Aerial Border Surveillance for Search and Rescue Missions Using Eye Tracking Techniques

Qianwen Wan
Dept of Electrical and Computer Engineering
Tufts University
Medford, MA, USA
wendyqianwenwan@gmail.com

Aleksandra Kaszowska
Dept of Psychology
Tufts University
Medford, MA, USA
kaszowska.a@gmail.com

Arash Samani
Dept of Electrical and Computer Engineering
Tufts University
Medford, MA, USA
arash.samani@tufts.edu

Karen Panetta
Dept of Electrical and Computer Engineering
Tufts University
Medford, MA, USA
karen@ece.tufts.edu

Holly A. Taylor
Dept of Psychology
Tufts University
Medford, MA, USA
Holly.Taylor@tufts.edu

Sos Agaian
Dept of Computer Science
The City University of New York
New York City, NY, USA
Sos.Agaian@csi.cuny.edu

***Abstract*—Aerial border surveillance is a crucial activity, which can assure the security of the country boarder and aid in search and rescue missions. This paper offers a novel "hands-free" tool for aerial border surveillance, search and rescue missions using head-mounted eye tracking technology. The contributions of this work are: i) a gaze based aerial boarder surveillance object classification and recognition framework; ii) real-time object detection and identification system in non-scanned regions; iii) investigating the scan-path (fixation and non-scanned) provided by mobile eye tracker can help improve training professional search and rescue organizations or even artificial intelligence robots for searching and rescuing missions. The proposed system architecture is further demonstrated using a dataset of large-scale real-life head-mounted eye tracking data.**

Keywords—Head-mounted eye tracking technology, Aerial border surveillance, and search and rescue missions

I. Introduction

Air, land and maritime border security is responsible for stopping unlawful entry to the country and aiding in search and rescue missions [1]. Modern techniques are proposed for the purpose of detecting, tracking and identifying threat events such as acts of terrorism, as well as searching and rescuing people in danger (including mountain rescue, land and urban rescue, and air-sea rescue) [2]. Autonomous robotic systems, unmanned aerial vehicles (UAVs), camera and radar sensors are mainly utilized for border security search and rescue missions.

This paper introduces a novel "hands-free" tool for aerial border surveillance search, and rescue using eye tracking techniques. The development of wearable eye trackers opened up a way to research in unconstrained environment setting by allowing researchers to monitor eye movements (saccades and fixations) using infrared cameras and to capture the field of vision by a front-facing scene camera. Eye tracking technology can be used to investigate human visual attention and provide the information of users' interest. To be more specific, the regions where humans look at are indicated by gaze points computed on the basis of pupil position.

The contributions of the proposed new visual analytics approach include: i) a gaze based object classification and recognition framework for aerial boarder surveillance; ii) non-scanned regions real-time object detection and identification system; iii) a discussion of how the scan-paths (fixation and non-scanned) recorded by mobile eye trackers can help improve training professional search and rescue organizations or even artificial intelligence robots for searching and rescuing missions.

This paper is organized as following: section II briefly reviews the related work of head-mounted eye tracking technology. In section III, the proposed system architecture for aerial boarder searching and rescuing is presented. Example experimental results will be shown in section IV. Finally, section V summarizes the contributions of this work and discusses the future directions.

II. Related Work

A. Mobile Eye Tracking Technology

Research has been done over the past decades to understand the nature of human attention by analyzing eye movements using eye tracking techniques [1]. Current eye tracking techniques are used to monitor "where" and "what" people look at, which can be categorized into stationary remote eye trackers and head-mounted mobile eye trackers. The head-mounted mobile eye tracker opens a way to monitor eye movements and capture the visual scene at the same time in a unobtrusive environment.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: TUFTS UNIV. Downloaded on October 06,2020 at 15:39:35 UTC from IEEE Xplore. Restrictions apply.

Head-mounted eye tracking devices have been adopted in various research contexts and scenarios. For instance, head-mounted eye trackers are widely used in cognitive psychology research, such as understanding human problem solving and spatial learning ability [3]. Wearable eye trackers were utilized in natural tasks designed by a group of researchers from the University of Rochester and the Rochester Institute of Technology, such as a geology trip tasks and a grocery shopping experience trip [4-6]. Moreover, mobile eye trackers were adopted in driver training strategies and accident countermeasures [7]; marketing and advertisement research [8], usability research [9]; accessibility applications [10]; and education applications [11]. Fig. 1 shows the SMI head-mounted mobile eye tracker, infrared camera monitored eye image and front-face scene camera recorded gaze-overlaid video.

