2012-06

Modeling Peripheral Vision for Moving Target Search and Detection

Yang, Ji Hyun; Huston, Jesse; Day, Michael; Balogh, Imre

Aviation, Space, and Environmental Medicine, Vol. 83, No. 6, June 2012
https://hdl.handle.net/10945/41686

Modeling Peripheral Vision for Moving Target Search and Detection

Ji Hyun Yang, Jesse Huston, Michael Day, and Imre Balogh

Yang JH, Huston J, Day M, Balogh I. *Modeling peripheral vision for moving target search and detection.* **Aviat Space Environ Med 2012; 83:585–93.**

Introduction: Most target search and detection models focus on foveal vision. In reality, peripheral vision plays a significant role, especially in detecting moving objects. ***Methods:*** There were 23 subjects who participated in experiments simulating target detection tasks in urban and rural environments while their gaze parameters were tracked. Button responses associated with foveal object and peripheral object (PO) detection and recognition were recorded. In an urban scenario, pedestrians appearing in the periphery holding guns were threats and pedestrians with empty hands were non-threats. In a rural scenario, non-U.S. unmanned aerial vehicles (UAVs) were considered threats and U.S. UAVs non-threats. ***Results:*** On average, subjects missed detecting 2.48 POs among 50 POs in the urban scenario and 5.39 POs in the rural scenario. Both saccade reaction time and button reaction time can be predicted by peripheral angle and entrance speed of POs. Fast moving objects were detected faster than slower objects and POs appearing at wider angles took longer to detect than those closer to the gaze center. A second-order mixed-effect model was applied to provide each subject's prediction model for peripheral target detection performance as a function of eccentricity angle and speed. About half the subjects used active search patterns while the other half used passive search patterns. ***Discussion:*** An interactive 3-D visualization tool was developed to provide a representation of macro-scale head and gaze movement in the search and target detection task. An experimentally validated stochastic model of peripheral vision in realistic target detection scenarios was developed.

Keywords: peripheral vision, target detection, recognition, search and target acquisition.

MODELING OF SEARCH and target acquisition (STA) has been a major concern for military simulations. For example, simulation models considering individual soldiers such as COMBAT XXI, OneSAF, and JSAF use the ACQUIRE algorithm for calculating visual detection probabilities. However, it has been shown that the ACQUIRE algorithm does not sufficiently reflect the performance of human observers (3,8,9), i.e., false positive detection and correct detection should be taken into account and modeled. Although frequencies of false positive detection can be modeled as a certain probabilistic property, the location of false positives still remains to be explored. Furthermore, since the ACQUIRE model was originally developed to represent imaging sensors, only a limited field of view is considered (16). As this model was extended to represent unaided human vision, this limitation was not addressed. Therefore, all combat simulations that use ACQUIRE-based models ignore what happens in the peripheral field of view of human observers. The motivation for this work is to address this deficiency in our current simulation by developing a model of detection outside of the foveal field of view that can be used in conjunction with the current methodologies to provide a better representation of unaided human vision.

Current target detection mechanisms in urban environments use the so-called 'windshield wiper' approach (4), where the visual field is split into several adjacent and non-overlapping fields of view and the target detection mechanism is applied to each field of view independently, generally in a sweep from left to right and back. The way of determining the locations to which the target detection mechanism is applied is far from actual human behavior. Jungkunz (8,9) investigated more likely fixation locations with respect to eccentricity, saliency, and distracters in the scene. He found that the maximum distracting capability is not tied to maximum saliency, but the distractor attracts the gaze less if its eccentricity from the initial fixation location gets longer.

Target detection algorithms, including those used in the above studies, have been mainly concerned with human eye movement, specifically where the foveal gaze moves and how long the gaze stays during the search task. For example, probably the best-known model of visual attention (6,7,11) uses saliency to determine the focus of attention. Improvements on this model employ machine-learning methods to train a detector on objects of interest (10) or gaze patterns of human subjects performing a similar task (13). While these models do perform better target detection than saliency-only in the indicated studies and are based to mimic human unaided search algorithms, they do not provide insight as to how peripheral vision influenced those movements.

The fovea provides high visual acuity within 2° of visual angle and this acuity decreases with higher eccentricity from the center of the visual field (5,14,15). On the other hand, peripheral vision is good at fast detection of

From the MOVES Institute, Naval Postgraduate School, Monterey, CA.

This manuscript was received for review in October 2011. It was accepted for publication in January 2012.

Address correspondence and reprint requests to: Ji Hyun Yang, Ph.D., MOVES Institute, 700 Dyer Rd., Bldg. 245, Rm. 379, Naval Postgraduate School, Monterey, CA 93943; jyan1@nps.edu.

DOI: 10.3357/ASEM.3230.2012

movement and seeing in dim lighting conditions (17). Considering many targets are non-static in realistic scenarios, peripheral vision is essential for detecting moving objects outside the fovea. However, little work has been focused on the role of peripheral vision in STA in realistic military scenarios. Junkunz (8,9) observed that subjects hardly ever waste even a single fixation if no distracting items are present, even if the target has very low contrast with the background. His study focused on foveal vision (in the vicinity of the gaze fixation point) and did not investigate effects of peripheral vision. Since peripheral vision is useful for detecting objects, it could affect how humans follow gaze fixations points.

Thus we designed a search and target detection task in a 180° horizontal, 60° vertical field-of-view virtual environment integrated with eye-tracking systems and performed a human-in-the-loop experiment. Our goal was to provide an improved model of unaided human vision, i.e., a stochastic model of target detection performance as a function of peripheral angle (eccentricity) and object entrance speed. As part of this project we developed an interactive visualization tool designed to provide a representation of spatial and temporal correspondence among features scanned in the virtual environment in relation to dynamic changes in the scenario, including peripheral objects (POs) and moving vehicles.

