

On the relationship between human search strategies, conspicuity and search performance

Maarten A. Hogervorst*, Piet Bijl and Alexander Toet
TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands

ABSTRACT

We determined the relationship between search performance with a limited field of view (FOV) and several scanning- and scene parameters in human observer experiments. The observers (38 trained army scouts) searched through a large search sector for a target (a camouflaged person) on a heath. From trial to trial the target appeared at a different location. With a joystick the observers scanned through a panoramic image (displayed on a PC-monitor) while the scan path was registered. Four conditions were run differing in sensor type (visual or thermal infrared) and window size (large or small). In conditions with a small window size the zoom option could be used. Detection performance was highly dependent on zoom factor and deteriorated when scan speed increased beyond a threshold value. Moreover, the distribution of scan speeds scales with the threshold speed. This indicates that the observers are aware of their limitations and choose a (near) optimal search strategy. We found no correlation between the fraction of detected targets and overall search time for the individual observers, indicating that both are independent measures of individual search performance. Search performance (fraction detected, total search time, time in view for detection) was found to be strongly related to target conspicuity. Moreover, we found the same relationship between search performance and conspicuity for visual and thermal targets. This indicates that search performance can be predicted directly by conspicuity regardless of the sensor type.

Keywords: Search performance, Field of Regard, Field of View, Modelling, Conspicuity, Scan path, Zooming, Detection time, Detection probability, Search strategy

1. INTRODUCTION

An important task in the military environment is to search and detect targets of military interest. In many cases the viewing area is limited due to the use of optics or a sensor system, e.g. a visual camera, image intensifier or thermal imager. In the current study we investigate which factors determine performance with a limited field of view (FOV).

A task directly related to the current study is surveying a designated area or Field of Regard (FOR). Military personnel is trained to scan an area in a specific manner, with scanning procedures depending on the task and function of the scanner. The question is whether these procedures are actually used in practice. That this is not (always) the case was shown in a large scale field experiment organized by NVESD in June 1995 called DISSTAF¹. In this experiment scan patterns of experienced gunners were recorded while performing a realistic search task. The scan patterns were very different from the prescribed procedure. In 90% of the time the Narrow Field of View (NFOV) was used instead of the overview image. Search strategy and search performance differed largely from person to person. The question arises whether the prescribed scanning procedure is optimal. It can also be questioned whether an optimal scanning strategy exists at all (and if it does, what it looks like).

In previous studies²⁻³ we have performed experiments with predefined scan paths to get insight into how search performance depends on basic scan parameters such as scan speed, zoom factor, target location and conspicuity of the target. The results suggest that search performance can be optimized (to a large extent) by taking expected conspicuity into account.

In the current study we recorded free scan paths of trained army scouts. This gives insight into the search strategies used by trained observers. One of the goals was to investigate what characterizes a good scan strategy. We determined how

* E-mail: maarten.hogervorst@tno.nl

search performance depends on the scan parameters and how search performance depends on target conspicuity. Previously, only visual images were used. In the current study we used thermal (IR) images as well. This allowed us to compare search performance for both sensor types.

2. METHODS

We ran a series of experiments in which trained military scouts from the Royal Netherlands Army ($N = 38$) searched for a target in a natural environment. We simulated a situation in which the observer scans the FOR with a visual EO- or an IR-sensor. The observers searched with a limited FOV through a large panoramic image. By using a joystick they could vary the direction and speed of the scan path. The experiments were run on a PC; the images were displayed on a 21 inch computer monitor (40 x 30 cm), with a resolution of 1024 x 768 pixels (pixel depth of 16 bits) running at a frame rate of 75 Hz.

2.1 Stimulus description

The panoramic images covered a search sector of 75 deg x 7 deg in the outside world. Panoramic images were created from a series of images. The visual images were taken with an analogue photographic camera with a 300 mm lens (FOV: 6.9 x 4.6 deg). The slides were digitized to images of 1536 x 1024 pixels, and made into a large panoramic image of 16384 by 1536 pixels. The thermal images were obtained with a digital LWIR-camera (FLIR SCI-2000, 7.5 to 13 μ m), 320x240 pixels, using a lens with horizontal field size of 7 deg. These images were stitched together into a panoramic image of 3505 by 452 pixels.