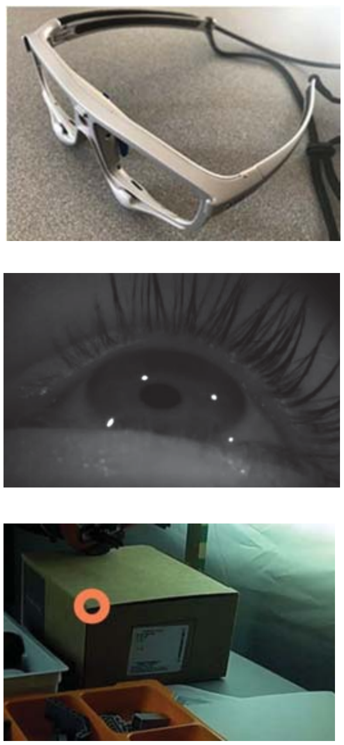

Fig.1 The SMI head-mounted mobile eye tracker(top) infrared camera monitored eye image (middle) and front-face scene camera recorded video with over-laid gaze(bottom).

Eye tracking technology offers possibility of a hands-free interaction with the physical world, which can be used for leveraging visual context for rescue missions, detection of unsafe behaviors or human action, and monitoring aerial border surveillance.

B. *Mobile Eye Tracker in Border security search and rescue mission*

Border security search and rescue missions in the United States are handled by professional law enforcement teams. They are usually accomplished by people scanning and searching from a certain distance in the sky. For example, air-sea search uses helicopters to search for survivors in the water or uses aircrafts to find suspicious or threatening people or boats crossing the country's border.

Oftentimes, sea rescuers are scanning areas that have the same background textures, such as sea rescues, which makes it difficult for the human eye to detect the object or person of interest. The individual visually surveying the scene from a helicopter may simply not recognize objects in the water as discernable objects, thus failing to detect them.

Eye tracking techniques, coupled with real-time video analysis, can aid in bringing object awareness to the viewer. In other words, the objects that are in the field of view but missed by the observer can still be detected and recognized.

This paper introduces a novel methodology utilizing the modern mobile eye tracking technology, which not only can provide aid in gaze based object detecting and recognizing but also identify areas in the field of view that are not being assessed by the viewer.

III. Proposed System architecture

The complete system architecture for aerial border surveillance for search and rescue missions using eye tracking techniques is shown in Fig. 2 below. The proposed system architecture consists of four parts: a head-mounted eye tracker; think aloud protocol, gaze to object classification; and eye movements monitoring. In this section, details about principle components of the proposed system are given.

A. *Head-mounted eye tracker component*

As noted before, head-mounted eye tracking technology is lightweight and therefore convenient to wear in daily activities [12]. Here, professional teams can be trained to wear this new type of cutting-edge equipment while performing searching and rescuing tasks.

Head-mounted eye trackers typically consist of a front scene camera capturing the participant's field of view, and infrared cameras directed at participants' eyes to monitor the pupil position by tracking the distance between the center of the pupil and the position of artificially illuminated corneal reflections. Current available head-mounted eye trackers can calculate and provide the information of where and what people look at without the necessity to restrict their movement [13], which allows researchers to expand their investigation to more ecologically valid experiments. In addition, a concurrent voice recorder can record human's voice while they are on a search and rescue task.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: TUFTS UNIV. Downloaded on October 06,2020 at 15:39:35 UTC from IEEE Xplore. Restrictions apply.

