\sim 290K$ samples to be analysed. It is a large computational problem to analyze eye movements from many users but not intractable. We begin with a description of the hierarchical clustering technique that we use to find the AoAs in the eye-tracking data before considering the interactive visualization methods that we have found beneficial for their exploration.

A. *Hierarchical Clustering of Eye-tracking Data*

We begin with a clustering technique applied to the raw eye tracking data. Clustering all the data over the whole measurement scenario as a single step would be self-defeating as the entire data set is so cluttered that few areas of



(a) level 0 (b) level 1 (c) level 2 (d) level 5 (e) level 8

Figure 2: The cluster hierarchy from (a) level-0, clusters found in the raw data, to the highest merged level, (e) level-8. The merging of the lower level clusters into the higher level AoAs can clearly be seen in the progression from left to right.

attention can be found and those that are found provide little meaningful information. It would also present a significant computational task. Instead, we cluster the data in a series of time windows of a user-specified duration. The windows are also allowed to overlap by a user-defined percentage. The duration of each window and of the overlap are chosen to match the study being carried out. In the case of our studies in ATC, individual tasks are typically around thirty seconds in duration so we use a window size of one minute and a 50% overlap to ensure that no single action by the controller goes unrecorded. The individual windows thus contain up to 3600 data points.

Level-0: clustering. The gaze points within each window can be clustered using the method developed by Zahn [18], where a minimum spanning tree (MST) is computed from the delaunay triangulation of the eye gaze points. The weight of the edges of the MST corresponds to the euclidean distance between two gaze-points and by removing the inconsistent edges in the tree, the MST can be split into multiple subgraphs that correspond to fixation clusters over an AoA. Thus, clusters of spatially proximate points are found. We also apply a filter which removes clusters with very few points (typically 5 although this can be varied), since these small clusters have a very short gaze duration that can be associated with a task and hence, can be regarded as noise. The result is a set of clusters, with start and end times, for each overlapping time window that represent an AoA associated with single tasks being conducted by the test subject. We refer to this set of clusters as the level-0 clustering and the original eye-tracking data can be then discarded as we do not refer to it again during the analysis phase, unless we wish to modify the level-0 time windows and restart the whole analysis. This cluster computation step can be processed in parallel on suitable hardware to speedup the analysis.

Level-1: cluster merging. The fixation clusters found at level-0 are generally of too high a resolution to be helpful since several of them may refer to a single action being carried out by the subject. We resolve this by examining the clusters found in each adjacent pair of windows to explore which of them are very closely associated. An adjaceny matrix is computed using the clusters from adjacent pair of time windows where for every pair of clusters we measure their association using a cost-function based on the degree of overlap of the clusters. From the adjaceny matrix, minimum spanning tree sub-graphs are calculated, where each sub-graph corresponds to connected fixation clusters. By removing inconsistent edges

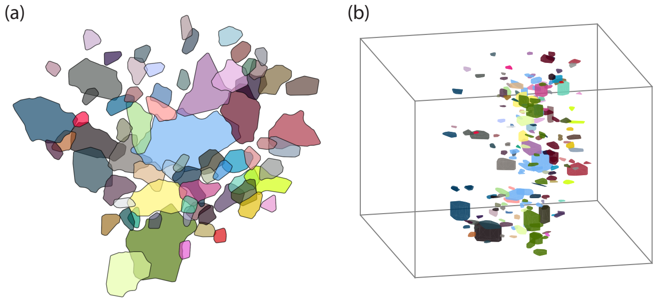


Figure 3: The areas of attention generated by an ATCO looking at their radar screen. (a) shows the AoAs using a 2D representation, where the temporal distribution is not visible while (b) uses a 3D space time cube display to allow clear visibility of the changes in the AoAs over time.

from the sub-graphs, highly connected clusters can be merged together into a single Area of Attention. The inconsistent edges from the sub-graphs are removed in the same way as for level-0, as discussed above using the Zahn's [18] algorithm. This edge deletion process can also be controlled by a user according to a user-defined selection parameter. This approach provides flexibility to an analyst to control the degree of merging among neighbouring AoAs. This merging produces the level-1 clusters or level-1 AoAs (regions of the display where the subject makes extended gaze fixations) which can be seen in Figure 2(b). These AoAs are associated with a new set of overlapping time windows at level-1, as can be seen in Figure 1. The merging process can again be easily parallelized for every pair of time-windows.

Hierarchical cluster merging. This merging of AoA sets from adjacent time windows can be repeated in a hierarchical fashion, as seen in Figure 1 until a level is reached where there remains only a single window. For our 90 minute data sets, with a one minute time window at level 0, this produces nine levels in the hierarchy. An analyst can freely navigate through the hierarchy with no recomputation required and so find levels which show specific behaviours. Subsets of time windows across the data, at any level of the hierarchy, can also be examined to compare behaviour across the duration of the study and identify changes.

B. Visualization Techniques

The AoAs found through the clustering technique can be examined, at any desired level, using a simple 2D display, as in the various sub-figures of figure 2, but with such long duration studies being considered this often produces