METHODS

Subjects

There were 23 subjects (21 men, 2 women), ages 24 to 44, who participated in the study. Subjects self-reported visual acuity. There were 17 who reported 20/20 vision or better, 4 who reported 20/30 vision, 1 who reported 20/40 vision, and 1 who reported 20/50 vision. As long as subjects could identify targets in the practice session, where they were allowed to ask questions and experimenters were allowed to provide answers and directions, they could continue with the main urban and rural scenarios. Subjects who were military personnel had served between 6 and 19 yr of active duty. This study was approved by the Naval Postgraduate School Institutional Review Board. Subjects were recruited through school-wide e-mails and fliers. All subjects provided Institutional Review Board approved informed consent to participate in this study and were made aware of their right to withdraw at any time without consequence.

Equipment

The basic elements of the apparatus included three sets of two stereo cameras and associated faceLAB 5.0 software (Seeing Machines Inc., Tucson, AZ) collecting eye and head movement data, three 94″ x 63″ screens on which a continuous simulated environment was projected, and one Bamboo touch pad (Wacom Co., Ltd., product #: CTT460) for recording subjects' responses. Each subject was positioned 7 ft away from the center screen to ensure having a 180° horizontal field-of-view in the simulated environment. One set of stereo cameras with 16-mm lenses was positioned in front of each projector display and adjusted for each subject individually, based on height. The faceLAB Link 2.0 software connected the three sets of stereo cameras to link individual systems to work as one and track 180° of head rotation from left to right. An Image Generator (IG) driven by the Delta3D game engine provided a view of urban and rural virtual environments on the display. Responses were recorded using four buttons on a Bamboo touch pad. Data from the touch pad and faceLAB Link 2.0 were sent to the IG and integrated into a data file synchronized with the virtual environment.

Procedure

The context for the experiment given to the participants was that they were manning a vehicle checkpoint in a location where there may be hostile activity. The main task was to monitor vehicles approaching the checkpoint for potential threats, but also to be vigilant for other possible threats. The focus of the experiments was to collect data on peripheral vision, but the subjects were not told this so as not to bias their behavior. They were told that the main vehicle monitoring task was the focus of the study. Two settings were used for the study: one was a complex urban setting with many buildings and other terrain features. The second one was a rural setting with only a small number of buildings and few other terrain features.

Subjects were instructed to detect objects in their periphery, i.e., pedestrians or unmanned aerial vehicles (UAVs). POs appeared at various angles and speeds. In the urban scenario, peripheral angles (eccentricity) were –80°, –70°, – 55°, +45°, +65°, and +80°, where 0° was on a line beginning at the subject's head and ending in the middle of the center display, and left and right rotation is shown in – and +, respectively. Eccentricity angles were non-symmetric in the urban scenario because the POs, pedestrians, needed to hide behind urban features such as a trash bins, trees, or pillars. POs appeared in slow or fast motion, which was $0.05° \cdot s^{-1}$ or $0.2° \cdot s^{-1}$, respectively. Inter-arrival times between POs were predefined via a Poisson process having $\lambda = 8$ s. POs in the urban environment, pedestrians, would slide out from behind an object located on the predefined eccentricity angle slowly or quickly and remain on screen for 2 or 10 s or until they were identified as friendly or hostile. Threats (hostile pedestrians) held guns, while non-threats (friendly) ones did not. In the rural scenario, POs were UAVs that appeared in the distance and moved directly toward the subject before reaching a maximum size and turning toward the closest edge of the display. UAVs would then fly at a constant speed for 4 s or until they were identified as friendly or hostile before disappearing. Peripheral angles were ±50°, ±65°, ±75°, ±80°, and ±85°. Threats (hostile UAVs, HESA Ababil, Iran) had a red canard, a wing at the back and an upward vertical stabilizer, while non-threats (friendly UAVs, RQ 1A Predator, United States) had a wing in the middle of the fuselage and downward-pointing tails. Subjects were told to press a "peripheral threat button" when recognizing threats and a "peripheral non-threat button" when recognizing non-threats.

Vehicle monitoring task: the subjects' main task was to monitor vehicles to keep their gaze predominantly in the vicinity of 0° of eccentricity, i.e., the center of the screen. The IG only placed POs when subjects' gaze was in the vicinity of the center of the field of view—the Poisson process mentioned before was paused if a subject's gaze was lingering outside this vicinity. Threats (hostile automobiles) were defined as dump trucks and utility trucks with a female driver and male passenger as well as sedans with a male driver and no passenger. The rest of the vehicles were considered non-threats or friendly. The vehicles only appeared on the central projector screen. They would appear in the distance and follow a road to the checkpoint, eventually coming to a full stop. At this point, the vehicle would be close enough for the subject to identify the genders of the people in the vehicles. Once the vehicle stopped, subjects pressed either the "threat vehicle button" or the "non-threat vehicle button" to identify whether the vehicle was hostile or friendly. Then, the vehicle would disappear and the next one in the queue would pull up to take its place. Vehicles would remain on the screen until identified as friendly or hostile. Concurrently with the vehicle monitoring task, subjects were performing the peripheral object detection task described above.