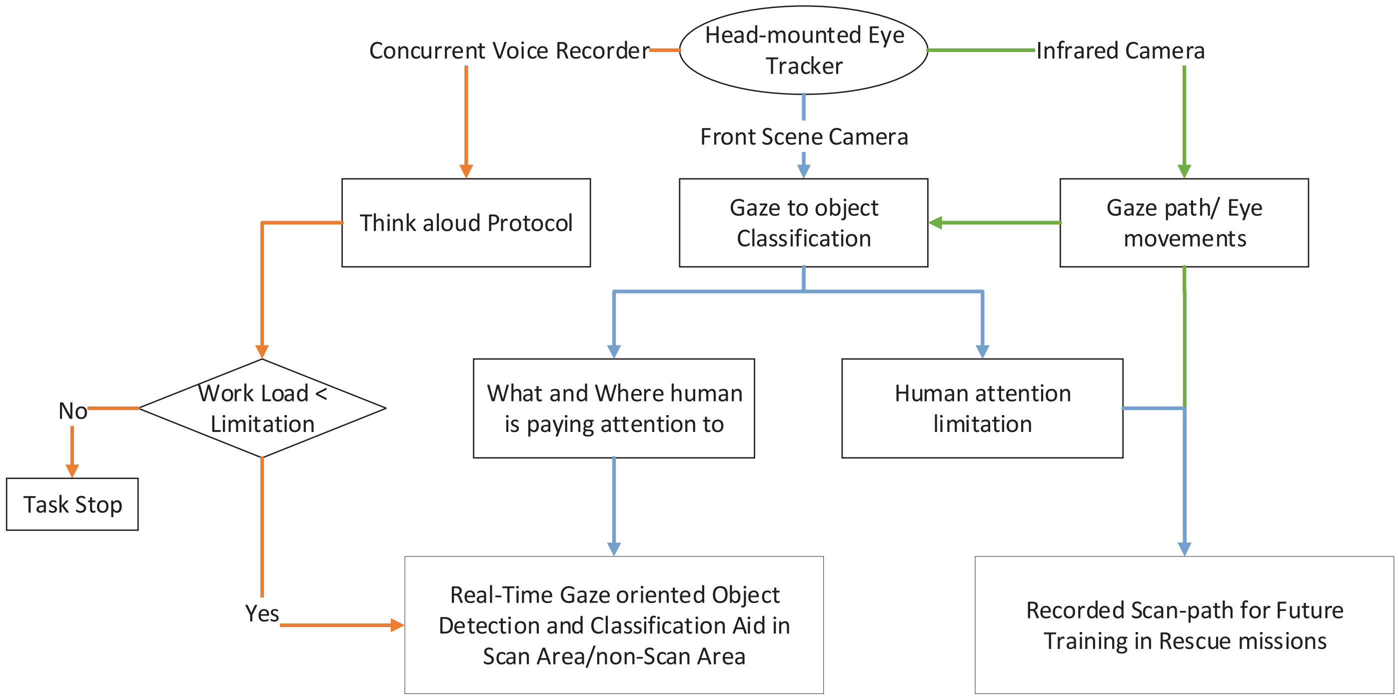

Fig.2 System Overview: the system includes an eye tracking component (professional people wearing eye tracker to act rescuing and searching mission), a real-time gaze-oriented object detection and classification aiding tool in scan and non-scanned region to help finding searching and rescuing targets, scan path and eye movements recording module is developed for training purpose, and think aloud speech process is proposed for workload estimation.

B. Gaze to object classification and recognition framework

1) The gaze oriented object classification and recognition is based on a pre-trained classifier using standard machine learning tools. Note that, the real-time searching and rescuing process is done by a trained person searching for targets within a specific area. Calibration and gaze estimation features are provided in the commercial eye tracing packages. Hence, real-time deployment of visual attention toward certain objects is highlighted. Real-time video analysis of the detected gaze pointed area can increase the search accuracy. The general steps are shown in the Fig 3 below:

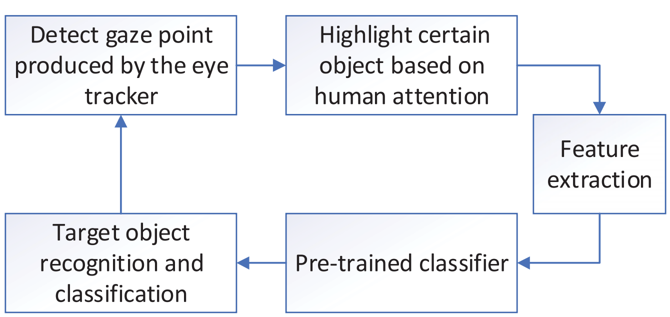

Fig. 3 General steps for attention oriented object classification and recognition framework.