The scene consisted of a heath scattered with trees. The weather was misty. For each viewing direction images were obtained with both cameras with and without target (a standing camouflaged person). This allowed us to make panoramic images with a target at various locations in the panorama. Copies of the panorama were made with the target blended in at (one out of a number of) different locations. This assured that the target surroundings were the same in the panoramas with and without target. In the visual conditions 45 target locations were used, while in the thermal conditions 51 locations were used (in some cases targets were blended in at other than the original target locations). We also downscaled the panoramas with a zoom factor of 0.25 (see Figure 1), amounting to a visual image of 4096 x 384 pixels and a thermal panorama of 876 x 113 pixels.

In the experiments the panoramas were viewed with a limited FOV. The following 4 conditions (see Figure 1) were run differing in window size and sensor type (visual or thermal):

1. *visL*: visual image, FOV size of 4.7 x 3.5 deg (1024 x 768 pixels), no zoom option
2. *visS*: visual image, FOV size of 2.3 x 1.8 deg (512 x 384 pixels), with zoom option (factor 1 or 0.25)
3. *thermL*: thermal image, FOV size of 4.7 x 3.5 deg (420 x 315 pixels), no zoom option
4. *thermS*: thermal image, FOV size of 2.3 x 1.8 deg (210 x 157 pixels), with zoom option (factor 1 or 0.25)

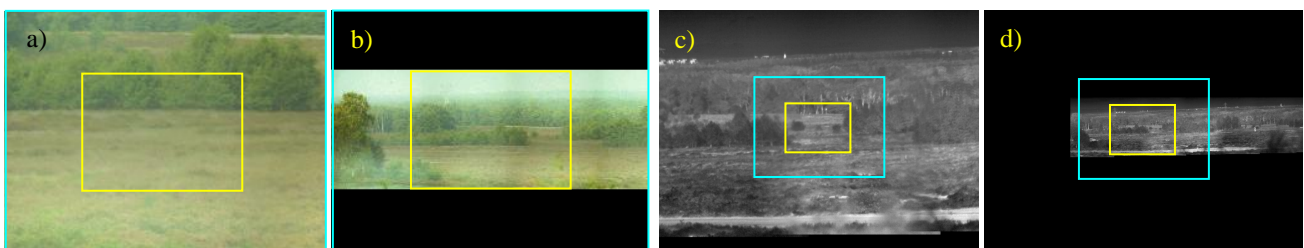


Figure 1. Example insets: showing the part of the panorama displayed on the monitor for zoom factors of 1 (a, c) and zoom factor of 0.25 (b, d) for the visual (a, b) and thermal (c, d) panorama. The small rectangle indicates the window size in the small window conditions (*visS* & *thermS*) and the large rectangle indicates the window size in the large window conditions (*visL* & *thermL*). In the visual conditions the window spans the whole screen in the large window condition. Note that the monitor area is smaller in the thermal conditions than in the visual conditions, but that the displayed outside areas are matched in corresponding conditions.

Examples are shown in Figure 1. In conditions with the zoom option observers could switch back and forth (from a zoom factor of 1) to the scaled panorama (with a zoom factor of 0.25) to get a better overview. Note that the resolution

of the thermal panorama was a factor of 0.41 smaller than that of the visual panorama, but that the FOV size of the covered areas (in degrees) was the same. All images were displayed at the same screen resolution; smaller FOV sizes (i.e. smaller than 1024 x 768 pixels) were created by adding a (cardboard) frame to the monitor that covered the borders of the screen.

2.2 Procedure

In each trial one (and only one) target was present in the scene (and the observer was aware of this). The observer was instructed to find the target as quickly as possible using the controls in an optimal way (zoom option, scan speed, scan path). In conditions with zoom option the observer could zoom in and out by pressing a button on the joystick. As soon as the target was detected the observer should press the fire button. After this, a blank screen appeared and the observer indicated the target location at detection by putting a white cross at the last target location. This was used to check whether the target was found correctly. Targets were regarded as correctly detected when the distance between the indicated position and the target position was less than 100 pixels. The observer was free to choose an optimal distance to the PC-monitor and vary this during the session. If the observer could not find the target the presentation was automatically ended after 80 seconds. Another possibility to the observer -in case he/she was certain the whole scene had been inspected without success- was to indicate the target at a random location. In retrospect, this was a bit unfortunate, since this made it difficult to discriminate between trials in which no target was found and trials in which a false target was indicated. The time between the onset of the stimulus and the button press was recorded as well as the whole scan path over time.

