Let's look at the Cockpit: Exploring Mobile Eye-Tracking for Observational Research on the Flight Deck

Nadir Weibel* Adam Fouse* Colleen Emmenegger* Sara Kimmich* Edwin Hutchins*
Distributed Cognition and Human Computer Interaction Laboratory, University of California San Diego

Abstract

As part of our research on multimodal analysis and visualization of activity dynamics, we are exploring the integration of data produced by a variety of sensor technologies within ChronoViz, a tool aimed at supporting the simultaneous visualization of multiple streams of time series data. This paper reports on the integration of a mobile eye-tracking system with data streams collected from HD video cameras, microphones, digital pens, and simulation environments. We focus on the challenging environment of the commercial airline flight deck, analyzing the use of mobile eye tracking systems in aviation human factors and reporting on techniques and methods that can be applied in this and other domains in order to successfully collect, analyze and visualize eye-tracking data in combination with the array of data types supported by ChronoViz.

CR Categories: H.1.2 [Information Interfaces and Presentation]: User/Machine Systems—Human Factors;
H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces—Evaluation/methodology, Theory and methods;

Keywords: Mobile Eye-Tracking, Ethnography, Video Analysis, ChronoViz, Observational Research, Aviation Human Factors

1 Introduction

Digital technology is transforming the way researchers conduct their experiments and plan their field work. The possibilities introduced by new digital devices and the provision of almost infinite storage for high-definition data open up new frontiers for observational research. It is now common to employ a range of digital devices to collect data on ongoing activity. Multiple HD digital video or still cameras, directional audio microphones, depth cameras, and digital pens have become typical tools for researchers in many disciplines. The techniques for the collection of rich multi-stream data sets are being extended into previously inaccessible real-world settings. Just as the availability of audio recording supported the development of conversation analysis [Sacks 1995] and the ethnography of speaking [Hymes 1962], the advent of a range of inexpensive digital recording devices and sensors promises to have a fundamental impact on a broad range of scientific and engineering activities.

While this abundance of data collection devices certainly helps in recording many of the important details of human activity and supports new methods based on what we now call *digital ethnography*, such a huge amount of information also introduces challenges for analysis and data integration. Moreover, raw data is often coded and transcribed in a wide variety of ways, creating new re-representations of the original events [Hutchins 1995]. Coordinating these multiple re-representations is a hard problem and it typically requires manual alignment and management of the different data sources. Recent work [Fouse and Hollan 2010] began tackling this issue through the introduction of a tool called ChronoViz [Fouse et al. 2011] for visualizing multimodal data that enables the integration of a number of different data streams, presenting a cohesive visualization of the collected data. ChronoViz introduced novel forms of interactive visualization and interactive navigation for time-series data. By supporting the analysis and management of multimodal information, ChronoViz opens up a new range of exploratory data analysis possibilities [Tukey 1977].

In our own research on aviation human factors we are interested in studying activity dynamics in the cockpit of commercial airplanes. Understanding pilot behavior in an existing flight deck or a flight deck under development requires a thorough understanding of the distribution and dynamic allocation of crew attention. One component of attention is *visual attention*. The ability to track what the crew is looking at throughout the phases of flight is invaluable in understanding the crews allocation of this attentional component. In order to study visual attention and understand its dynamic allocation, we decided to experiment with the integration of a mobile eye tracking system into the ChronoViz tool.

*e-mail: {weibel,afouse,cemmeneger,skimmich,ehutchins}@ucsd.edu

ETRA 2012, Santa Barbara, CA, March 28 – 30, 2012.
© 2012 ACM 978-1-4503-1225-7/12/0003 $10.00

2 Related Work

Eye tracking research and applications are diverse and growing in number as eye tracking technology becomes more robust, wearable, and transportable. Eye tracking is now commonly used in a variety of disciplines including neuroscience, marketing and advertising, psychology and computer science [Duchowski 2002; Duchowski 2007]. Much of the research is concerned with the neurological substrate of the human visual system. How to exploit the physical characteristics and visual psychophysics of the eye and retina in the development of eye tracking glasses and object registration on a visual scene, as well as what constitutes a moment of visual attention concern much of the contemporary research [Duchowski 2007].

Promising and relevant research is coming out of work in Attention Aware Systems (AAS) and similar efforts. This work is attempting to design ubiquitous computing systems capable of assessing the current user's focus of attention and make predictions about where their attention may be in the near future, and when their attention may need to be redirected. Developers have started using eye trackers for point of gaze as an alternate computer interaction paradigm [Roda and Thomas 2006; Roda 2011]. For example, Zhai is developing ways to move a cursor and select an object on the computer screen with a user's eye movements [Zhai 2003].

More interestingly to us, eye trackers are going into even more complex and natural settings, with unconstrained eye, head and hand movements, including daily activity and professional work environments. Studying eye scan behavior in these settings is crucial as findings from studies performed in non-natural settings such as detecting images in static pictures [Ramanathan et al. 2010] are unlikely to inform analysis of complex behavior in more realistic settings [Tatler et al. 2011]. Realistic settings that have been studied with eye tracking include hand movements while cooking [Land and Hayhoe 2001], visual attention during reading tasks in a foreign language [Hyrskykari 2006], and attention on other speakers during virtual group meetings [Vertegaal 1999]. In aviation, researchers are taking eye trackers into high-fidelity space shuttle and airplane simulators. Huemer performed eye movement analysis in the Intelligent Spacecraft Interface System (ISIS) simulator [Huemer et al. 2005]. This study found that astronauts modified their instrument scan strategies depending on the flight segment and the number of shuttle malfunctions occurring during a trial. Researchers are using eye trackers to study pilots' monitoring strategies and how visual attention allocation is impacted by changes in the flight display technology [Wickens et al. 2004; Thomas and Wickens 2004; Sarter et al. 2007; Steelman et al. 2011]. Others are using double eye tracking systems on both pilots to monitor automation mode call-outs [Björklund et al. 2006]. Interestingly, this research outlined how mode awareness is a more complex phenomenon whose description requires more than eye gaze and communication. In another interesting domain, researchers are using visual attention measurements to monitor drivers' observations and intervene if necessary [Fletcher and Zelinsky 2009; Hammoud et al. 2008].

