UAV Sensor Operator Training Enhancement Through Heat Map Analysis

Ashish Amresh, John Femiani, Jason Fairfield, Adam Fairfield
Arizona State University
amresh@asu.edu, john.femiani@asu.edu, adam.fairfield@asu.edu, jason.fairfield@asu.edu

Abstract

Heat map based data visualization and mining is an emerging area in game engine design and architecture. Employed by many state of the art game engines and popular commercial games, this technology helps populate and collate player activity and behavior to better inform the system for further action. Simulation and serious games can tremendously benefit by applying heat map based visualization for the purposes of analyzing and tracking player behavior. Heat maps are time varying texture maps that represent a chosen activity over a certain grid at any particular interval of elapsed time. In this paper results of applying a real-time heat map data capture and generation tool on two military simulations: 1) Ground-based combat scenario and 2) Unmanned Aerial Vehicle sensor operator scenario is presented. The research showcases several real-time visualization techniques developed into the simulation with the main goal of understanding participant behavior. Novice and expert data is populated as part of the experiment to validate the effectiveness of our methods.

Keywords—**Serious Games, Heat Maps, Visualization, Training**

1 Introduction

Recently, unmanned aerial systems (UAS's) have been increasingly important for both government and commercial applications such as intelligence gathering or map making. However, unlike a traditional aircraft, a UAS is operated by a team of people including pilots and sensor operators who are remotely located, using an interface to control the system that is similar to a video game. This can introduce additional challenges to the UAS operator, but it also introduces an important opportunity to use tools, such as heat map visualizations which have been used in the past to analyze video game play, in order to provide enhanced training for UAS operators.

In this work we propose a novel use of heat-maps for training by comparing novice users to expert users in order to learn where they differ, and how they evolve over time. The proposed use of heat maps will allow trainers to understand how sensor operators learn, and how variations in sensor operation correlate with experience. Our contribution is a visualization tool that can demonstrate the differences between expert and novice sensor operators of a UAS.

1.1 Related Work

Heat maps have been a key instrument employed by game developers to understand and track player behaviors. Typically employed during the testing phase they have become a useful tool to fine-tune games and remove bottlenecks originating from bad game mechanics and design decisions. Heat maps were employed in the original Halo [8] game launched with the first XBOX console, where they provided a deep understanding of optimal weapon placement.

Heat maps can be used to visualize a large range of measured information, but one very common use is in mapping user gaze with eye tracking technology. In a 2 dimensional environment (such as a website or image) it is quite simple for overlaid heat maps to provide information on areas users find interesting as well as ignored areas. For added effect, these 2D heat maps have also used the heat data to affect image opacity to reveal visited areas and obscure overlooked sections with either shadow or fog [6].

The use of heat maps for a three dimensional scenes introduces additional challenges. Stellmach, Nacke, and Dachselt suggest that traditional 2D methods for using heat maps would result in data loss when used on a 3D scene, since multiple parts of the scene could project to the same point in a 2D map [7]. Three methods were suggested to overcome this challenge: Object-based maps accumulate heat data per object and show each object in a single color. Surface-based maps work in a similar way, but they accumulate heat data for every mesh polygon in the scene and assign a single color to each face. Finally, the type of map used in this paper is a projected heat map. This approach uses a texture map and projects each point on the object's surface to it corresponding texture coordinates in order to accumulate data onto a 2D texture map.

Heat maps have been used for quick visualization of data in analytics since their inception. However, they have not been widely used to facilitate streamlined training as is proposed in this paper. Similar methods for heat map compo-

1550-6037/13 $26.00 © 2013 IEEE
DOI 10.1109/IV.2013.60

sition and analysis are used to study such things as crowd motion analysis [3]. This study uses accumulated motion heat maps to quickly identify regions of interest for further analysis. Different sets of heat maps were used for different conditions and were accumulated over long runs [2]. Training analysis has also been studied using various means, but none have fully taken advantage of the power heat maps can bring to the field. An example is a study that compared novice and expert eye gaze patterns during a laparoscopic surgery simulation [4]. Similarly to our project, this study tasked novice and expert surgeons to perform a simulated surgery while various data points were collected. This is where heat maps could have shown differences in novice/expert gaze positions over time, much like the several heat maps that were populated in our project.