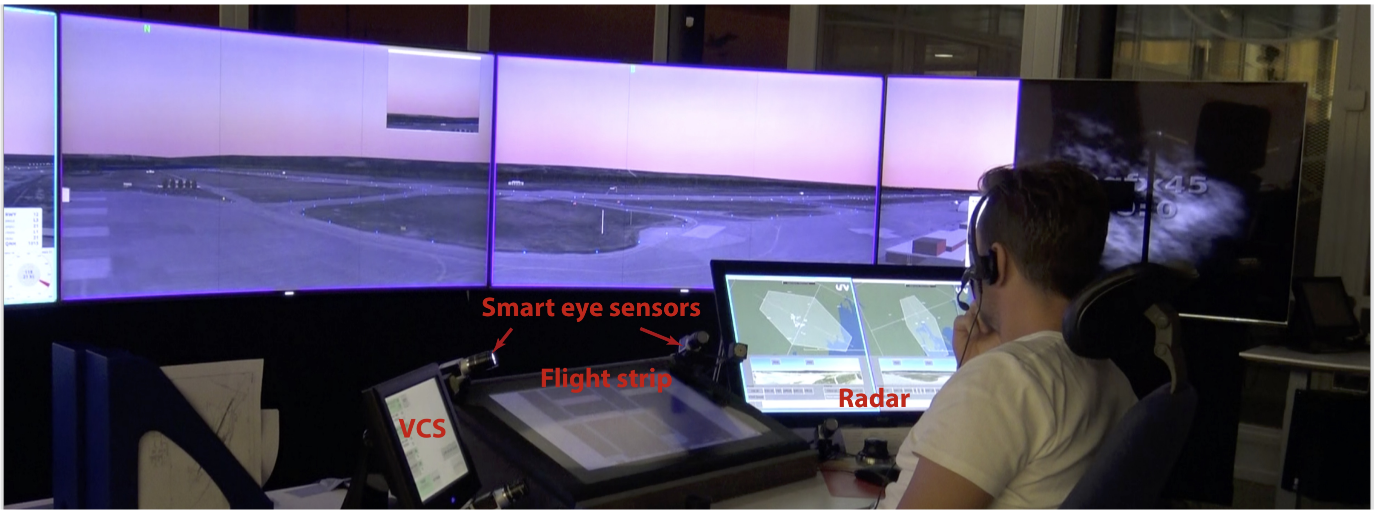


Figure 4: The experimental setup used in our case study consists of six screens making up the simulated tower view with three supporting screens placed on the desk. The six cameras of the SmartEye eye tracking system can be seen strategically positioned around the subject's workstation.

a cluttered display which is difficult to interpret and where it is hard to recognize change over time and so identify the source of the behavioural change. Instead, we have used a three dimensional representation of the AoAs, known as a 'space-time cube' [19], [20], which was originally developed to examine the movements of individuals or populations over time. In this display a 2D 'map' is presented at the base of a cuboid frame, and the temporal data then displayed above it with time going upwards, away from the map surface. As a static display this is often ineffective as spatial relationships can be lost between temporally separated points but, when interactive rotation is permitted the analyst can generally identify the spatial relationships present and draw meaning from them. An example can be seen in Figure 3(b).

To enhance the space-time cube and make it even easier to identify spatial relationships, we are able to exploit the hierarchical clustering approach to make it even clearer. The AoAs found at the higher levels of the hierarchy are spatially proximate, having been formed by merging those at the lower levels based on their degree of overlap. Thus, we can use colour information from the higher level AoAs, to recolour those found at the lower levels right down to level-0 if required. Several AoAs which are separated in time at, for example, level-3 may inherit the same colour information from a single AoA into which they have been merged at level-8. Their spatial proximity is thus made much more clear in the 3D representation and so the data is more easily interpreted. Examples can be seen in Figures 3(a) & (b) where the colour information (labels) from level-8, the highest level, have been propagated down to level-4 so groups of AoAs at that level will appear with the same colour because they are spatially proximate but temporally disparate.

IV. Case Studies

The system developed is designed to enable the analyst to more easily explore long-duration eye-tracking recordings in complex recording environments. In this section, we present a number of practical examples of how the visualization tool can be used with real data sets recorded using experienced controllers performing standard training scenarios. The specific use cases described in this section illustrate how the tool can be used to perform:

1) An overall analysis of a single recorded session.
2) Comparison between a number of different controllers.
3) Analysis of visual attention during specifically selected events, such as single landings or simultaneous aircraft activity in two remotely controlled airports.

The specific data sets have been recorded within a virtual tower environment (see Figure 4) using a set of scenarios run through an air traffic simulator. In each scenario, the controller is operating on two airfields simultaneously, with varying levels of traffic over time and with weather and other events predefined within the scenarios. The ATCO has standard radar, strip and VCS (Voice Com System) displays on the desk in front of them and the two airfields are visible through six large displays arranged in a semicircle in front of the controller. This is referred to as the 'out-the-window' (OTW) view. The six large displays show both airfields simultaneously, three being used for each. Each airfield also uses a 'picture-in-picture' view within the OTW view, from a 'Point-Tilt-Zoom' (PTZ) camera under the ATCOs control. The PTZ was, in all cases, used to display a magnified view of the approach end of the runway and remained unchanged throughout the simulation. The radar and strip displays each include information for both airfields simultaneously. The SmartEye gaze tracking system used in these experiments used six cameras, arranged around the controller, providing a wide field of view and so a comfortably large recording volume. Thus, the controller was able to move in a quite natural manner during the recording sessions but was required to remain seated.

The recorded data is in the form of separate screen coordinates for each of the display screens at which the controller may be looking. Consequently, the data is preprocessed before analysis to create a single volume of data which includes the data from each of the individual screens. The base (top-down)