After a brief introduction to the study, subjects were first asked to read and sign an informed consent form before the experiment started. Then they filled out anonymous demographic data, which included level of visual acuity and number of years of military service. The three faceLAB™ eye tracking systems were then calibrated for the subject. Following the calibration the subjects were briefed on how to identify hostile vehicles and were instructed that their primary task was to assess whether vehicles would be allowed to pass through a security checkpoint. Their secondary task would be to watch for and identify pedestrians in the urban scenario or UAVs in the rural scenario. Subjects were told to emphasize accuracy over speed. All subjects used binocular vision at all times. After the task was explained, a practice scenario was administered using the urban environment. This scenario would go on for 10 min or until the subject said that they felt confident in their ability to correctly identify targets. During the practice task, if the subject asked a question, it would be answered. When the practice had finished, the experimental urban scenario was given, which lasted for 10 min. During the experimental run, experimenters remained silent though the subjects were allowed to speak. At the completion of the urban scenario, the faceLAB™ calibration was checked to ensure that the eye trackers were still operating correctly. Once this was confirmed, the rural practice scenario was given, followed by the rural main scenario. After this was finished, screen calibration data were collected again to compare eye tracking data quality with that of the beginning of the experiment.

Statistical Analyses

Signal detection theory was used to indicate overall performance on the target detection, recognition, and vehicle monitoring task. Missed Detection and False Alarms represent negative performance metrics, whereas Correct Rejection and Correct Detection are positive metrics. A repeated measure ANOVA was performed as a preliminary analysis and a mixed-effect model (10) was used to develop mathematical models for the prediction of target detection performance. Mixed-effects modeling can determine both fixed effects and random effects, which are intertwined in experimental data and can provide a model in a proper and parsimonious way (12). The general form of a mixed-effect model can be described as below (10):

$$y_{ij} = f(\boldsymbol{\varphi}_i, \mathbf{x}_{ij}) + e_{ij}$$

where y_{ij} is the j_{th} response of the i_{th} individual, $\boldsymbol{x}_{ij}$ is the predictor vector for the j_{th} response of the i_{th} individual, f is a nonlinear function of the predictor vector and a parameter vector φ_i of length r, and e_{ij} is a normally distributed noise term. There are no restrictions on the predictor vectors $\boldsymbol{x}_{ij}$. The parameter vector can vary from individual to individual. This is incorporated into the model by writing φ_i as

$$\varphi_i = \mathbf{A}_i\beta + \mathbf{B}_i\mathbf{b}_i, \ \mathbf{b}_i \approx N(\mathbf{0}, \sigma^2\,\mathbf{D})$$

where β is a p-vector of fixed population parameters, $\mathbf{b}_i$ is a q-vector of random effects associated with individual i, the matrices $\mathbf{A}_\mathrm{i}$ and $\mathbf{B}_\mathrm{i}$ are design matrices of size $r \times p$ and $r \times q$ for the fixed and random effects, respectively, and $\sigma^2\mathbf{D}$ is a covariance matrix. A second order model for the prediction of target detection performance was used in this study.

RESULTS

The significance level α for testing hypotheses was set to 0.05. Spearman's rank correlation is shown in ρ and the corresponding P-value is shown in P. A total of 50 POs appeared in the urban and rural scenarios. Among the 50 PO appearances, subjects did not respond 2.48 times on average (median = 2, SD = 1.88) in the urban scenario and 5.32 times on average (median = 5, SD = 3.28) in the rural scenario. Of the 50 Pos, 4 were threats, i.e., pedestrians lifting guns or hostile UAVs with red canards. In the urban scenario, on average, subjects missed 0.35 threats (median = 0, SD = 0.57), whereas they falsely identified non-threats as threats 0.26 times (median = 0, SD = 0.54). In the rural scenario, on average, subjects missed 0.18 threats (median = 0, SD = 0.39), whereas they falsely identified non-threats as threats 0.55 times (median = 0, SD = 0.80). **Table I** summarizes overall PO detection performance for all subjects.

TABLE I. DESCRIPTIVE STATISTICS OF PERIPHERAL OBJECT DETECTION AND VEHICLE MONITORING PERFORMANCE.

	Urban Scenario			Rural Scenario		
	Mean	Median	SD	Mean	Median	SD
Peripheral object detection performance						
Peripheral Objects Missed (Missed Detection)	2.48	2	1.88	5.32	5	3.28
Targets Missed (Missed Recognition)	0.35	0	0.57	0.18	0	0.39
Non-Targets Recognized (False Alarm)	0.26	0	0.54	0.55	0	0.80
Targets Detected (Correct Recognition)	3.30	3	0.71	3.05	3	0.79
Non-Targets Rejected (Correct Rejection)	44.60	45	1.90	40.91	42	3.37
Vehicle Monitoring Task Performance						
Vehicles Checked	128.0	126.5	5.28	147.4	144	10.18
Vehicles in Queue	0.072	0.039	0.086	0.068	0.012	0.16
Vehicles in View	1.73	1.70	0.14	3.27	3.23	0.15
Threats Missed (Missed Detection)	0.14	0	0.35	0.41	0	0.67
Non-Threats Recognized (False Alarm)	3.36	1.5	4.96	0.82	0	1.74

POs appeared at various angles and speeds. **Table II** summarizes the mean, median, and SD of button reaction time for each peripheral angle and speed. Similarly, saccade reaction time (first fast eye movement followed by each PO's appearance) was analyzed in the same way as shown in Table II. In the urban scenario, a two-factor, within-subject repeated measures ANOVA showed that subjects had faster saccade reaction time (SRT) and button reaction time (BRT) for fast-moving POs than for slow-moving POs [F(1,22) = 43.52, $P < 0.001$ and F(1,22), = 236.66, $P < 0.001$]; bigger peripheral angles had slower SRT and BRT in both the left and right directions [F(5110) = 13.06, $P < 0.001$ and F(5110), = 11.29, $P < 0.001$]. **Fig. 1** shows the mean SRT and BRT and visually confirms the statistical findings. Effects of PO eccentricity and speed were not as apparent in the rural scenario as in the urban scenario.