In order to study many of the outlined domains, research relies on a wide range of visualizations of eye tracking data. This is a critical component that allows researchers to explore and comprehend the raw data captured through the eye tracking devices. Many approaches have been proposed and it is common practice to exploit a number of them to analyze the data from different points of view. Fixation, heat and fog maps [Špakov and Miniotas 2007] are classic variants of a simple visualization of the number or proportion of fixations that landed on a specific region of a fixed display. While these visualizations allow the analysts to have an overview of the subject's fixation, it is not easy to compare multiple subjects or sessions. The eSeeTrack visualization [Tsang et al. 2010] tries to overcome this problem by combining tree-structured visualization of the areas of interest with a timeline, facilitating comparison of sequential gaze orderings. Often the temporal order of visiting areas of interest is critical and gaze paths rather than gaze points need to be analyzed. Räihä et al. introduced a technique that treats time and gaze path as the prime attribute to be visualized [Räihä et al. 2005].

Research based on eye tracking is challenging and requires sophisticated handling of the collected data. While a lot has been done, we feel that no current system adequately supports capturing and mapping visual attention *in the wild*, in naturalistic and complex environments, and allowing for consideration of attention that is distributed across multiple parties and integrated with other modalities (such as haptic and auditory attention). In this paper we will show how we have begun to tackle eye tracking in such complex ecologies, using our research on aviation human factors as an example. We will show how we integrated mobile eye tracking into ChronoViz – a tool for visualizing and navigating multimodal time-synchronized data – to reach this goal.

3 ChronoViz

ChronoViz [Fouse et al. 2011] is a system designed to address the analysis bottleneck resulting from collecting multiple data streams by providing synchronized interactive visual representations of the multiple data streams. While several systems have been developed to support various aspects of this analysis challenge, ChronoViz is unique in focusing on navigation and visualization of multiple diverse data sources. For example, although numerous systems exist for coding and annotation of video, such as ELAN [Wittenburg et al. 2006], VCode [Hagedorn et al. 2008], and Diver [Pea et al. 2004], they are either not designed for analysis or visualization of multiple types of data, or they do not support easily extensible visualization and navigation facilities. Even in the aviation human factors domain, the Cognitive Avionics Tool Set (CATS) is looking at analyzing flight data in combination with physiological and neurological markers [Schnell et al. 2008]. Although this system organizes multiple data streams, its main effort is to classify pilot workload state and not visualize and explore their activity.

ChronoViz allows researchers to visualize time-based data from multiple sources, navigate this data in flexible ways, and manually or automatically code the data with structured or unstructured text-based annotations. The data sources can include multiple video files, audio files, computer logs, sensor readings, paper notes, and transcriptions (see Figure 1). Since these recordings are often from different devices that can be difficult to synchronize, ChronoViz offers interactive mechanisms for aligning data sources. Researchers can create annotations, define and evolve category structures for annotations, and construct visualization filters. Annotations can also be created automatically by employing an analysis plug-in framework that supports custom scripts [Fouse and Hollan 2010].

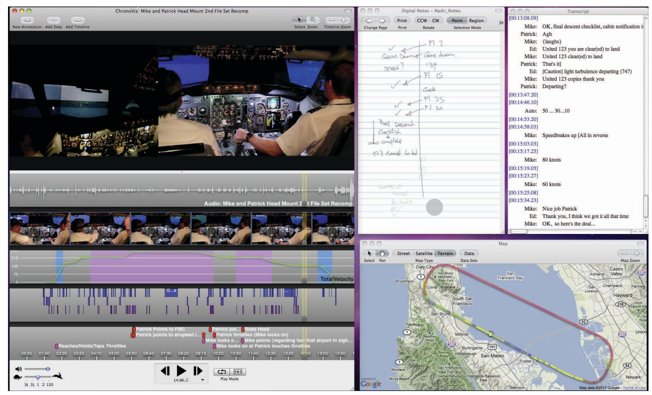


Figure 1: *The main ChronoViz interface used to analyze crew coordination during a simulated flight in a Boeing 737. Left: main window consisting of the video data, timelines and annotations. Right: three visualizations of the additional collected data (top-left: paper notes, top-right: speech transcription, bottom: GPS).*

ChronoViz exploits digital pen technology[1,2] to support integration of paper-based digital notes. It enables paper to be turned into an interactive interface to link notes to digital media, or trigger specific digital applications on a separate computer, providing a natural integration of paper and digital facilities [Weibel et al. 2011]. With our study of aviation human factors, this capability allows an observer to record flight events in their notes on paper, then automatically link these observations directly to video and eye tracking data for analysis of attention during the events.

While the array of data streams supported by ChronoViz before this effort greatly supported behavioral research and resulted in important advances in many of the research domains in which it had been used, we were interested in understanding the dynamics of visual attention on the flight deck and therefore decided to integrate support for mobile eye tracking into this analysis and visualization tool. In the next section we briefly introduce mobile eye tracking systems, while in the remainder of the paper we focus on the integration of eye tracking data into ChronoViz. We focus our discussion particularly on the aviation human factors domain, but we believe that our findings are valid beyond this domain in many of the challenging research settings where eye tracking is analyzed in combination with other data sources.