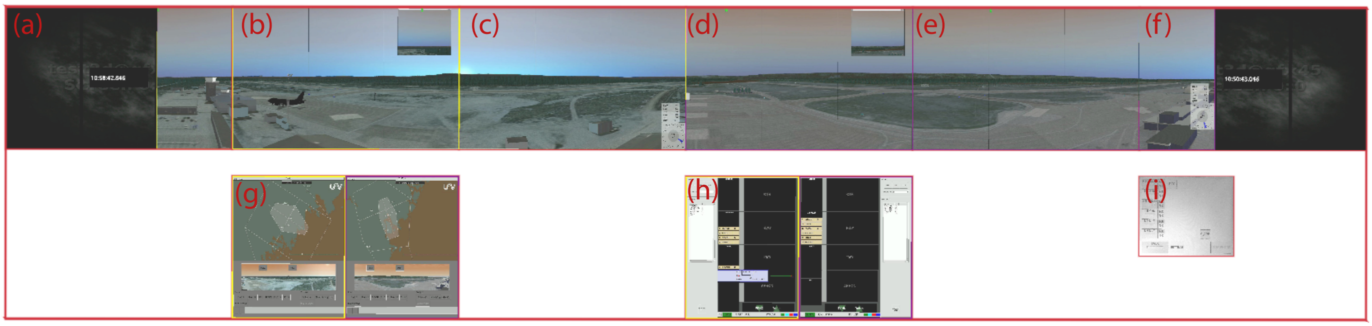


Figure 5: Screen setup: First airport - (a) to (c), Second airport - (d) to (f), Radar screen for both airports - (g), Flightstrips for both airports - (h), Voice Com System - (i)

view of the preprocessed data can be seen in Figure 5 with the tower OTW views (a-c and d-f), the radar (g), strips (h) and VCS (i) remapped into a single 'screen-time' volume.

In this work, we distinguish between the notions of areas of interest and of areas of attention. Areas of interest (AoIs) are specific areas that contain items or instruments of analytical interest, such as the e-strip area, the PTZ camera and so on. AoIs are pre-defined by the analyst. Areas of attention (AoAs) are areas to which the operator actually pays attention, as captured by the eye tracker, and then identified by our analysis tools. Typically the AoAs are at considerably higher resolution than the anticipated AoIs. The identified AoAs can then help the analyst to refine the AoIs.

Each of the cases presented here, together with the corresponding analyses, have been reviewed by a professional ATCO.

A. *Case 1: Overall analysis of a training session*

Figure 6 shows the identified AoAs for one ATCO and the AoIs, at the screen level, as predefined by the analyst. The X and Y axes correspond to the 2D-coordinates of gaze points during the entire 90 minute session. We can immediately see a clear correspondence between the AoAs and AoIs for the controller carrying out the tasks of the scenario. For example the AoAs in the radar views are neatly divided between AoIs corresponding to the two airports being controlled. The same can be seen for the flight strip display.

Several AoIs can be naturally associated with the flight strip tool (e-strip) such as the areas for apron, taxiway (TWY), runway (RWY), and control zone (CTR), visible in Figure 8. Observation of AoAs over the flight strips on Figure 8 leads to two conclusions. Firstly, we see a strong correspondence between the different AoIs and the AoAs on the flight strips. Secondly, the ATCO paid much more attention to the flight strips of the left airport compared to the right airport.

On the OTW view for both airports in Figure 6, we see an accumulation of AoAs at both ends of the runway, with fewer AoAs in the middle of each runway. This indicates that the ATCO pays most attention to the ends of the runway, while the center part of the runway is only glanced at more sporadically. This is also what one could expect from ATCO work. We can thus use this overarching picture to determine whether the ATCOs are working roughly as expected, or not. We see no clear deviations from our expectations in this case.

We can also observe several AoAs on the PTZ camera, for each airport, indicating that the ATCO has paid attention to this camera. However, Figure 6 gives no information about when the ATCO looked at the PTZ camera. This information can be obtained by showing the time component through the space-time cube display, as shown in Figure 7. Time is associated with the vertical axis. For the left airport, for example, we observe that the ATCO paid attention to the PTZ camera regularly, when a lot of attention was also focused at the approach end of the runway, and also at the end of the session although little attention was paid to the runway at that time (see annotated Figure 7).

It should also be noted that, while the single images included in this paper may appear cluttered due to the complex nature of the data being displayed, by rotating the $3D$-view, as seen in Figure 7, the three dimensions (X, Y, and t) of the AoAs are visible and the user can minimize occlusion of

Figure 6: 2D 'top-down' view with AoIs. The two tower (OTW) views along the top. The two radar displays, the two strips displays and the VCS display are arranged along the bottom.

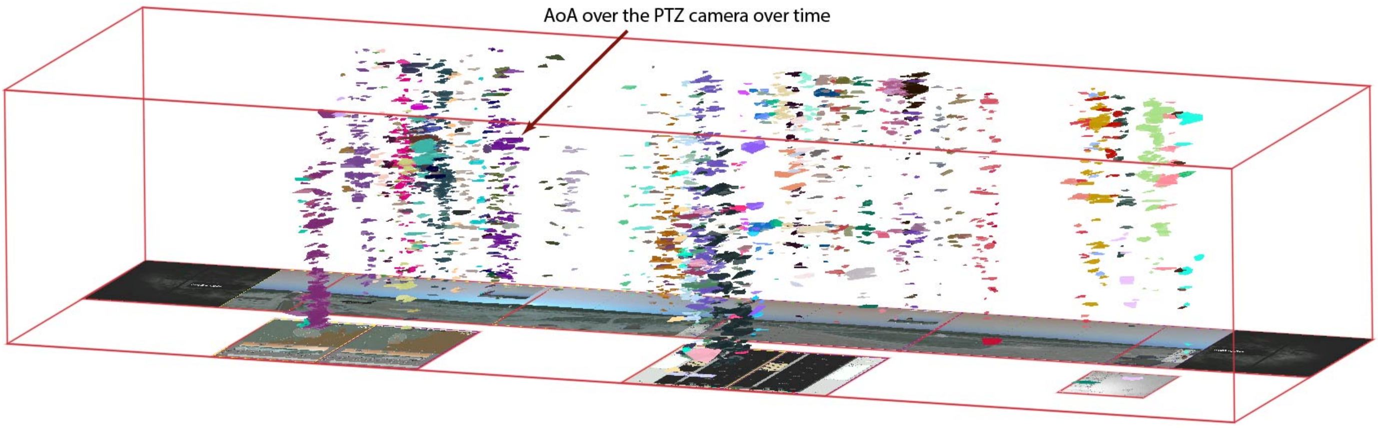


Figure 7: ATCO-B side view: AoA over PTZ camera over time

Figure 8: Flight strips: Top view.

specific AoAs in which they are interested, for example the series of purple instances denoting the AoA over the PTZ camera.