1.2 Organization

The remainder of this work is organized as follows: In section 2 we give a detailed description of the software architecture and technology developed to create several types of heat maps. In section 3 we discuss how the heat maps are used by the trainer to help improve training and understand the data coming out of the heat maps. In section 4 we give a complete breakdown of the software tools used to achieve the various needs of the project and discuss the production pipeline for developing a scenario. In section 5 we mainly discuss the results of having novices and experts interact with our UAS sensor operator scenario and the resulting improvements in the training process. In section 6 we look at future directions for the project and additional features that are being planned to continue the effort.

2 Heat Map Performance Tracking Architecture

Our approach concentrates on enhancing the training methodology for UAS camera operators by capturing trainee performance over time in a number of various heat maps. Furthermore, our system proposes keeping track of a separate set of training data performed by a separate class off users considered "expert". We refer to this distinction of expert heat maps as the oracle heat maps.

2.1 Heat Map Generation

Each heat map is stored as a 32 bit floating point raster image $hm_{(s,t)}$. After each frame is rendered, the texture coordinates at each rendered pixel are identified and the value stored in the heat map is incremented. The heat map is updated with each frame, so that the value at each texel is proportional to the amount of time that texel is visible. If the user has an eye-tracker, then pixels that are rendered on the screen but not included in the user's gaze are not updated. In order to visualize the heat-map, a pseudo-color image is generated from the recorded heat data, the scene can be redrawn with the pseudo-color image as a texture. The colors are chosen so that portions of the scene that receive less attention appear blue, whereas the portions of the scene that receive the most attention appear more red.

2.2 Heat Map Types

As a user navigates the simulation, our system automatically maintains different heat map visualizations on various aspects of their surveillance habits.

Positional: The simplest heat map is the positional map, which records the UAS's location relative to the simulated environment in the ground plane. This heat map is not of much significance for the current project as the pilot is on a rail or pre-programmed to follow a set path.

Pitch Angle: A group of heat maps provide views of the the amount of time spent with UAS camera at various pitch angles. Each heat map is populated by a different interval of the pitch angle $pitch_i$. Separating the pitch angle into multiple maps improves the resolution quality of the data and at the same time allows trainers to set calibrated intervals to gauge operator surveillance. Rather than writing a single pixel value for each frame, pitch angle heat maps store the camera's viewing frustum projected onto the environment below.

Zoom/FOV: Another set of heat maps record the camera's zoom level fov_i. This is accomplished in a similar manner that the pitch angle is stored where the zoom level is divided into strict intervals and the camera's projected viewing frustum represents the UAS.

Eye Tracking - Environment: A single projected heat map that provides a visualization of the locations most viewed by the player using pupil gaze tracking information. Instead of updating the heat map for every pixel rendered to the screen, the updates are limited to pixels that fall within a limited distance from the user's gaze. This provides a more specific view of the areas surveyed by the player than the pitch or FOV heat maps can show.

Eye Tracking - Screen: In addition to heat-maps that track which part of the 3D surface were visible, a single heat map is kept that shows information on the player's gaze throughout the session. However, this map leaves the pupil gaze location in screen space to determine where on the screen the player is looking most often. This can be used to show if the player is checking head's up display (HUD) elements frequently and how they scan the screen during surveillance.

2.3 Oracle Accumulation

For a given user, every successful simulation attempt outputs the above set of heat maps, set U as a unique attempt. This is in contrast to the centralized oracle set, O, which accumulates multiple simulation outputs. Instructors are provided with a supplementary tool that allows them to manage the oracle heat map set by adding results of a given attempt to the existing oracle set. The heat-maps in the oracle set are added together element-wise to produce a single cumulative oracle heat-map, which can be converted to a pseudo-color texture and rendered in a display alongside a student's heat-map.

3 Behavioral Tracking Architecture

In our UAS sensor operator scenario we modeled an area that was a sixty square kilometer grid near Baghdad, Iraq. We then developed a system that can easily drape building, trees, landscape over GIS data and create a 3D representation of this world. We developed a simple system for populating the scene with people and vehicles. This resulted in a rich and lively environment where we can embed many patterns of life situations for the sensor operator. These patterns of life scenarios were populated by polling domain experts in the United States military and collecting information on what they look for in sensor camera data feeds and what they consider to be suspicious. These patterns of life situations were translated into agent behaviors and the world was programmed to exhibit these behaviors. Since the training scenario was for sensor operators and not pilots, the UAS was on a rail and the sensor operator had a twenty minute cycle for surveying six hot spots in the area and identify and tag suspicious behavior.