A mixed-effect model was used to develop mathematical models for the prediction of target detection performance. Polynomial bases were selected since continuous functions can be estimated by polynomials (e.g., Taylor series). As we increase the order of the polynomial, we can account for more nonlinearity in the system. Therefore, we decided on a second order polynomial as an initial default model, which should not be oversimplifying as a linear model would be, nor overfit the data as

TABLE II. MEAN, MEDIAN, AND SD OF BUTTON/SACCADE REACTION TIME FOR EACH ANGLE AND SPEED COMBINATION *P*-VALUES COMPARING GROUP MEANS BETWEEN PASSIVE AND ACTIVE GROUPS.

	Urban Scenario (Pedestrian detection: mean/median/SD)				Rural Scenario (UAV detection: mean/median/SD)			
	Button Reaction Time		Saccade Reaction Time		Button Reaction Time		Saccade Reaction Time	
	Slow	Fast	Slow	Fast	Slow	Fast	Slow	Fast
–85°	N/A	N/A	N/A	N/A	2.56/2.52/0.66 (**0.0017**)	2.47/2.41/0.86 (**0.029**)	1.17/0.82/0.88 (**0.0024**)	0.94/0.72/0.80 (**0.017**)
–80°	2.28/2.13/0.46 (0.57)	1.42/1.28/0.42 (0.12)	1.16/1.11/0.38 (<**0.001**)	0.88/0.89/0.28 (**0.046**)	2.17/2.16/0.72 (0.18)	2.30/2.18/0.55 (<**0.001**)	0.96/0.76/0.69 (0.33)	0.94/0.70/0.58 (**0.0018**)
–75°	N/A	N/A	N/A	N/A	N/A	2.08/1.83/0.69 (0.065)	N/A	0.84/0.58/0.46 (<**0.001**)
−70°	2.17/2/08/0.43 (0.22)	1.36/1.26/0.24 (**0.04**)	1.07/1.08/0.47 (**0.0091**)	0.78/0.80/0.23 (<**0.001**)	N/A	N/A	N/A	N/A
–65°	N/A	N/A	N/A	N/A	2.08/1.98/0.70 (0.14)	2.24/2.06/0.59 (0.25)	0.92/0.79/0.52 (**0.0012**)	0.88/0.57/0.41 (**0.0039**)
–55°	1.67/1.51/0.24 (0.43)	1.17/1.10/0.21 (0.76)	0.88/0.83/0.31 (<**0.001**)	0.66/0.72/0.22 (0.052)	N/A	N/A	N/A	N/A
–50°	N/A	N/A	N/A	N/A	1.89/1.74/0.43 (0.61)	2.36/2.25/0.85 (0.60)	0.79/0.75/0.33 (<**0.001**)	0.91/0.68/0.69 (0.21)
45°	2.05/1.84/0.50 (0.56)	1.33/1.22/0.23 (0.10)	0.99/0.98/0.42 (**0.01**)	0.71/0.72/0.21 (**0.001**)	N/A	N/A	N/A	N/A
50°	N/A	N/A	N/A	N/A	1.90/1.75/0.54 (0.30)	2.26/2.11/0.54 (0.39)	1.06/1.10/0.70 (0.12)	0.91/0.58/0.46 (**0.016**)
65°	1.79/1.60/0.34 (0.06)	1.42/1.32/0.32 (0.41)	0.92/0.98/0.35 (<**0.001**)	0.77/0.72/0.24 (**0.025**)	2.09/2.09/0.44 (**0.012**)	2.24/2.27/0.57 (**0.060**)	1.01/0.67/0.61 (**0.0024**)	1.18/1.25/0.52 (**0.0035**)
75°	N/A	N/A	N/A	N/A	2.71/2.53/0.79 (**0.044**)	1.85/1.92/0.74 (0.26)	1.15/0.66/0.68 (<**0.001**)	0.96/0.57/0.88 (**0.0028**)
80°	2.24/1.98/0.65 (0.07)	1.74/1.51/0.77 (0.20)	1.37/1.31/0.60 (0.005)	0.90/0.96/0.28 (0.016)	3.36/3.27/1.36 (**0.048**)	2.66/2.33/0.89 (**0.02**)	1.53/1.19/1.35 (0.12)	1.07/0.93/0.86 (**0.03**)
85°	N/A	N/A	N/A	N/A	2.27/2.05/0.54 (0.065)	N/A	0.84/0.53/0.58 (0.12)	N/A

Bold font indicates a significant difference.

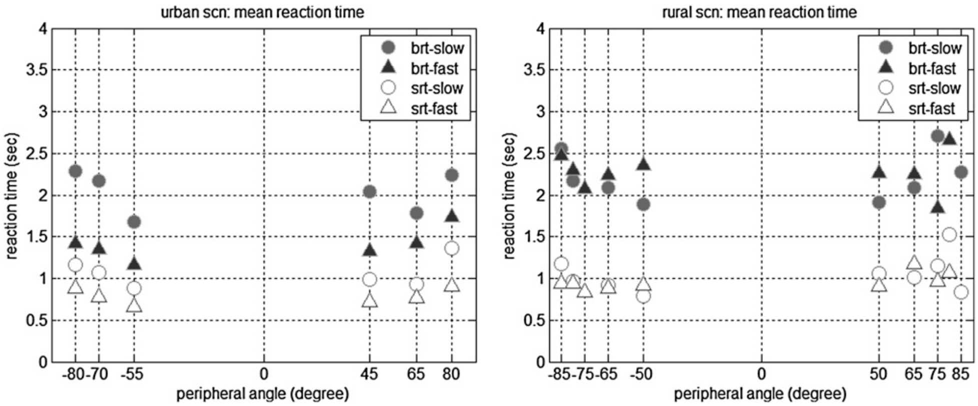

Fig. 1. Mean button reaction time (BRT) and mean saccade reaction time (SRT).