It is possible to further enhance the data analysis by adding the trajectories of moving objects (such as aircraft) with indication of specific events (such as touch down), as shown in Figure 10. In this figure, we can clearly see that the "empty" part of the display on the right side corresponds to a time period without traffic. We also observe that there is a correspondence between the attention put on both ends of the runway and the existence of moving objects. These observations can certainly be useful to assess whether the ATCO is roughly working as intended.

B. Case 2: Comparison between different controllers

In this case we consider the comparison of the actions of a number of different controllers, in this case four, performing the same scenario. This comparison allows the analyst to explore the variation in behaviour by different controllers under similar circumstances. The dual remote tower scenario is predefined and incorporates a variety events over time including the arrival of new approaching flights and scheduled departures, variations in visibility due to weather, and ground vehicle movements. These events are initiated at the same times in each run of the simulation but there can be some variation as a result of different behaviour by the various ATCOs–perhaps routing an aircraft slightly earlier or later.

Comparing the four examples shown in Figure 9, we can see the fundamental behaviour is very similar, driven by the conditions imposed by the scenario and the regulations and guidelines with which the four ATCOs comply, with all informational elements (radar, estrips, VCS and tower views) being monitored, however, the overall behaviour of each controller is very different. ATCO-C, whose areas of attention are shown in figure 9(c), has a significantly more sparse series of AoAs than the others, with a lot of time where his attention is not on the informational displays at all. The controller in Figure 9(a) also has a quite sparse sequence of AoAs but clearly spends far more time focused upon the radar display than ATCO-C, particularly for the leftmost of the two airports which has a considerably higher level of traffic. ATCO-A spends the least time monitoring the OTW tower views and the most on the radar display as shown by the large AoAs present over the radar display, shown in pink in the figure. ATCO-B is similar in behaviour to ATCO-D but spends even more time on the radar and strips. Clearly ATCO-C and ATCO-D have a recurring eye gaze attention at different areas of the OTW screen.

It is thought that the ability to analyse and compare a number of long-duration recordings of this type will be of benefit in defining, demonstrating and checking adherence to 'best practice' during the training process. It might also be of value in the process of incident investigation, where the mis-steps and failures of attention leading to the incident might be visible when compared with other controllers operating in similar scenarios. To enhance this it would be beneficial to be able to directly compare two or more recordings through a computational comparison and this is being developed but is currently regarded as future work.

C. Case 3: Event selection

In this case, we illustrate now how the visualization tool can be used to select and analyse behaviour associated with specific states (or events) in the dataset. Using a temporal slider and the $3D$ view, we can easily filter out data to show only those data recorded during a specific time period associated with specific events within a scenario, and so show only the corresponding visual attention. We also stress that being able to create temporally filtered views helps the analyst to cope with occlusion of AoAs on the unfiltered view showing the whole session.