In order to collect meaningful training data, an adequate training simulation must be concerned with not only the characteristics of the pilot/operator pair but also the traits of the accompanying environment and its occupants i. e. civilians, vehicles, buildings, etc. To this end, we maintain a set of behaviors that can be assigned to a given entity. entity. Behaviors are defined as small XML files and are complimented by various skeletal animation sets. Behaviors are first classified as one of two base types, either non-suspicious or suspicious. Behaviors are further distinguished by the explicit action being performed by the exhibitor, e. g. "running a wire spool along the roadside". Additionally, multiple behaviors may be chained together to add more realism. For instance, a suspicious person may walk normally with a group of civilians, but then break away to dig a hole for a roadside explosive device before returning to a crowd.

During training, the operator is tasked with identifying individuals on the ground who are performing suspicious acts. The trainee may tag individuals whom they believe are performing sufficiently suspicious behaviors. Once selected, the user chooses a behavior from a list of suspicious and non-suspicious visual cues that the suspect in question may be exhibiting. After a trainee has successfully completed the training simulation, their choices are compared with the correct results to grade users on their behavioral interpretation. While some users will have difficulty detecting certain behaviors, others may mark a non-suspicious behavior as suspicious. This lightweight system provides a flexible architecture for training observers to distinguish between the cultural patterns of civilians and the potentially combative stances of targets.

4 Technology Breakdown

This project involved development of two software applications: 1) A simulation that can be installed on any Windows PC that ran for twenty minutes and can be repeated several times to gain familiarity and improve the overall score. Scoring was a combination of correct identification of suspicious activity along with appropriate tagging. Figure 1 shows the simulation and the tagging interface used by the player. The square on the right corner represents the hot spots that the simulation is going to visit and its current location. The tagging menu is divided into four areas: IED placement, individuals, groups and other. Each area has several other sub choices to select from. 2) A viewer that trainers can use to look at player heat maps at a group or individual level that can further lead to improving certain errors that trainees might be consistently repeating at an individual or group level. Figure 2 shows the viewer tool used for comparing expert and novice heat maps. The numbers 1-8 represent the toolbar, the heat map type, directory to load the maps from, map navigation, cycle the novice and expert maps together, deviation of the novice from expert, the heat map image and lastly the tagging stats.

Figure 1: Top view of the game showing the tagging interface.

The world was created using a combination of a 3D terrain created from Google Earth [3], 3D models developed for the landscape elements, people and vehicles exhibiting patterns of life behavior using custom created animation and a simple interaction control set. The Delta 3D [1] game engine and its Stage editor were modified to suit the needs of the project for developing the scenario and rendering the final scene. A low cost eye tracker Gaze Tracker [5] was retrofitted to a baseball hat and used by the participant to provide eye tracking data. Gaze tracker has a documented accuracy of 89 percent which was more than sufficient for this project.

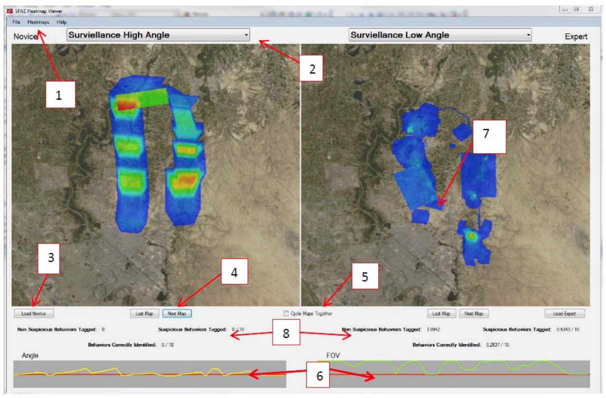

Figure 2: The Viewer application that is used to compare expert and novice maps.

5 Results and Analysis

The UAS scenario was tested by two sets of participants, a group of four sensor operators working for the air force (experts) and another group of 25 aviation undergraduate students (novices). Both the groups played the game three times and all their heat maps were accumulated and scores averaged for comparison purposes. Figure 4 shows the difference in surveillance captured from various senor camera angles between experts and novices. The colors indicate how long the camera spent on a certain point and as the time spent increases it transitions from blue to red. From these maps we can tell where the player is positioning the camera, for how long and at what angle. The expert heat maps indicate that trained operators spend more time at a high angle while scanning cities than novices. Experts also spend very little time in low angle and only over important areas while novices tend to spend more time at non critical areas and at low angles.