[1] http://www.anoto.com

[2] http://www.livescribe.com

4 Mobile Eye Tracking System

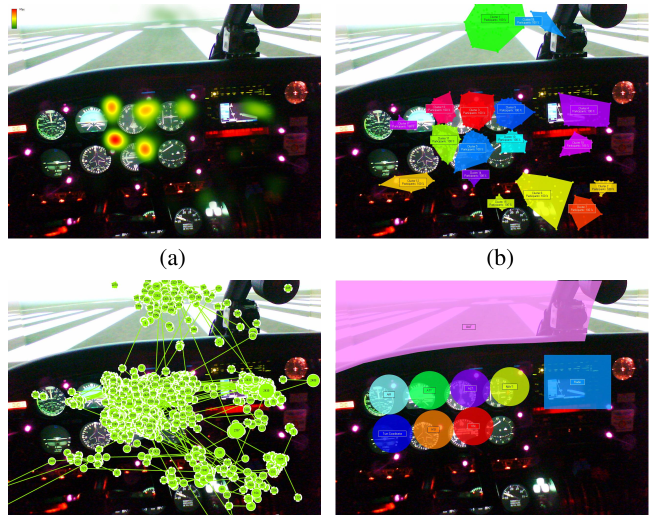
On-going research collaborations with a local flight training school and a major US-based airframe manufacturer provided us access to high-fidelity fixed-based Piper Seneca and wide-body airliner simulators. Because we needed to move the data collection apparatus from simulator to simulator and were not allowed to modify the simulators, we chose a wearable eye tracking system – the Tobii Glasses System [3]. The comfortable and unobtrusive form factor of the glasses made this system particularly suitable for use in a flight simulator. The wearable platform consists of a pair of lightweight glasses connected to a small support processor that records the data. The glasses have a forward-facing scene camera that records the participant's field of view and a microphone to record audio. The lenses are made of infrared reflective glass, so light reflected off the retina and through the pupil can be detected by the eye tracking camera. The position of the pupil reflection is registered with respect to the scene camera, giving an (x, y) coordinate for where the eye gaze is directed at any given time. A pilot can wear the glasses and fly the simulator without being distracted by them. In a motion-based simulator, the glasses do not move around on the pilots head and there is no significant mass far from the face that might cause unexpected forces on the pilots head while flying. The one important limitation on the use of this system is that the subject cannot wear other glasses while wearing the eye tracking glasses.

In addition to the glasses and support processor, the Tobii system makes use of an infrared (IR) marker system. The IR markers are not worn, but are placed in the environment in which the subjects will be working. The markers are small, highly portable, and can be installed and removed in minutes. IR markers enable the eye tracking system to achieve registration of the eye data onto multiple planes of the visual scene in the flight deck without making modifications to the flight simulator. An infrared camera faces forward and is located below the scene camera. Through this camera, the glasses can detect and identify IR markers in real time, and record their position with respect to the scene camera. In post-processing, the positions of the IR markers can be used to transform the gaze locations from coordinates defined with respect to the scene camera to coordinates defined with respect to an "Area of Interest" (AOI) that has been identified by the researcher in the space defined by the positions of the IR markers. The analysis software provided with the Tobii Glasses system – called Tobii Studio – can export both the raw eye tracking data and the AOI-mapped data for further processing and visualization, e.g. in ChronoViz.

5 Visualizing Eye Tracking Data

The Tobii glasses system produces a record of eye position and can superimpose a representation of eye gaze location on the output of the built-in scene camera. The system samples eye position at 30 cycles per second. As soon as the data is loaded into Tobii Studio, it is possible to view a "playback" video that shows raw eye gaze position superimposed on the view from the scene camera. A number of interesting visualizations are available in Tobii Studio including heat maps (Fig. 2a), clusters (Fig. 2b) and bee swarms (Fig. 2c).

The heat map visualization shows the accumulated gaze data over the course of a whole flight. This visualization illustrates a gaze offset error that is further discussed in Section 5.1: there is a "hot" region just to the right of and above each of the instruments that received large amounts of visual attention. The bee swarm representation shows individual eye gaze fixation events with transition lines between successive fixations. This shows that most eye gaze



(a) (b) (c) (d)

Figure 2: *Eye tracking data visualization tools in Tobii Studio. (a) Heat maps. (b) Clusters. (c) Bee swarm. (d) AOI positioning.*

fixations were directed to the flight instruments or out the front window. The cluster representation uses the statistics of the eye gaze fixations and transitions to compute clusters of associated eye gaze events. The fact that this clustering algorithm found distinct clusters corresponding to each of the flight instruments is evidence that this pilot's visual attention is reasonably well adapted to scanning the flight instruments. This representation picks out and renders visible relations among eye gaze events that are not visible in the heat map because they are not frequent enough and that cannot be discerned in the bee swarm representation because they fall in regions that are crowded with fixations.

5.1 AOIs

The IR markers provide stable landmarks in the scene so that the system can compute where the eye gaze falls in the world that is viewed. Either before or after the data collection session, a still image is captured using the scene camera. AOIs can then be defined on this image. Figure 2d shows the AOIs defined for analyzing visual attention on the flight instruments. In this session, we defined nine named areas of interest, each one covering a meaningful target of visual attention. Because these still images also contain the IR markers (bright dots in the image), the AOIs can be associated with regions in the local space. As long as four or more IR markers are in the field of view of the infrared camera, the system can locate the eye gaze relative to the frame of reference defined by the markers.