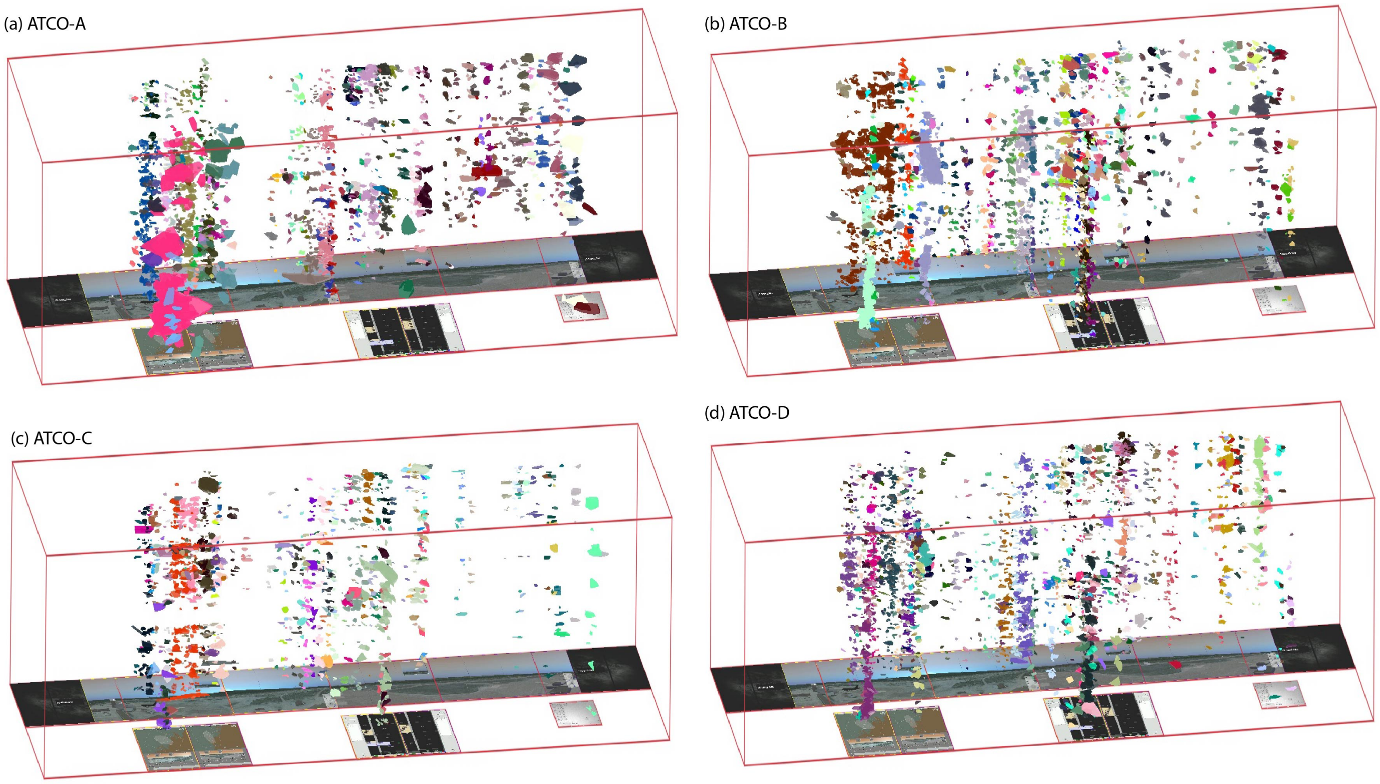


Figure 9: Four air traffic controllers carrying out the same scenario. Despite all regulations and guidelines being complied with by these experienced controllers, and no significant differences developing during the studies, the behaviour of each controller is clearly quite unique. Note that the colours assigned to AoAs differ between the four different displays since they are assigned during the data analysis phase.

1) Simultaneous aircraft in two monitored airports: As a first example we choose to focus on a situation with simultaneous aircraft movements at both airports. On the overview Figure 11, we see attention (AoAs) corresponding to all AoIs that require attention for checking a movement: both PTZ cameras on the OTW, both wind widgets on the OTW, both radar views, and the flight strips. A smaller AoA over the radio indicates that some attention was also given to this instrument while there were simultaneous aircraft in both airports. On the OTW, we observe that attention is appropriately allocated to follow each aircraft position, with a focus on both ends of the runways. This confirms that ATCO's attention was appropriately given to the most critical areas.

Figures 14 and 15 show the selected movement in the right airport. On the former, we can see the AoAs detected by the tool for the flight strip corresponding to the right airport, while the latter figure shows the same AoIs with the time component. For instance, one can clearly observe on Figure 15, the attention to the flight strips and attention to the aircraft movement. Attention is also given to the weather tool before the aircraft is visible. The ATCO looks at PTZ camera when the aircraft is in the end of the runway and has almost disappeared from the rightmost screen. Similar conclusions are valid for the AoAs on the PTZ camera and weather tool of the left airport, as shown in Figures 12 and 13. We can also see in this figure that the ATCO focused their attention on the radar monitor in advance of visual contact with the aircraft. Hence, using the temporal space time cube, we can

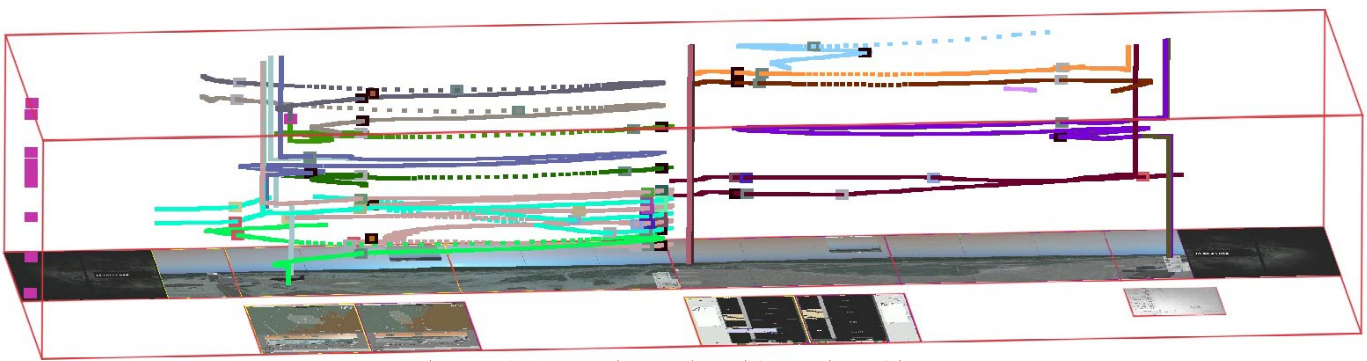


Figure 10: Space time cube with moving objects.

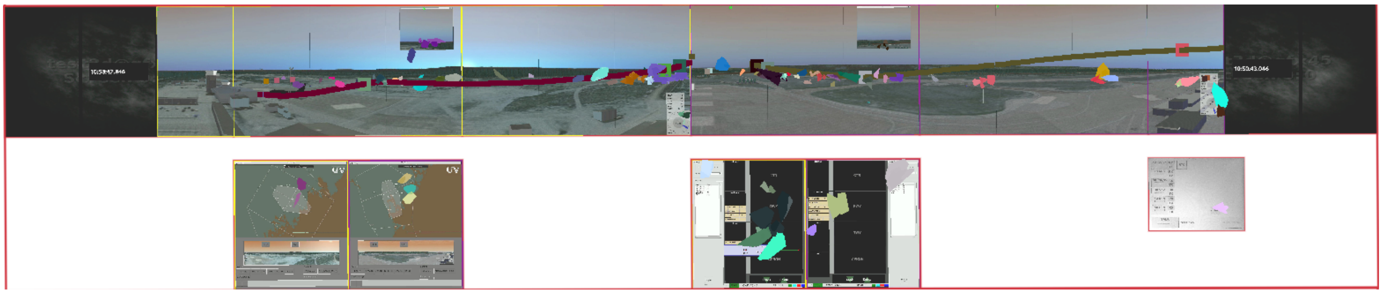


Figure 11: ATCO-D top view: AoAs and aircraft in both airports.