Figure 5 shows the camera'v view at different zoom levels. From these maps we can tell where the player is zooming in and how long they spend at each zoom level. The UAS camera should be zoomed in only when a suspicious activity has been spotted and the player needs to confirm and tag the activity. While the experts only zoom into the hot spots and areas of interest, the novices are zoomed in at many unimportant areas. Between cities and areas of interest the player should be zoomed out to maintain a maximum view of the terrain, however blank areas in the novice map show that they do not follow this rule. Figure 3 shows where the player is looking at the screen and if there were HUD elements then bloom spots should appear in those locations. Experts continually view the center of the screen and deviate for very short amounts of time where as novices have the tendency to look up and lack focus. Figure 3 also shows where the player's eyes are looking in the world using the eye tracker, this helps the trainers to find out where the player is actually looking as opposed to the camera. We observed that experts have a good coverage of the major hot spots and the places with suspicious activities while the novices lack focus and are all over the world.

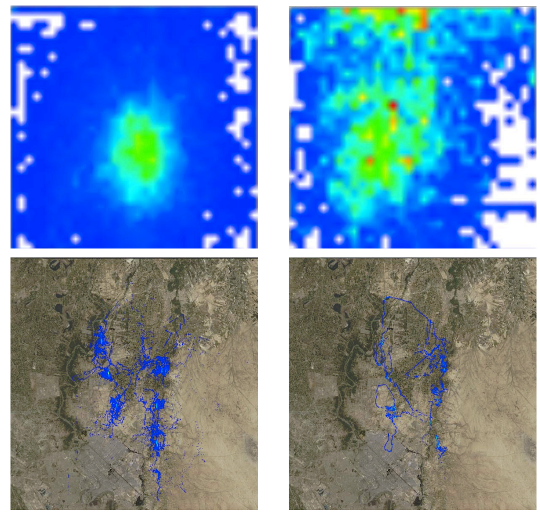

Figure 3: Gaze tracker and world gaze data for experts and novices

6 Conclusion and Future Work

We have presented our research in understanding how best to use heat maps for improving training for applications that rely heavily on multi function abilities of participants. We specifically chose the two scenarios, so that we could prototype and develop our methods in one and then apply those methods in the other. The UAS sensor operator is a multi function personnel who has to have a quick ramp up in abilities to perform critical tasks. The methods developed under the project can easily be incorporated into existing simulations and training systems. In future work we will develop additional tools and visualizations that can give immediate feedback to the trainees and guide their behavior on the fly. We are also developing methods that require group collaboration and teamwork. In the UAS world the pilot is not on a rail and is an integral part of tracking and identification of threats and we are hopeful to further develop our application to accommodate these needs.

7 Acknowledgements

We would like to thank Alion Sciences Corporation for their support in funding this project.

Figure 4: Surveillance angle at expert-high, expert-mid and novice-high.

Figure 5: Field of view for expert-zoomed out, expert-zoomed in and novice-zoomed out.

References

[1] Rudy Darken, Perry McDowell, and Erik Johnson. Projects in vr: the delta3d open source game engine. *Computer Graphics and Applications, IEEE*, 25(3):10–12, 2005.

[2] Nacim Ihaddadene and Chabane Djeraba. Real-time crowd motion analysis. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[3] John R Jensen. *Remote Sensing of the Environment: An Earth Resource Perspective 2/e*. Pearson Education India, 2009.

[4] Benjamin Law, M Stella Atkins, AE Kirkpatrick, and Alan J Lomax. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, pages 41–48. ACM, 2004.

[5] Richard R Ragland. Point-of-gaze tracker, November 28 1995. US Patent 5,471,542.

[6] O Špakov and D Miniotas. Visualization of eye gaze data using heat maps. *Electronics and electrical engineering*, 2:55–58, 2007.

[7] Sophie Stellmach, Lennart Nacke, and Raimund Dachselt. 3d attentional maps: aggregated gaze visualizations in three-dimensional virtual environments. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 345–348. ACM, 2010.

[8] Clive Thompson. Halo 3: How microsoft labs invented a new science of play. *Wired Magazine*, 15(9), 2007.