the higher order models do. A second order model for the prediction of target detection performance was set as shown below:

$$\begin{aligned} rt &= \phi_1 \cdot x_1^2 + \phi_2 \cdot x_1 + \phi_3 \cdot x_2 + \phi_4 \\ &= (\beta_1 + b_1) \cdot x_1^2 + (\beta_2 + b_2) \cdot x_1 + (\beta_3 + b_3) \cdot x_2 + (\beta_4 + b_4) \end{aligned}$$

where rt = reaction time, x_1 = angle (in degrees), x_2 = speed (1 = slow, 2 = fast), β_i = fixed effects for x_i, b_i = random effects for x_i. Estimated variables using MATLAB for both urban/rural scenarios and button/saccade reaction time are shown in **Table III**. Fixed parameters (β_1, β_2, β_3, and β_4) stayed in the model if they were significant ($\alpha = 0.05$). Similarly, random parameters (b_1, b_2, b_3, and b_4) were removed from the model if the corresponding covariance matrix element was negligible, i.e., < 0001. For all cases, random effects b_1, b_2, and b_3 were negligible, so using only fixed effects (i.e., β_1, β_2, and β_3) provided a better model fit. On the other hand, the random effect in the constant term (b_4) remained in the model. The corresponding random effect b_4 for each subject varied from −0.7525 to 1.0524 (The corresponding random effect b_4 for each subject is available upon request). For the rural scenarios, the speed of a peripheral object, x_2, was not a significant predictor of the model in the rural scenario, thus it was removed from the final model. All fixed effects were significant ($P < 0.05$) except a fixed effect of x_1^2, i.e., β_1 of rural SRT ($P = 0.062$).

Both BRT and SRT were positively associated with the second order terms of the peripheral angle, such that the reaction time graph describes a parabola facing upward as shown in **Fig. 2**. This upward parabola shows that reaction time tends to slow as peripheral angle increases. The first order term of the angle is negatively associated with both BRT and SRT, showing that the symmetry point of the parabola is not exactly on the center, but rather skewed toward the right about 3.5°−5.5°. This consistent minor skewness to the right side of the screen could just be a modeling error or subjects' general preference to the right side of the screen (or the seat might have been slightly off center or slightly rotated to the right). The symmetry line was on 17.7° for rural SRT; however, the fixed effect was not statistically significant and this number is not as reliable as previous ones. On the other hand, speed was negatively associated with both BRT and SRT, representing faster POs tending to lead to shorter reaction times in general. Fig. 2 represents data and the mixed-effect model of Subject 12 as an example.

The above mixed-effect model predicts reaction time only when subjects detected POs. Whether subjects detected POs on a given peripheral angle and speed can be obtained from the data. Table I showed overall PO detection performance whereas **Fig. 3** shows PO detection probability for each peripheral angle and speed given for the urban and the rural scenario, respectively. For instance, on average, subjects detected only 62.6% of peripheral targets that appeared at −85°, but 100% at ±50° in the rural scenario. The PO detection probability shows an inverse U-shape describing how the detection probability decreases as peripheral angle increases. Subjects failed to notice POs in the rural scenarios more than

TABLE III. MIXED-EFFECT MODEL PARAMETER ESTIMATION.

	Mixed-Effect Model	Remarks
Urban BRT	$brt_{urban} = .000088^{**} \cdot x_1^2 - .00081^{*} \cdot x_1 - .63^{**} \cdot x_2 + (2.3^{**} + b_4)$	Random effects b_1, b_2, and b_3 are negligible and thus are not included in the model.
Urban SRT	$srt_{urban} = .000067^{**} \cdot x_1^2 - .00053^{*} \cdot x_1 - .28^{**} \cdot x_2 + (1.05^{**} + b_4)$	
Rural BRT	$brt_{rural} = .00010^{**} \cdot x_1^2 - .0011^{*} \cdot x_1 + (1.80^{**} + b_4)$	x_2 was not a significant predictor ($P > 0.01$) and thus was not included in the model.
Rural SRT	$srt_{rural} = .000031\dagger \cdot x_1^2 - .0011^{*} \cdot x_1 + (.85^{*} + b_4)$	Random effects b_1, b_2, and b_3 are negligible and thus are not included in the model.*

BRT = button reaction time; SRT = saccade reaction time.
† $P < 0.01$; * $P < 0.05$; ** $P < 0.01$.