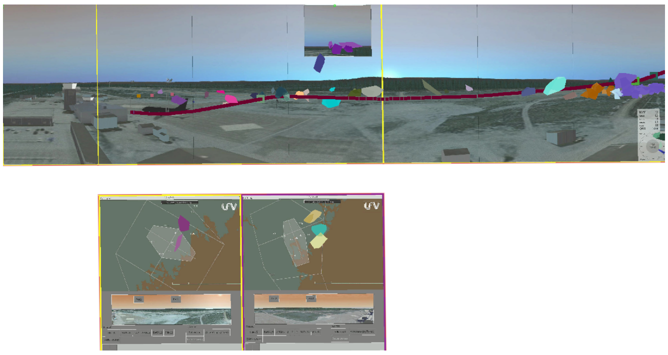


Figure 12: ATCO-D top view: AoAs on left airport.

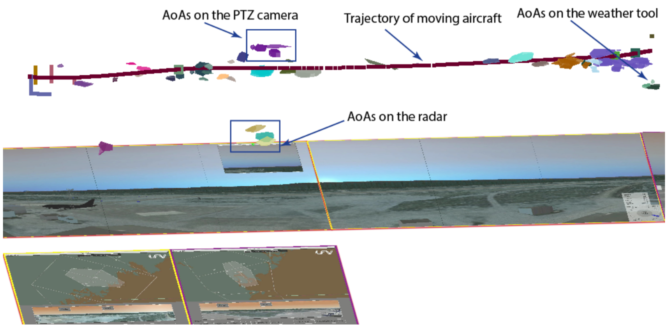


Figure 13: ATCO-D side view: AoAs on left airport with time component.

confirm the order of visual attention of the ATCOs while they monitor the critical parts of the screens during an event.

2) Landing events: From Figure 9(d) we can observe that ATCO-D has a periodic scanning behaviour at specific areas of the airport and hence there are multiple AoAs that are repeatedly active during different time intervals. After localizing the landing events in the right airport we were able to observe a strong similarity in the scanning behaviour, where attention is paid repeatedly at specific areas on the display and this is highly similar for different landing events as shown in Figure 16(a)-(c).

3) Scanning behaviour: We can also observe, as seen in figure 17, that this controller exhibits a strong pattern of scanning behaviour both during a scheduled departure (Figure 17(a)) and during a period of no activity on the airport (Figure 17(b)).

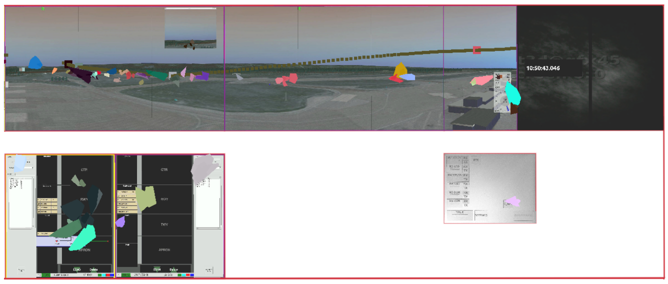


Figure 14: ATCO-D top view: AoAs on right airport.

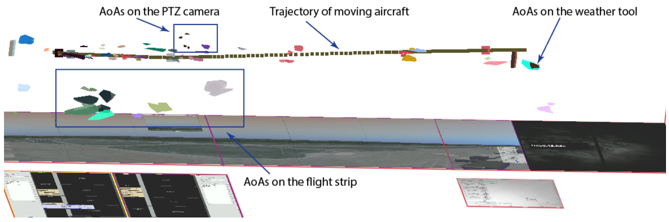


Figure 15: ATCO-D side view: AoAs on right airport with time component.

V. Discussion

Using the set of cases that are discussed above, we have demonstrated our method which is designed to shift the focus from the Areas of Interest (AoI) that are typically specified by the analyst and, instead, to explore the Areas of attention (AoA) which are contained within the data itself. The AoAs are usually much more focused than the AoIs and so can tell the analyst more about how the subjects' attention is focused while undertaking complex tasks.

The 'top-down' (overview) of the AoAs can provide a certain insight into how the subjects' attention is being directed but the 3D view can provide great insight into the frequency with which different tools and displays are being attended to as well as when during the scenario the subject regards them as important. Patterns of behaviour can be observed and occasions when attention on a tool or region of a display is absent may be apparent.

The 3D space-time cube allows the very different behaviour of the individual subjects to be made clear. While their AoAs may be spatially very similar, when observed with the temporal element made clear the way in which they are performing the tasks may be shown to be quite different. It