We noticed that even after a successful calibration of the glasses there can be a systematic difference between the actual eye gaze location and the measured eye gaze location. In order to get a measure of this error we created a dynamic visual stimulus as shown in the left image of Figure 3. The subject tracks a moving target – a green dot – that moves around a computer display. It can be seen in the heat map example shown in the right panel of Figure 3 that the recorded gaze positions can be offset about one degree of visual angle (in this case to the right). A bias toward gaze events on the right side of the display is also apparent. This seems to be caused by the fact that eye gaze is recorded for the right eye only, and the cameras are located on the right side of the glasses.

While working with AOIs, we use the gaze offset error determined by the subject's tracking of the dynamic visual stimulus to offset the AOIs in a way that matches the gaze offset error. Thus, for the

[3] http://www.tobiiglasses.com/scientificresearch

subject whose gaze data is shown in Figure 2, the AOIs are offset to the right and above (see Figure 2d). When this is done, the AOIs will register gaze events that fall on the correct regions of interest rather than on the measured locations.

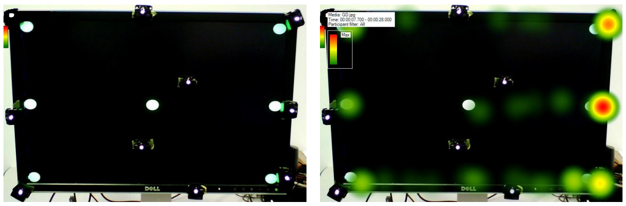


Figure 3: *Calibration procedure on a desktop display. Left: dynamic visual stimulus shown at fixed positions on the screen. Right: heat maps of a subject performing the calibration; note the small, but significant gaze offset error to the right of the visual stimulus.*

6 Temporal Dynamics of Visual Attention

Flying an airplane is a complex dynamic activity that can require hours of sustained attention. Pilots must continually solve problems of attending to the appropriate information in the flight deck environment and disregarding that which is not relevant. Our research therefore focuses on the temporal dynamics of visual attention. The visualizations generated by Tobii Studio provide static summaries of the collected eye tracking data, but they do not support the visualization and quantification of the dynamics of visual attention.

We experimented with importing the Tobii Studio visualizations into ChronoViz as separate video files. However, at present, this process is not well supported. The glasses system is not designed to collect data over time spans exceeding an hour and the Tobii Studio software was only recently configured to support the analysis of data produced by the mobile eye tracker. In order to overcome these obstacles, we developed dedicated eye tracking tools and visualizations as part of ChronoViz. In this section we show how these tools support data analysis and visualization ultimately allowing us to explore the temporal dynamics of visual attention on the flight deck.

6.1 Overlaying Eye Gaze

Instead of using Tobii Studio to process the eye tracking data, we used it as a simple tool to export the recorded data from the mobile eye tracking system into a universally readable text file[4]. The exported data file contains the location of the eye gaze with respect to several frames of reference including that of the glasses, the one of the scene in front of the glasses, and one defined by the locations of the IR markers.

A difficult challenge for visualizing eye tracking data is integrating the spatial and temporal dimensions of the data. A straightforward visualization of the spatial dimension is to simply overlay the gaze points on an image of some portion of the field of view of the participant – such as the instrument panel. This will produce a representation that lends itself to investigation of where the participant looked, but does not lend itself to understanding the temporal pattern of gaze fixations. In order to understand how visual attention is correlated with other behavior, the temporal dimension of the data must be considered. In this case, it may not be enough to know how much time was spent looking at a particular instrument, or even when that instrument was the focus of gaze. We also need to know where the pilot looked before and after that particular instrument, what else the pilot was doing when looking at the instrument, and whether this is a common or uncommon pattern.

We developed a visualization in ChronoViz that, like the basic visualization of Tobii Studio, supports superimposing the gaze data on the video from the eye tracking scene camera. This is accomplished by first importing the scene-camera video into ChronoViz, then importing the gaze point data from the data file exported from Tobii Studio. Figure 4 shows the resulting visualization, a dynamic representation of the participant's gaze for any point in time during the flight. The video with superimposed gaze data can be controlled via the ChronoViz timeline, allowing the researcher to control the gaze data in the same way as a video by playing, pausing and dragging on the timeline.

The great advantage over the existing tools is that the visualization of gaze points in ChronoViz is configurable for a range of different analysis possibilities. The default representation shows the gaze point for the given video frame, represented with an outlined dot, as well as a half-second of gaze data leading up to the current frame, as represented by a line leading to the dot. However, the researcher can change the duration of the line leading to the current point, and specify whether gaze data is accumulated within the visualization, and can modify the color and opacity of the gaze data visualization. For example, longer lead times can be used to get an overall impression of the gaze pattern of the pilot.

6.2 AOI Time Series

Another approach to visualizing eye tracking data is to focus on activity as it relates to the AOIs. If AOIs have been defined in Tobii Studio, in addition to the raw gaze data, the exported data file also includes every eye gaze hit within the defined AOI. We created a simple application that can read the exported gaze data file, and create a file (in comma-seperated-value format) containing a set of binary time series for each of the AOIs defined in Tobii Studio.

Once activity over pre-defined AOIs has been calculated, this data can be visualized in ChronoViz for each AOI, indicating when gaze

Figure 4: *Eye gaze visualization in ChronoViz. The bright red square indicates the current position of the eye gaze, while the matte traces represent the path of the eye gaze during the preceeding 5 seconds.*

[4]While we used Tobii Studio to do this, in the near future we expect to export the raw data via the newly released Glasses SDK, thus completely bypassing Tobii Studio.

is directed toward that area. The time-series representations are designed to allow the researcher to get an overview of the pattern of gaze over a number of AOIs. When imported into ChronoViz, the exported data produces a timeline for each AOI and puts a series of color bars on the time line. With this colored bar representation, bars of color represent a point in time where a participant directed his gaze toward that specific area (see Figure 5).