In this step, we analyze image patches around the user's gaze position to determine the class for the object of interest. Meanwhile, the sequences of attention events that provide border security targeting information is built up. In other words, a border security interest gallery is collected, maintained and updated continuously, which is useful to train the classifier for real-time researching and rescuing.

Moreover, real-time features of using head-mounted eye tracker component can be extended and developed in the short future. Potentially, other state-of-art image processing and computer vision tools can be coupled with wearable eye tracker. For instance, image enhancement and image de-noising can be used to give a better vison view for search and rescue team members.

2) Video real-time object detection and identification algorithms in non-scanned region

The most outstanding contribution of using wearable eye tracker is that it enables us to identify which regions within the visual scene were not attended to by the viewer. Survivors might be missed by the rescue team simply because nobody noticed them.

By potentially real-time processing the front-facing camera captured video, we can perform object detection and classification on the non-scanned areas using object recognition tools, such as deep convolutional neural networks. The built up border security interest targets gallery is used for feature extraction and classifier training. This step, which we believe, can increase the possibility of finding missing targets.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: TUFTS UNIV. Downloaded on October 06,2020 at 15:39:35 UTC from IEEE Xplore. Restrictions apply.

C. Scan path and eye movements recording

The proposed system also enables human performance analytics procedure by recording the scan path and eye movements, which will be instrumental in developing training methodologies to enhance effectiveness in researching and rescuing missions. This can accelerate knowledge and skill acquisition into practice.

D. Discussion on think aloud protocol

Another important feature of the proposed system provides is the mechanism of an illustrative think aloud protocol: the system has capabilities to distinguish between noise and silence, as well as return sound measures such as voice pitch and prosody. It can be used to automatically analyze and recognize cognitive states including: emotion detection, workload stress estimation, inside feeling indicators, what the person is thinking while completing a task. In a specific search and rescue task, if a searching team member's workload is over the limitation, he or she should terminate the task immediately.

IV. Example Experimental Results for Gaze to Object Classification and Recognition Framework

Since we could not find an existing aerial border surveillance head-mounted eye tracking dataset, we choose to evaluate the proposed system architecture using an large-scale eye-tracking dataset that was developed and collected by a group of psychologists. The dataset was a case study explored how engineers discovered possible solution paths in aiding completion of a mundane and repetitive task, detailed information can be found in [14].

For each object category, 70% images were selected as training images and 30% images were used as testing images. Recognition rates are generated using different algorithms: Bag of visual words for gazed object classification, Histograms of Oriented Gradients (HOG) with a flat machine learning classifier (SVM), and transfer learning.

The steps for bag-of-visual-words (BoVW) classification [15] are: 1) Extract SURF [16] (fast speed scale-invariant feature transform) features from all training images; 2) Construct the visual words vocabulary by K-means clustering [17] (K=500); 3) Generate a histogram to present each image; 4) Train the classifier and predict the object's category.

Hand-crafted feature extraction using Histograms of Oriented Gradients (HOG) is tested for the proposed system against the eye tracking dataset. Dalal and Triggs [18] [19] introduced HOG features, which serves to encode the edge orientation features in an image.

Recent astonishing work in the field of deep convolutional neural network (DCNN) has demonstrated its ability to extract unique features from input images and its superior performance over a wide variety of tasks [20]. However, they have few limitations such as: large-scale amount of training data is needed; expensive computational power is required. To solve the challenges listed above, transfer learning is proposed to optimize the system to solve a task with limited dataset and reduce computational power by using pre-trained models from a similar task. In this work, the pre-tained model is based on ImageNet, which is mainly developed for object recognition purposes. And we removed the last fully connected layer and and fine-tuned the DCNN to classify different objects by adding customized layers.

The table below shows the recognition and classification rate using different algorithms. We can see that, transfer learning achieves higher recognition and classification rates.

TABLE 1 RECOGNITION RATE USING DIFFERENT ALGORITHMS

	HOG+SVM	**BoVW**	**Transfer Learning**
Recognition rate	70.3%	83.4%	91.7%

V. Conclusion

This paper proposed a novel "hands-free" methodology for aerial border search and rescue mission utilizing the modern mobile eye tracking technology, which not only can provide aid in real-time gaze based object detecting and recognizing but also identify areas in the field of view that are not being assessed by the viewer.