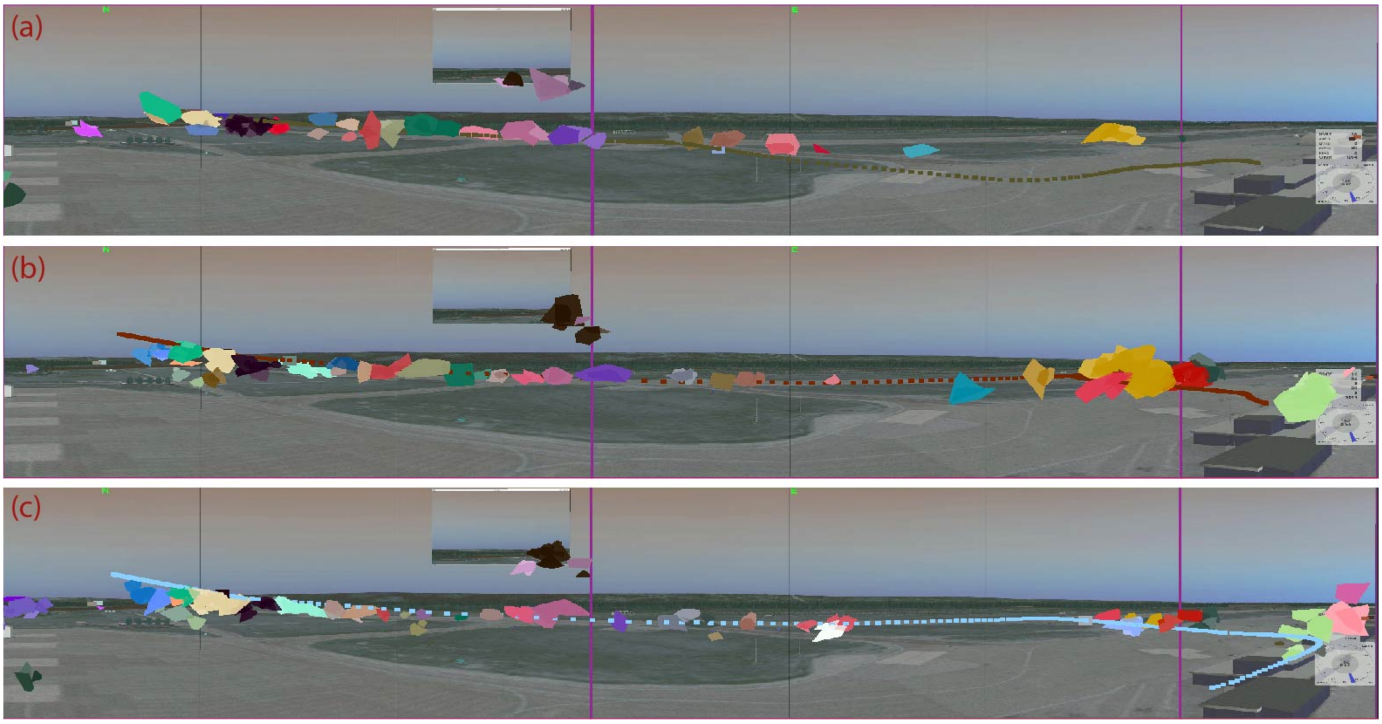


Figure 16: Similarity in AoAs of an ATCO monitoring different landing simulations.

is hoped that this will allow good and bad behaviours to be observed and their causes to be identified. In this way, it may be possible to accelerate the training processes and show the trainee where failures in attention are leading to errors.

The 3D space time cube display is a powerful tool but does suffer from cluttering, when many small AoAs have been identified. While interaction (rotating and zooming the display, and cropping the data space) can alleviate this, it is impossible to remove it completely. We are currently exploring other displays to avoid this problem using transparency and volumetric rendering approaches.

The system currently identifies the AoAs within each subject's recorded eye-tracking data and this sometimes makes cross comparison between users, as seen in our second example case, quite difficult. We are currently working on methods to automatically cross-correlate the identified AoAs between sets of users to make this process simpler and to allow for automatic measures of similarity and difference to be extracted from the several data sets.

VI. Conclusions

The analytical method we have developed and implemented within an analysis application is a powerful tool for the extraction of multi-dimensional clusters in temporal data. It is computationally efficient and can handle the very large data sets extracted from long duration eye-tracking experiments with ease. Thus, it is well suited to the analysis of user attention within the experiments being carried out within the field of air traffic control research.

The identified AoAs can be explored by the analyst from an overview level to a very fine grain level, while the unique

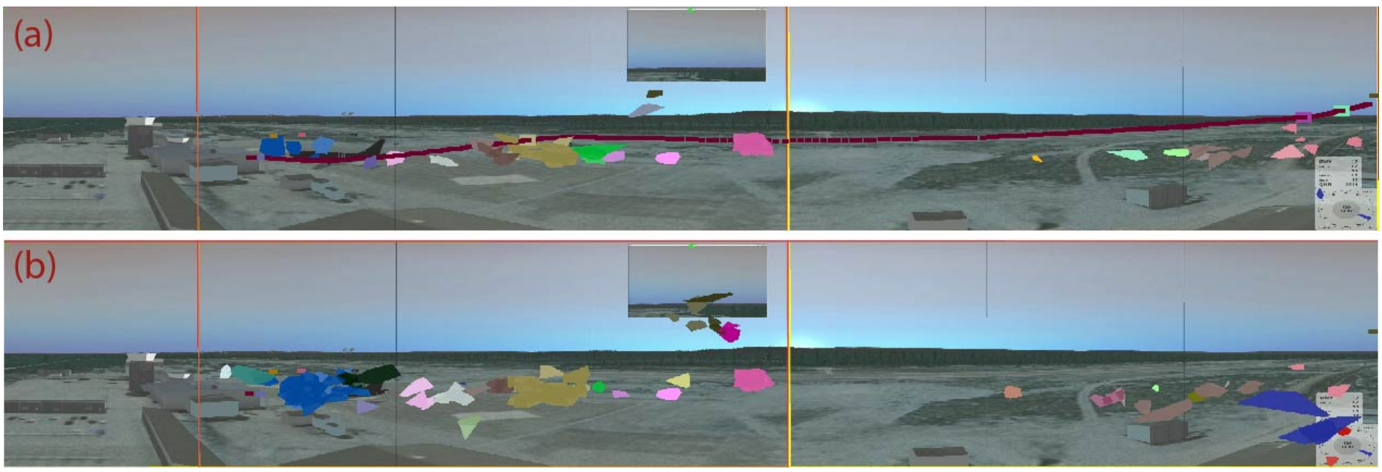


Figure 17: Scanning behavior with similar AoAs in the presence and absence of aircraft on the screen.

AoAs identified using their color information, can help an analyst to identify similar gaze behaviour over time. The 3D space-time cube allows the analyst to quickly gain visual insight into the overall behaviour as well as to identify behavioural patterns and anomalies. Exploring such events is then easily achieved by selecting small time windows around specific events and examining them in detail. The unique color labels of AoAs can be used in future with sequential data mining algorithms to find repeating patterns and anomalies in user attention computationally.