Sustained gaze is represented by larger areas of color, and due to the integrative nature of the human perceptual system these appear brighter than the quick glances represented by a single column of colored pixels. The colored bar representation is a spatially efficient representation, which allows the time-series for multiple areas to be shown together. This allows researchers to easily and flexibly get access to different patterns of allocation of visual attention, such as looking for periods of time when there was gaze activity in related areas.

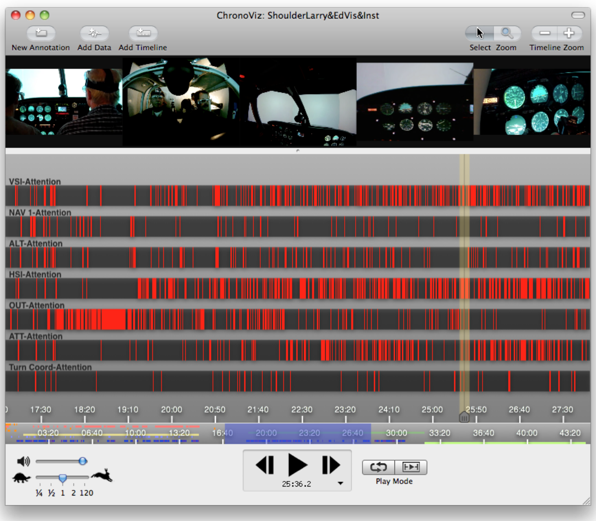


Figure 5: *AOIs defined in Tobii Studio are visualized as time-series in ChronoViz. Each AOI is represented on a different timeline. The red bars represent the single hits of the eye gaze on this AOI. Each timeline bar provides an impression of the intensity of visual attention to an AOI defined region across the duration of the recording.*

6.3 Dynamic AOIs

One of the limitations of defining AOIs in Tobii Studio is that they are fixed, and redefining or simply moving them requires an additional round of processing of the eye tracking data. If AOIs have to be visualized in ChronoViz, this is further complicated by an additional long export process, which discourages performing exploratory analysis. We therefore introduced a novel way of creating *dynamic* areas of interest directly within ChronoViz. By exporting a fixed frame from the scene camera that contains the visual field under analysis (e.g. the instrument panel) and the IR markers used to map the scene camera to that visual field, it is possible to automatically overlay the gaze data on top of the exported frame. The frame has to be defined only once in Tobii studio prior to exporting the raw gaze data. This data file also contains a *mapped* (x,y) position with respect to the selected frame, and ChronoViz is then able to automatically overlay the gaze position on top of this frame for any given time in a dedicated panel (Figure 6). This representation combines spatial selection and colored bar timelines. The analyst can select spatial regions over the imported frame in which the AOIs are defined, simply by using the mouse to drag a rectangular selection around that region. This dynamically creates a colored bar

Figure 6: *Dynamic AOIs. The analyst selected eye gaze hits out the front window for a flight in instrument meteorological conditions (no view except on takeoff and landing). The timeline shows intense visual attention early in the flight, during the takeoff roll, and late in the flight, during the last moments of approach and the landing.*

timeline below the image, showing the times when the participant's gaze entered the selected region. The linkage between visualization elements provides additional benefit because the timeline is interactive. Users can click any of the gaze occurrences shown in the timeline to navigate to the point in time when this happened, as reflected in linked video and raw gaze position visualizations. This interactive visualization, combining spatial and temporal visualization, gives researchers the ability to quickly see a temporal gaze pattern for any region, investigate details of activity when gaze was active in that region, and fine-tune the selection based on the results.

6.4 Computer Vision

Ultimately, what matters to our research is not simply where the pilot was looking, but what the pilot was looking at. Even with the introduction of dynamic querying capabilities, the mapping of gaze data to particular regions of interest is still dependent on IR markers. In our research we often faced challenges in placing the markers in the field of view of the subject in such a way that they cover the plane that we wanted to analyze. The IR markers also suffer from other problems, such as partial coverage of the visual field (i.e less than four IR markers visible), irregular surfaces (such as the instrument panel of an airplane) that make secure attachment of IR markers difficult, and motion of the surface to which the markers are attached (as happens in flight simulators and actual airplanes; IR markers can shake off the instrument panel). Furthermore, identifying where the pilot is looking is sufficient to establish what the pilot is looking at only in environments where particular objects of interest have fixed locations in space. In environments with dynamically changing display content (as encountered on the displays in modern airplanes, but also like a normal computer display), knowing where the pilot is looking does not necessarily provide information about *what* the pilot is looking at. In all of these cases the IR marker system cannot provide a complete solution to the problem of automatically detecting when gaze falls on particular objects of interest.

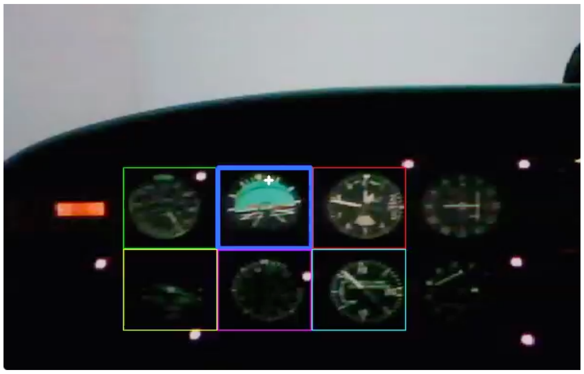


Figure 7: *Using computer vision to achieve eye gaze-to-object registration. Template matching routines find the flight instruments in each scene camera video frame and draw the outlines of the objects of interest (in this case, the six outlined instruments). Eye gaze coordinates are superimposed on the image (shown as a white cross at the top of the top center instrument).*