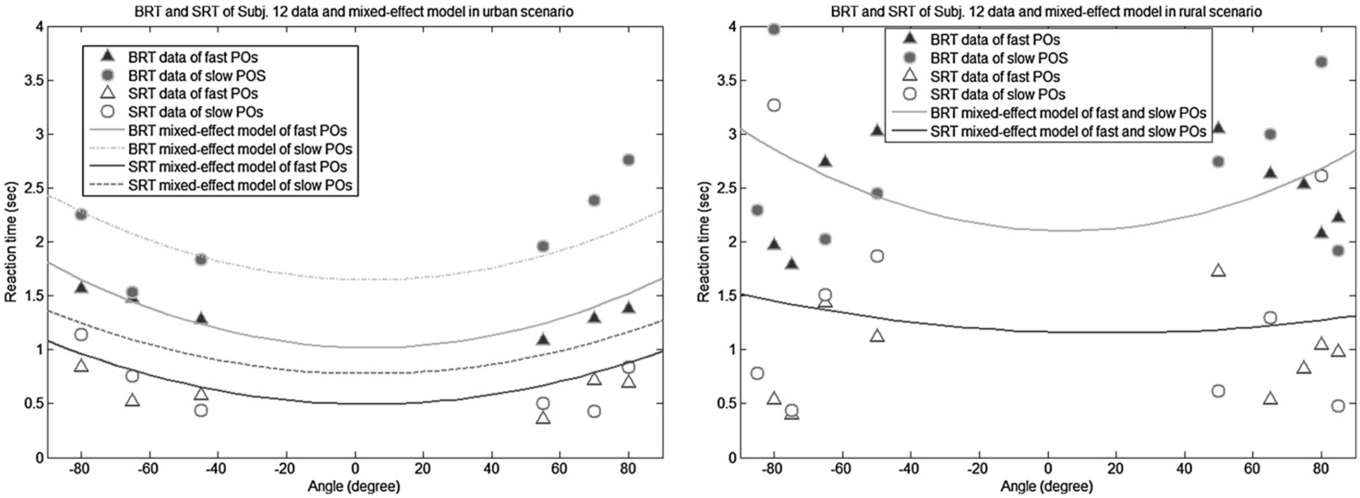

Fig. 2. Mixed-effect model shown with BRT and SRT data of subject No.12.

twice as much as in the urban scenarios, i.e., "Num. of peripheral objects missed (Missed Detection)" was 2.48 times and 5.32 times in the urban and rural scenarios, respectively. This can be explained by the fact that rural scenarios had higher eccentricity angles, e.g., 85°, and subjects indeed showed lower detection probability at those angles.

The vehicle monitoring task was the main task given to the subjects. Reviewing subjects' performances on the main task could provide workload-related measures. Table I summarizes key performance variables in the task. Vehicles Checked is the total of vehicles subjects monitored/checked during the urban scenario, Vehicles in Queue is fully stopped vehicles waiting for clearance, Vehicles in View is vehicles shown on the screen, Threats Missed (Missed Detection or MD) is hostile vehicles identified as non-hostile vehicles, and Non-Threats Recognized (False Alarm or FA) is the number of non-hostile vehicles indentified as hostile vehicles. FA was significantly higher in the urban scenario than in the rural scenario.

As a preliminary step to see whether performances on the main task (i.e., the vehicle monitoring task) and the secondary task (i.e., the peripheral target detection task) affected each other, correlation coefficients between peripheral object detection performance and vehicle monitoring performance variables were calculated as shown in **Table IV**. MD_P is missed detections of peripheral objects, MR_P is threat peripheral objects identified as non-threats, FA_P is non-threat peripheral objects identified as threats, MD_v is threat vehicles identified as non-threats, FA_v is non-threat vehicles identified as threats, $Total_v$ is checked vehicles, $Queue_v$ is vehicles in the queue, and $View_v$ is vehicles displayed on the screen. Correct Detection (CD) and Correct Rejection (CJ) are not included because they can be derived from MD and FA respectively, i.e., $p(CD|threat) + p(MD|threat) = 1$ and $p(CJ|non\text{-}threat) + P(FA|non\text{-}threat) = 1$.

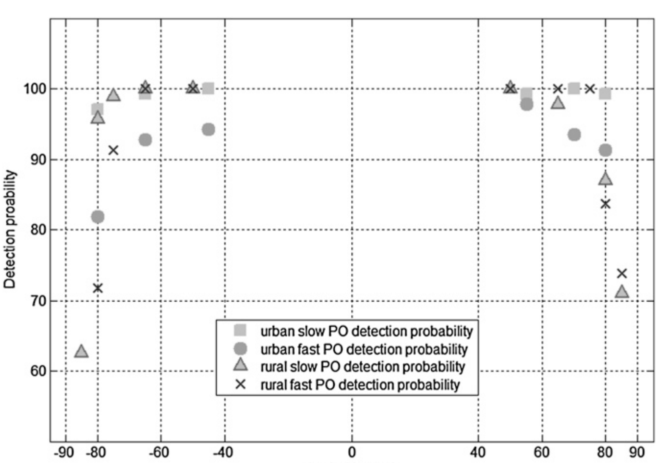

Fig. 3. PO detection probability with respect to peripheral angle.

In the urban scenario, MD_P and MD_v are positively correlated ($P < 0.05$), suggesting that subjects who missed POs more often incorrectly identify foveal threat vehicles as non-threats, or vice versa. MD_P and $Total_v$ are positively correlated ($P < 0.01$), suggesting that subjects who missed POs more had more vehicles examined, or vice versa. If subjects were more focused on the vehicle monitoring task, which could result in examining more vehicles, they were more likely to miss objects that appeared in their periphery. MD_v and $Total_v$ are positively correlated ($P < 0.1$), which shows the tradeoff between accuracy and speed. Subjects who had more vehicles examined (i.e., emphasis on speed) had a higher number of missed detections (i.e., accuracy deterioration).

In the rural scenario, FA_v positively correlated with MD_p ($P < 0.10$), suggesting subjects tend to miss peripheral targets when they falsely identify vehicles as threats. MD_v and MR_p were negatively correlated ($P < 0.10$), suggesting a performance tradeoff between foveal and peripheral vision tasks, i.e., when subjects miss threats in foveal vision, they tend to have less missed recognition in peripheral vision. On the other hand, the positive correlation between FA_v and FA_p ($P < 0.05$) suggests that the more of a tendency there is for identifying targets in the periphery as threats, the stronger the tendency to do the same in foveal vision.