The behaviour of small numbers of test subjects can easily be compared and some similarities and differences in behaviour can be readily identified. In future work, we will extend this capability through automatic methods to cross-correlate and compare the AoAs of different users and how they are accessed over time. This ability to cross-correlate the AoAs between subjects will have an additional benefit since it will permit the identification of a common set of AoAs within which sequence identification will be possible. Using this approach, it is to be hoped that common behaviours, both positive and negative, can be found and used to create predictors of potential future incidents.

ACKNOWLEDGMENT

Funding for this research project was provided by the Swedish Transport Agency, Trafikverket, and the Swedish Science Council, Vetanskapsrådet.

REFERENCES

[1] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 455–470, 2002.

[2] P. K. Muthumanickam, C. Forsell, K. Vrotsou, J. Johansson, and M. Cooper, "Supporting exploration of eye tracking data: Identifying changing behaviour over long durations," ser. BELIV '16. New York, NY, USA: ACM, 2016, pp. 70–77.

[3] P. K. Muthumanickam, K. Vrotsou, A. Nordman, J. Johansson, and M. Cooper, "Identification of temporally varying areas of interest in long-duration eye-tracking data sets," *IEEE transactions on visualization and computer graphics*, 2018.

[4] J. H. Goldberg and J. I. Helfman, "Scanpath clustering and aggregation," in *Proceedings of the 2010 symposium on ETRA*. ACM, 2010, pp. 227–234.

[5] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[6] 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*, vol. 3, no. 1, pp. 5–24, 2011.

[7] C. M. Privitera and L. W. Stark, "Algorithms for defining visual regions-of-interest: Comparison with eye fixations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 9, pp. 970–982, 2000.

[8] K. Kurzhals, M. Hlawatsch, F. Heimerl, M. Burch, T. Ertl, and D. Weiskopf, "Gaze stripes: Image-based visualization of eye tracking data," *IEEE TVCG*, vol. 22, no. 1, pp. 1005–1014, 2016.

[9] 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*. ACM, 2014, pp. 43–50.

[10] G. Andrienko, N. Andrienko, M. Burch, and D. Weiskopf, "Visual analytics methodology for eye movement studies," *IEEE TVCG*, vol. 18, no. 12, pp. 2889–2898, 2012.

[11] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, "State-of-the-art of visualization for eye tracking data," in *Proceedings of EuroVis*, vol. 2014, 2014.

[12] ——, "Visualization of eye tracking data: A taxonomy and survey," in *CGF*. Wiley Online Library, 2017.

[13] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[14] Z. Kang and S. J. Landry, "An eye movement analysis algorithm for a multielement target tracking task: Maximum transition-based agglomerative hierarchical clustering," *IEEE Trans. Human-Machine Systems*, vol. 45, no. 1, pp. 13–24, 2015.

[15] Z. Kang, S. Mandal, J. Crutchfield, A. Millan, and S. N. McClung, "Designs and algorithms to map eye tracking data with dynamic multielement moving objects," *Computational intelligence and neuroscience*, vol. 2016, 2016.

[16] S. N. McClung and Z. Kang, "Characterization of visual scanning patterns in air traffic control," *Computational intelligence and neuroscience*, vol. 2016, 2016.

[17] Y. Wang, W. Cong, B. Dong, F. Wu, and M. Hu, "Statistical analysis of air traffic controllers eye movements," in *The 11th USA/Europe ATM R&D Seminar*, 2015.

[18] C. T. Zahn, "Graph-theoretical methods for detecting and describing gestalt clusters," *IEEE Transactions on computers*, vol. 100, no. 1, pp. 68–86, 1971.

[19] T. Hägerstrand, "What about people in regional science," *Regional Science Association Papers*, vol. 14, pp. 6–21, 1970.

[20] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of time-oriented data*. Springer Science & Business Media, 2011.

AUTHOR BIOGRAPHY

Prithiviraj Muthumanickam is currently a PhD student at Linköping University, Sweden. His current research activities focus on visualization and analysis of very long duration eye tracking datasets and his research interests include multivariate temporal data analysis and visualization.

Aida Nordman is a Lecturer in Computer Science at Linköping University, Sweden where she works with knowledge discovery in multidimensional event-based data. She received her PhD in Computer Science in 2010 from Linköping University. Her research interests include data mining, knowledge representation and reasoning, description logics, temporal logic, and artificial intelligence.

Lothar Meyer is a safety engineer at Luftfartsverket (LFV), the Swedish Air Navigation Service Provider. He holds a degree in electrical engineering and a doctorate in air traffic services from the Technische Universität Dresden, where he worked as a research associate in the field of aviation safety. His areas of expertise covers socio-technical systems, risk assessment and airport surveillance technologies.

Supathida Boonsong has 16 years of experience in Air Traffic Control specializing in aerodrome and terminal control. From 2004 to 2017, she worked at Bangkok Terminal Control, AEROTHAI. Currently, she is working at Luftfartsverket (LFV), in the RESKILL project.

Jonas Lundberg has a PhD in Computer Science from Linköping university, 2005, in the area of Interaction Design, with Human-Automation Collaboration as the current main area of research. He has also conducted research in Safety Science since 2006. Since 2016 Associate Professor in Information Design at Linköping University, Sweden.

Matthew Cooper received his B.Sc. in Theoretical Physics in 1986 from the University of Exeter, UK, and his M.Sc. in Computer Science and Ph.D. in Chemistry from the University of Manchester, UK, in 1987 and 1990, respectively. He is associate professor in Information Visualization at the University of Linköping, Sweden. His focus is in visual data analysis of multivariate temporal data.