For this reason we introduced an alternative approach that applies computer vision techniques [5] to achieve eye gaze-to-object registration. This approach uses only the unadorned scene camera video and the raw gaze data. A template of a desired object is derived from a selected single frame of the scene camera video. This template is then matched against every frame in the entire scene camera video. If the match of the template to a region of the scene camera frame exceeds a threshold, then the object is detected and its coordinates in the current video frame are noted. The template matching mechanism we used is tolerant of changes in illumination and small rotations, and combined with the constraints of the cockpit environment enable the detection of the desired object even when the head is moved. Very rapid head motion can cause blurring in the captured frames, which may cause the template match to fail. As soon as the video stabilizes, however, the match will again succeed. A figure (rectangle in this case) can be drawn on the frame to surround the detected object. If the raw eye gaze coordinated for that video frame fall within the bounding figure of the object, then the gaze of the pilot is on the instrument. We increase the thickness of the bounding figure to emphasize the eye gaze-to-object registration. The annotated video frames are then written to a new movie file that includes the eye gaze and object bounding figures (see Figure 7). This video can be imported to ChronoViz, or the data for hits in each AOI can be exported to a data file and then visualized in ChronoViz in exactly the same way as IR marker derived AOI data are visualized as described in Section 6.2.

The first advantage of this approach is that it does not require IR markers, and therefore avoids the cost of IR markers and all of the problems associated with the installation of the markers and the processing of IR marker data. Second, the technique does not require that objects of interest have fixed locations. The computer vision template recognition routines work with dynamically changing display content, and even with moving displays. We have used this technique to detect a moving hand-held Apple iPad displaying navigation charts and checklists in a flight deck. Finally, the computer vision techniques are quite general and could be used to detect a wide range of mobile objects of interest including the faces of other workers in the environment.

[5] We used the open source library of computer vision routines provided in OpenCV: http://opencv.willowgarage.com.

7 A Moment of Flight

There are many dimensions to pilot cognition. We have been mapping pilot attention because it is a key component of cognition and because many aspects of attending activities, unlike some other aspects of cognition, are directly observable. The great contribution of eye tracking technology to this kind of work is that it provides a rough index of visual attention. Looking is an activity and eye gaze data provide a way to measure the activity of looking. Of course, attention is multimodal. In addition to visual attention, people attend to the settings of their activity via the other senses. The video recordings allow us to observe many of the ways the body is used to attend to the world. The audio recordings allow us to observe verbal attention in the form of speaking about objects and events. We cannot directly observe auditory attention, but we can often infer that it was active by observing what happened after an opportunity to hear something.

We recently observed and recorded two pilots performing a simulated thirty-six minute flight from Seattle to Portland in Boeing's new 787 Dreamliner. In order to capture data on the dynamics of flight crew attention during the flight we employed two mobile eye tracker systems, several digital pens, three HD video cameras and two microphones. This resulted in a ChronoViz session with four video streams documenting the activity of the pilots, two additional video streams representing attitude of the airplane and the instrument console, two eye tracking panels, seven paper-digital notes panels, and the digital airplane state data collected from the simulator. Just minutes after data collection was completed, we were able to begin the analysis of the pilot's activities and their visual attention by integrating the collected data in ChronoViz. We were then able to use any of the data streams to navigate the entire data set to particular points of interest. To illustrate the utility of integrating eye tracking into ChronoViz for observational research on the flight deck we present an analysis of about six seconds in duration from the thirty-six minute flight. Following Boeing usage, we refer to the pilots by their primary roles: Pilot Flying (PF) and Pilot Monitoring (PM). In this brief excerpt we examined in detail how the crew distributed their attention among three tasks: monitoring the flight path, programming the flight management computer (FMC) for the arrival at the destination airport, and configuring the flight deck for ongoing activity.

Let us begin with a consideration of the allocation of multimodal attention by an individual pilot. At time 30:01.5, the video recordings show that the PF lightly touched the front edge of left thrust lever with the side of the pinky finger on his right hand, bumping the lever lightly in the direction of reduced thrust. What does this action by the PF mean? We know that this is not an instrumental action because the touch is too light to move the lever. The simulator data stream shows that both autopilot and auto-thrust systems were engaged and the video of the flight instruments shows that the airplane was just reaching and capturing the cruise altitude at the moment in question. Our rich ethnographic background tells us that every pilot knows that thrust must be reduced when a jet airplane transitions from climb to level flight. We know that the airplane was capturing the altitude and the PF was attending to thrust, but lacking evidence of pilot attention to the altitude capture, we cannot know that the PF's bodily attention to thrust lever position was related to the altitude capture event. The eye tracking data fills this gap, clearly showing the PF looking at three key indications of capturing the cruise altitude. He looked at the altimeter and vertical speed indicator while touching the thrust lever, and he looked to the airspeed indicator just after. The eye tracking data thus establish visual attention to the altitude capture event while the video shows bodily attention at the same moment related to thrust reduction. Without the eye tracking data, or lacking a means to integrate and coordinate that data with other data streams, there would be no way to be sure what this bodily action is. With that integration we can not only say what activity the gesture is part of – monitoring the level off – we can also identify its meaning in the context of that activity – an expression of expectation concerning the reduction of engine thrust.