TABLE IV. SPEARMAN'S RANK COEFFICIENT ρ BETWEEN PERFORMANCE VARIABLES.

	MD_P	MR_P	FA_P	MD_v	FA_v	$Total_v$	$Queue_v$
Urban							
MD_P	--						
MR_P	−0.03	--					
FA_P	0.03	−0.14	--				
MD_v	0.42*	−0.30	0.29	--			
FA_v	−0.31	0.10	−0.07	0.01	--		
$Total_v$	0.62**	−0.34	0.23	0.38†	−0.13	--	
$Queue_v$	0.09	0.31	−0.29	−0.02	0.12	0.05	--
$View_v$	0.07	0.20	0.15	0.10	0.09	0.16	0.77**
Rural							
MD_P	--						
MR_P	−0.01	--					
FA_P	0.34	0.30	--				
MD_v	0.24	−0.38†	0.21	--			
FA_v	0.37†	0.12	0.50*	−0.05	--		
$Total_v$	−0.21	0.15	−0.14	−0.21	−0.02	--	
$Queue_v$	0.14	−0.34	−0.05	0.05	0.01	−0.32	--
$View_v$	0.07	−0.30	−0.19	0.10	−0.08	−0.14	0.77**

MD_P = missed detections of peripheral objects, MR_P = threat peripheral objects identified as non-threats, FA_P = non-threat peripheral objects identified as threats, MD_v = threat vehicles identified as non-threats, FA_v = non-threat vehicles identified as threats, $Total_v$ = checked vehicles, $Queue_v$ = vehicles in the queue, and $View_v$ = vehicles displayed on the screen.
† $P < 0.10$, * $P < 0.05$, ** $P < 0.01$.

As part of this project, we developed an interactive 3-D visualization tool, Interactive Virtual Environment and Eye Head Movement Visualization, designed to provide a representation of spatial and temporal correspondence among features scanned in a virtual environment in relation to dynamic changes in the scenario, including peripheral objects and moving vehicles. The visualization replays the scene at various speeds, maps the gaze on the screen features, rotates the head according to head direction measurements, changes the gaze vector color depending on data confidence level, and shows button press actions. By using this tool as shown in **Fig. 4**, we can observe subjects' local gaze points as well as their gaze and head movement patterns. Gaze and head movement data were collected to obtain a better idea of human search strategies, e.g., if there is any pattern for choosing foveal scan area, how actively subjects engage their head and gaze, etc. Observations of participants' macro head and gaze movement showed that subjects employed either "active" or "passive" search strategies during the experiment. We define active search as subjects actively looking for POs by rotating their heads to the left and the right periodically, while we define passive search as subjects focused on the main vehicle monitoring task until they thought they saw POs in their view. There were 11 subjects who showed the active search strategy, whereas 12 showed the passive search strategy. Our data showed no significant differences in expertise levels, i.e., years of service, between subjects who applied active search strategy and passive search strategy. However, do the two different search strategies show any difference in target detection performance?

In the urban scenario, the button reaction time of both the passive and active groups did not show a significant difference for most angle and speed combinations. However, the SRT of the passive search group was significantly greater than that of the active search

Fig. 4. Interactive Virtual Environment and Eye Head Movement Visualization.

group. Although subjects who used the active search strategy had faster saccade detection time than those with the passive search strategy, their overall detection response time, i.e., BRT, was not different from those who applied the passive search strategy. Thus, an active search strategy did not provide observably faster performance than a passive search strategy. However, in the rural scenario, the BRT of the active search group was significantly faster than those of the passive search group for some combinations of entrance speed and angle. Table II shows the P-value in both scenarios whether BRT and SRT are significantly different between subjects who applied active search strategy and passive search strategy. These contradicting results between urban and rural environments need to be investigated more in a future study—for example, how the environmental complexity affects human visual performance.

In terms of accuracy, there was no significant difference between the two groups in the urban scenario. Missed Detections, Missed Recognition, and False Alarms did not differ. However, the passive group had significantly ($P < 0.05$) fewer Missed Detections in their main task, i.e., the vehicle monitoring task. Thus, we can conclude that using a passive search strategy is better for the urban scenario because subjects showed better performance in the main task than when using the active strategy while they did not compromise their peripheral target detection performance. In the rural scenario, the passive search group had more Missed Detections and False Alarms in the peripheral object detection/recognition task ($P < 0.05$) while the main task performance was not different between the two groups. The rural and urban environments resulted in conflicting performances between passive and active search strategies, which requires follow-up research to investigate environmental factors (e.g., scene complexity) influencing target detection performance.

DISCUSSION

Statistical results from human response data (both ocular movement and button response) confirmed that PO detection performance can be predicted by eccentricity (peripheral) angle and entrance speed of the PO. Subjects self-reported their own visual acuity before they started. Since the purpose of the study was not to see the correlation of visual acuity and target detection tasks, we did not collect objective visual acuity. Subjects only had to have sufficient visual acuity to perform the target detection tasks and this was verified during the practice session. It is evident that our subjects had unequal visual acuity. Intersubject differences were modeled via the mixed effect model as shown in the previous section. Regardless of whether intersubject differences come from unequal visual acuity or unequal muscle response time, the mixed effect model is able to handle differences within or between subjects.