Each observer participated in two sessions, in a visual condition (visL or visS) and a thermal condition (thermL or thermS). The sessions were balanced with respect to temporal order and each condition was performed an equal number of times (each by 19 observers). Within each session the targets appeared in random order. On each trial the presentation started at a random location within the panoramic image with a zoom factor of 1.

3. RESULTS OF THE SEARCH EXPERIMENTS

3.1 Raw data

Figure 2 shows an example of a scan path. Indicated is the centre of the field of view for a zoom factor of 1 (black) or 0.25 (white) over time. Figure 2 shows a successful trial. After initially missing the target (with a zoom factor of 0.25) it was eventually detected with a zoom factor of 1. The target (in the centre of the black rectangle) is relatively small in this case, which explains why it was missed during the first transition.

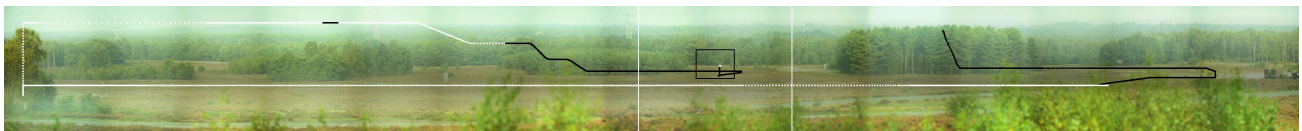


Figure 2. Example of a scan path in the visS condition (visual image, small window) with the target centred in the black window, zoom factor of 1 in black and zoom factor of 0.25 in white. The displayed areas are indicated by the rectangles in corresponding colours.

3.2 Total viewing duration

Figure 3a shows the fraction of detected targets as a function of the viewing duration (taking all targets and observers into account). The viewing duration is the amount of time that has passed since the start of the search. In the visual conditions it takes longer to detect the targets with a smaller window (visS vs. visL). The fraction of targets that is found eventually is about the same in both conditions.

By multiplying the durations of the visS data by a factor of 0.75 (a shift in Figure 3) the data coincide with the visL data. This means that it takes 25% less time to find the same number of targets in condition visL as in condition visS. This is not surprising, since in visS the window is smaller. It therefore takes longer to scan the area with the same zoom factor. In condition visS the observers had the option to switch to a zoom factor of 0.25, which brings a larger (outside) area into view (see Figure 2). However, by zooming out the resolution reduces, which makes it difficult to discriminate the targets from the background (especially small ones). This probably explains why performance is lower when the smaller window is used. Figure 3b shows the fraction of detected targets per second (the detection rate). The peak of

visL is shifted towards lower durations relative to the peak of the visS condition, corresponding to the fact that most of the targets are found earlier in this condition.

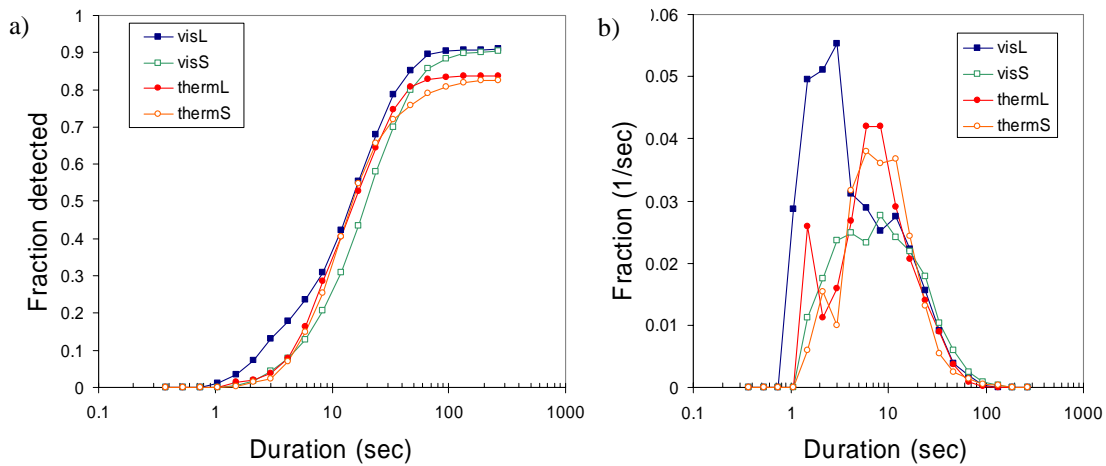


Figure 3. The fraction of detected targets (a) and the fraction of detected targets per time slot (b) as a function of the duration of the trial for all stimulus conditions.