Because we recorded this information for both pilots, we can also characterize and visualize multiparty multimodal attention. Ensuring the airplane follows the planned flight path is the responsibility of the PF and is primarily a matter of visual attention. However, the eye gaze data show the PM also monitored the flight path by directing his gaze to the altitude, vertical speed and airspeed early in the segment. Programming the future flight path is the responsibility of the PM. However, we can show the PF also contributed visual and verbal attention to this task. Because the airplane is being flown by the autopilot and auto-thrust systems, which are following a three-dimensional path programmed into the FMC, the crew does not need to attend to the control of the airplane's trajectory at all times. At time 29:55.8, the PM was simultaneously handling flight charts, directing visual attention to a particular flight chart, and talking about the data on the chart. While holding the arrival chart in front of his face, the PM said, "After Helens, then we are going to go to Battleground." The eye gaze data reveal a within-individual dynamic configuration of visual and verbal attention in the inspection of the chart and the construction of the utterance. The PM's eye gaze hovered around the depiction of HELNS waypoint when he said "helens." His eye gaze moved down the chart along the line depicting the route of flight to the Battleground waypoint when he said, "then we are going to go." Finally, his eye gaze hovered around the data block adjacent to the chart depiction of Battleground when he said, "to battleground."

The dual eye tracking data integrated into ChronoViz also permit us to describe the allocation of visual attention by the flight crew as a system. We use the video record to navigate to the moment of interest and then inspect the eye tracking displays for both pilots, which ChronoViz maintains synchronized with all other data streams. While the PM was reading the chart, the PF was directing visual attention to the display where the PM will enter the navigation information into the FMC. In particular, the PF looked at the identifier of the waypoint named by the PM. The PM's spoken utterance thus coordinated the allocation of visual attention by the two pilots to two different representations of the same entity. This illustrates a multiparty multimodal attention configuration. While the PM's visual attention was on the source of navigation information, the PF's visual attention was already on the destination for that information. This allocation of attention to the next locus of action in the ongoing navigation activity is evidence of the pilots' joint participation in, and construction of, a shared activity. The PF's eye gaze anticipates the PM's next action in the activity, which is entering the identifier for the waypoint called Battleground into the route. Further evidence of anticipation and participation in a jointly constructed activity comes from the PF's next utterance which he produced as the PM repositioned his body and hands to touch the keypad to enter the waypoint identifier. The PF spoke the three letters that compose the identifier for the Battleground waypoint. These are the letters subsequently entered into the FMC by the PM.

In both of these examples, the integration of eye gaze data with other data in ChronoViz permits us to see phenomena that would otherwise be invisible and to perform analyses that would otherwise be impossible. Human factors research has long sought a way to conceptualize the cognitive properties of operational teams rather than of individual operators. By characterizing who attends to what, when and in which modalities we have at last achieved a rough description of the cognitive properties of the flight crew system. Eye tracking technology has made the visual component of attention available for integration with indices of other modalities of attention. We expect that developing the characterizations of multiparty multimodal attention will contribute to the refinement of theoretical conceptualizations of the cognitive properties of working groups.

8 Discussion and Conclusion

Mobile eye trackers and analysis toolkits are the first step in mapping visual attention. However, in order to quantify and visualize the *dynamics* of visual attention and to integrate this mapping with other aspects of attention such as haptic attention or auditory attention, the handling of the eye tracking data must be extended beyond the capabilities of current technologies. More flexible tools supporting exploratory data analysis are needed. For example, in aviation human factors most eye tracking research has focused on the visual activities of a single pilot monitoring flight instruments. This is necessary, of course, but in highly automated modern airliners it is of decreasing relevance. In the automated flight deck eye gaze is much more variable and more concerned with coordination among the crew than simply with monitoring the flight trajectory. Understanding eye gaze in the modern flight deck *requires* the integration of eye tracking data with other data streams in the context of a wide variety of flight deck activities.

We integrated an off-the-shelf eye tracking system into an experimental tool for supporting the simultaneous visualization of multiple streams of time series data. The pairing of mobile eye tracking with ChronoViz resulted in an extremely powerful tool to analyze complex dynamics during long-duration distributed and collaborative activities. While integrating eye tracking data into ChronoViz, we also developed new ways to visualize the dynamics of visual attention and developed interactive tools that free the researcher from some of the constraints of automatic gaze logging based on predefined AOIs.

Of course, eye tracking can contribute to observational research without automated eye gaze-to-object registration. For example, the analysis presented in Section 7 relied on human coding because the objects of interest, characters on a chart or on a computer display, cannot be automatically recognized using the present technology. In the long run, however, automated eye gaze-to-object registration will be an important element of observational research for several reasons. First, it makes possible quantification of gaze events over large data sets. Second, when combined with ChronoViz, it supports the visualization of patterns of visual attention on multiple time scales, thus facilitating the navigation of complex data sets and supporting exploratory analysis. Exploratory analysis is particularly valuable in light of the absence of good theoretical understandings of the new phenomena that are revealed by juxtaposing multiple streams of behavioral data. Finally, automatic coding of eye gaze-to-object registration can provide input to algorithms that filter large data sets for moments of interest to a particular analysis project. The analysis of a moment of flight presented in Section 7 is an initial hand-built attempt to map the multiparty multimodal allocation of attention in the flight deck. We are excited about the prospect of developing a robust suite of tools and techniques to support this important class of behavioral analysis.

Observational researchers are exploiting the growth of sensor technology to produce dense time-series measurements of many aspects of behavior including eye gaze. Eye tracking data has a special place in research on expert human performance because knowing where to look, when to look, and what to see when looking are key elements of expertise. We hope we have shown the value of eye tracking data in observational research by integrating it with, and visualizing it in the context of, many other behavioral data streams.

Acknowledgements

We thank all of the member of the DCog-HCI lab for their help in performing this research. Many thanks also go to our collaborators in the aviation domain for giving us access to simulators and cutting edge research on concept flight decks. Support for this research was provided by NSF award #0729013, and by UCSD-Boeing project agreement 2011-012.