The goal of this work was to produce a human performance data based model of target detection in the peripheral field of view that can be used in combat simulations. As a result, a second-order mixed-effect model was applied to provide each subject's prediction model for peripheral target detection performance as a function of eccentricity angle and speed in both urban and rural environments. It was expected that target detection performance in an urban environment would be worse than that in a rural environment due to scene complexity (i.e., number of objects shown in the scene). To the contrary, subjects showed better target detection performance such as higher detection rate and shorter response time in the urban environment. A confounding variable could be the eccentricity angle. POs were placed behind urban features such as trash bins in the urban scenarios, which resulted in a lower eccentricity angle range (±80°) than those of rural environments (±85°). Since our model predicted increased eccentricity angles would decrease the target detection rates, the inconsistent eccentricity angles could affect overall performance. Even when we removed inconsistent eccentricity angles from data analysis and compared target detection performance using only the same eccentricity angles between urban and rural scenarios, e.g., at 65°, this discrepancy still held. There are two possible explanations for this result. First, the entrance motion of POs was different, i.e., translational vs. radial movements. In the urban scenario, pedestrians stepped out from hidden features, which was a translational motion with no size change. In the rural scenario, UAVs appeared from a great distance and increased in size as they approached the subject, which was a radial motion that involved size changes. Either PO sizes or movement type seemed to affect target detection performance. Secondly, many subjects commented that our UAV detection scenario in the rural environment was less realistic than the pedestrian detection in the urban environment. It may be that subjects perform better with familiar tasks than uncommon tasks.

Since the objective was to produce an "effects" model where we can reproduce the net effect of this phenomenon, we have not looked into what the underlying causes may be of our observations. In addition, it was critical for our use to collect data in realistic settings; therefore, we could not constrain the stimuli in a way to facilitate experiments that can tease out the underlying mechanisms. Furthermore, much of the work on the effect of eccentricity looks at relatively small angles (1,2) when compared to the target locations used in our experiments, so it is not clear how those results can be applied to the work done here. As future work it may be beneficial to see how the methodology developed for looking at the effects of eccentricity can be applied to settings where the targets are at extreme angles as in our work.

Our mixed-effect model on peripheral vision effects on target detection will be included in COMBAT XXI to construct more realistic human behavior in that military simulator. Our model, supported by human data, could be compared with existing visual detection models such as the inverse cube law of sighting to enhance understanding of target detection in general. Additionally, it may be used to help inform ways to train soldiers to use search strategies in combat environments. Future study will include overall STA performance comparison between using foveal vision only vs. both foveal and peripheral vision.

ACKNOWLEDGMENTS

This work is funded by the Marine Corps Combat Development Command and the Naval Modeling Simulation Office. Prof. Ron Fricker reviewed the statistical analysis in this paper and we are very thankful. We are grateful to Noah Llyod-Edelman for helping us in calibrating the experimental device.

Authors and affiliation: Ji Hyun Yang, Ph.D., Michael A. Day, M.S., Jesse Aragon Huston, B.A., and Imre L. Balogh, Ph.D., Naval Postgraduate School, Montery, CA.

REFERENCES

1. Carrasco M, Frieder KS. Cortical magnification neutralizes the eccentricity effect in visual search. Vision Res 1997; 37: 63–82.
2. Carrasco M, McLean TL, Katz SM, Frieder KS. Feature asymmetries in visual search: effects of display duration, target eccentricity, orientation and spatial frequency. Vision Res 1998; 38:347–74.
3. Darken C, Jones B. Computer graphics-based target detection for synthetic soldiers. In: Proceedings of Behavior Representation in Modeling & Simulation; March 26-29, 2007; Norfolk, VA. Orlando, FL: Simulation Operability Standards Organization; 2007.
4. Darken C, McCue D, Guerrero M. Realistic fireteam movement in urban environments. In: Proceedings of the Sixth AAAI Conference on Artificial Intelligence and Interactive Digtal Entertainment; October 11-13, 2010; Stanford, CA. Palo Alto, CA: AAAI; 2010:9-14.
5. Henderson JM. Human gaze control during real-world scene perception. Trends Cogn Sci 2003; 7:498–504.
6. Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 1998; 20:1254–9.
7. Itti L, Koch C. Computational modeling of visual attention. Nat Rev Neurosci 2001; 2:194–203.
8. Jungkunz P, Darken CJ. A computational model for human eye-movements in military simulations. Comput Math Organ Theory 2011; 17:229–50.
9. Jungkunz P, Darken CJ. Modeling human eye movements for military simulations. In: Proceedings of Behavior Representation in Modeling and Simulation (BRIMS); March 22-25, 2010; Charleston, SC. Norfolk, VA. Orlando, FL: Simulation Operability Standards Organization; 2010.
10. Lindstrom ML, Bates DM. Nonlinear mixed effects models for repeated measures data. Biometrics 1990; 46:673–87.
11. Navalpakkam V, Itti L. Modeling the influence of task on attention. Vision Res 2005; 45:205–31.
12. Olofsen E, Dinges D, Van Dongen H. Nonlinear mixed-effects modeling: individualization and prediction. Aviat Space Environ Med 2004; 75(3, Suppl.):A134–40.
13. Peters RJ, Itti L. Beyond bottom-up: incorporating task-dependent influences into a computational model of spatial attention. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2007; Minneapolis, MN. Piscataway, NJ: IEEE; 2007.
14. Rayner K. Eye movements in reading and information processing: 20 years of research. Psychol Bull 1998; 124:372–422.
15. Rayner K, Pollatsek A. Eye movements and scene perception. Can J Psychol 1992; 46:342–76.
16. Maurer T, Driggers RG, Wilson DL. Search and detection modeling of military imaging systems. Proc SPIE 2005; 5784:201–15 (Bellingham, WA.).
17. Wolfe J, Kluender K, Levi M, Bartoshuk L, Herz R, et al. Sensation and perception. Sunderland, MA: Sinauer Associates, Inc.; 2006.