The contributions of the proposed new visual analytics approach are: i) a gaze based object classification and recognition framework for aerial boarder surveillance; ii) non-scanned regions real-time object detection and identification; iii) a discussion of how the scan-paths (fixation and non-scanned) provided by mobile eye tracker can help improve training professional search and rescue organizations or even artificial intelligence robots for searching and rescuing missions.

The architecture has been tested using large-scale real-life head-mounted eye tracking data, which achieved remarkable accuracy and demonstrate its promising value for aerial border search and rescue application.

In the future, the system's accuracy should be examined and improved by using an aerial border surveillance dataset collected by head-mounted eye trackers.

Acknowledgment

This work is supported by the Tufts University School of Engineering, Center for Applied Brain & Cognitive Sciences (CABCS) under award numbers A452001 ARM981. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the CABCS.

References

[1] C. Bolkcom, "Homeland security: Unmanned aerial vehicles and border surveillance," 2004.

[2] D. Erdos, A. Erdos, and S. E. Watkins, "An experimental UAV system for search and rescue challenge," IEEE Aerospace and Electronic Systems Magazine, vol. 28, pp. 32-37, 2013.

[3] A. Duque and C. Vázquez, "Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: TUFTS UNIV. Downloaded on October 06,2020 at 15:39:35 UTC from IEEE Xplore. Restrictions apply.

tracking study," Journal of Behavior Therapy and Experimental Psychiatry, vol. 46, pp. 107-114, 2015/03/01/ 2015.

[4] J. B. Pelz, T. B. Kinsman, and K. M. Evans, "Analyzing complex gaze behavior in the natural world," in Human Vision and Electronic Imaging XVI, 2011, p. 78650Z.

[5] T. P. Keane, N. D. Cahill, J. A. Tarduno, R. A. Jacobs, and J. B. Pelz, "Computer vision enhances mobile eye-tracking to expose expert cognition in natural-scene visual-search tasks," in Human Vision and Electronic Imaging XIX, 2014, p. 90140F.

[6] K. M. Evans, R. A. Jacobs, J. A. Tarduno, and J. B. Pelz, "Collecting and analyzing eye tracking data in outdoor environments," Journal of Eye Movement Research, vol. 5, p. 6, 2012.

[7] M. Sodhi, B. Reimer, J. Cohen, E. Vastenburg, R. Kaars, and S. Kirschenbaum, "On-road driver eye movement tracking using head-mounted devices," in Proceedings of the 2002 symposium on Eye tracking research & applications, 2002, pp. 61-68.

[8] M. Wedel and R. Pieters, "A review of eye-tracking research in marketing," in Review of marketing research, ed: Emerald Group Publishing Limited, 2008, pp. 123-147.

[9] A. Poole and L. J. Ball, "Eye tracking in HCI and usability research," in Encyclopedia of human computer interaction, ed: IGI Global, 2006, pp. 211-219.

[10] A. Al-Rahayfeh and M. Faezipour, "Eye tracking and head movement detection: A state-of-art survey," IEEE journal of translational engineering in health and medicine, vol. 1, 2013.

[11] S. Sridharan, Gaze Guidance, Task-Based Eye Movement Prediction, and Real-World Task Inference using Eye Tracking: Rochester Institute of Technology, 2016.

[12] M. Barz and D. Sonntag, "Gaze-guided object classification using deep neural networks for attention-based computing," presented at the Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, Heidelberg, Germany, 2016.

[13] G. Andrienko, N. Andrienko, M. Burch, and D. Weiskopf, "Visual analytics methodology for eye movement studies," IEEE Transactions on Visualization and Computer Graphics, vol. 18, pp. 2889-2898, 2012.

[14] Q. W. Karen Panetta, Aleksandra Kaszowska, Holly A. Taylor and Sos Agaian, "(Submitted, under review)A Software Architecture for Automating Cognitive Science Eye Tracking Data Analysis and Object Annotation."

[15] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in Workshop on statistical learning in computer vision, ECCV, 2004, pp. 1-2.

[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, pp. 91-110, 2004.

[17] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967, pp. 281-297.

[18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, pp. 886-893.

[19] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," Pattern Recognition Letters, vol. 32, pp. 1598-1603, 2011.

[20] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2014, pp. 806-813.

978-1-5386-3443-1/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: TUFTS UNIV. Downloaded on October 06,2020 at 15:39:35 UTC from IEEE Xplore. Restrictions apply.