References

BJÖRKLUND, C., ALFREDSON, J., AND DEKKER, S. 2006. Mode monitoring and call-outs: An eye-tracking study of two-crew automated flight deck operations. *The International journal of aviation psychology 16*, 3, 257–269.

DUCHOWSKI, A. 2002. A breadth-first survey of eye-tracking applications. *Behavior Research Methods 34*, 4, 455–470.

DUCHOWSKI, A. 2007. *Eye tracking methodology: Theory and practice*. Springer-Verlag New York Inc.

FLETCHER, L., AND ZELINSKY, A. 2009. Driver inattention detection based on eye gazeroad event correlation. *The International Journal of Robotics Research 28*, 6, 774.

FOUSE, A., AND HOLLAN, J. D. 2010. DataPrism: A tool for visualizing multimodal data. In *Proc. Measuring Behavior 2010*, 1:1–1:4.

FOUSE, A., WEIBEL, N., HUTCHINS, E., AND HOLLAN, J. D. 2011. Chronoviz: a system for supporting navigation of time-coded data. In *Extended Abstracts of CHI 2011*, 299–304.

HAGEDORN, J., HAILPERN, J., AND KARAHALIOS, K. 2008. Vcode and vdata: illustrating a new framework for supporting the video annotation workflow. In *Proc. AVI 2008*.

HAMMOUD, R., SMITH, M. R., DUFOUR, R., BAKOWSKI, D. L., AND WITT, G. J. 2008. Driver distraction monitoring and adaptive safety warning systems. In *Proc. SAE Commercial Vehicles Engineering Congress and Exhibition*.

HUEMER, V., HAYASHI, M., RENEMA, F., ELKINS, S., MCCANDLESS, J., AND MCCANN, R. 2005. Characterizing scan patterns in a spacecraft cockpit simulator: Expert versus novice performance. In *Proc. Human Factors and Ergonomics Society Annual Meeting*, vol. 49, 83–87.

HUTCHINS, E. 1995. *Cognition in the wild*. The MIT Press.

HYMES, D. 1962. The ethnography of speaking. *Anthropology and human behavior 13*, 53, 11–74.

HYRSKYKARI, A. 2006. Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading. *Computers in human behavior 22*, 4, 657–671.

LAND, M., AND HAYHOE, M. 2001. In what ways do eye movements contribute to everyday activities? *Vision research 41*, 25-26, 3559–3565.

PEA, R., MILLS, M., ROSEN, J., DAUBER, K., EFFELSBERG, W., AND HOFFERT, E. 2004. The diver project: Interactive digital video repurposing. *IEEE MultiMedia 11*, 54–61.

RÄIHÄ, K., AULA, A., MAJARANTA, P., RANTALA, H., AND KOIVUNEN, K. 2005. Static visualization of temporal eye-tracking data. *Proc. INTERACT 2005*, 946–949.

RAMANATHAN, S., KATTI, H., SEBE, N., KANKANHALLI, M., AND CHUA, T. 2010. An eye fixation database for saliency detection in images. *Computer Vision–ECCV 2010*, 30–43.

RODA, C., AND THOMAS, J. 2006. Attention aware systems: Theories, applications, and research agenda. *Computers in Human Behavior 22*, 4, 557–587.

RODA, C. 2011. *Human Attention in Digital Environments*. Cambridge Univ Pr.

SACKS, H. 1995. *Lectures on conversation*, vol. 1. Blackwell.

SARTER, N., MUMAW, R., AND WICKENS, C. 2007. Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *HUM FACTORS 49*, 3, 347.

SCHNELL, T., KELLER, M., AND POOLMAN, P. 2008. Neurophysiological workload assessment in flight. In *Proc. IEEE Digital Avionics Systems Conference, 2008.*, 4.B.2–1 – 4.B.2–14.

ŠPAKOV, O., AND MINIOTAS, D. 2007. Visualization of eye gaze data using heat maps. *Electronics and electrical engineering 2*, 55–58.

STEELMAN, K., MCCARLEY, J., AND WICKENS, C. 2011. Modeling the control of attention in visual workspaces. *HUM FACTORS 53*, 2, 142.

TATLER, B., HAYHOE, M., LAND, M., AND BALLARD, D. 2011. Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision 11*, 5.

THOMAS, L., AND WICKENS, C. 2004. Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. In *Proc. Human Factors and Ergonomics Society*, vol. 48, 223–227.

TSANG, H., TORY, M., AND SWINDELLS, C. 2010. eSeeTrack – Visualizing Sequential Fixation Patterns. *IEEE Transactions on Visualization and Computer Graphics 16*, 6, 953–962.

TUKEY, J. 1977. Exploratory data analysis. *Addison-Wesley Series in Behavioral Science: Quantitative Methods, Reading, Mass 1*.

VERTEGAAL, R. 1999. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In *Proc. CHI '99*, 294–301.

WEIBEL, N., FOUSE, A., HUTCHINS, E., AND HOLLAN, J. D. 2011. Supporting an integrated paper-digital workflow for observational research. In *Proc. IUI 2011*, 257–266.

WICKENS, C., ALEXANDER, A., HORREY, W., NUNES, A., AND HARDY, T. 2004. Traffic and flight guidance depiction on a synthetic vision system display: The effects of clutter on performance and visual attention allocation. In *Proc. Human Factors and Ergonomics Society*, vol. 48, 218–222.

WITTENBURG, P., BRUGMAN, H., RUSSEL, A., KLASSMANN, A., AND SLOETJES, H. 2006. Elan: a professional framework for multimodality research. In *Proc. Language Resources and Evaluation Conference (LREC)*.

ZHAI, S. 2003. What's in the eyes for attentive input. *Communications of the ACM 46*, 3, 34–39.