In the thermal conditions the fraction of detections is about the same for both conditions. This pattern of results is different from that observed in the visual conditions. It is probably related to the fact that in these conditions it is relatively easy to detect potential targets even when a zoom factor of 0.25 is used. Note however that in the thermal conditions it is often difficult to eliminate false targets. The targets are displayed with high contrast, but at a low resolution.

After about 40 s the fraction of detected targets is at about 90% of its final value for all conditions. This means that, in our situation, only a small (but in some cases important) increase in performance can be obtained when the search time is increased beyond 40 s.

3.3 Time in view

There is obviously no chance to detect a target before it comes into view. Once the target appears the probability of detection will depend on various factors, such as the time passed since appearing in view, scan speed and zoom factor. Here, we investigate how detection performance depends on these factors. From each scan we extracted the frames in which the target was visible and divided these into periods (transitions) in which the target was continuously visible (e.g. the scan in Figure 2 contains 3 transitions). The first and last transitions are analysed further*. The first transitions give insight in the probability of detection for given scan parameters. The last transitions (that contain more detections) give insight in the time it takes to detect a (detected) target.

3.3.1 Modelling fraction detected

First, we look into how detection performance depends on the time passed since it appeared in view (time in view) and how this can be modelled. Figure 4 shows the cumulative fraction of detected targets as a function of the time a target is in view (the “detection time”) for the visual and thermal conditions, in which all trials (scans and observers) were taken into account, for the visual and thermal conditions. Separate curves are shown for the first and last transitions for conditions with a small and large window.

The fraction of detected targets can be described by⁴:

$$P(t) = P_{\infty} \left(1 - \exp\left(-\frac{(t-t_o)}{\tau}\right)\right) \quad , \text{ for } t > t_o \quad (1)$$

* In some cases the first and last transitions coincide.

in which $P(t)$ is the fraction of detections after duration t , t_0 is an offset in time; τ indicates the steepness of the curve, and P_∞ corresponds to the fraction that is detected after long inspection times. We fitted the parameters P_∞ , t_0 and τ to the data of the various conditions (using a common least squares method) and tested the following hypotheses for each image type (separate for visual and thermal images):

1. t_0 is the same for all conditions
2. τ is the same for the *last* and the *first* transitions for each sensor for a given condition
3. P_∞ for the large window conditions is the same as for the small window conditions for a given transition

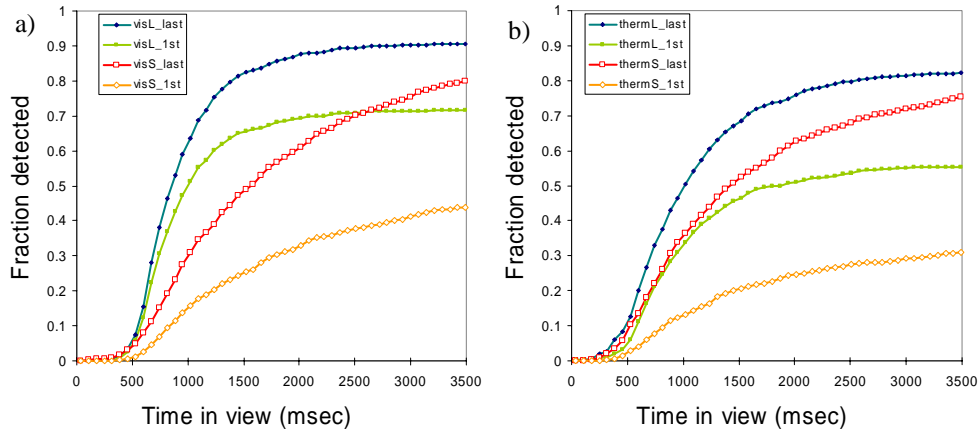


Figure 4. Fraction of detections as a function of the time a target is in view, for the visual conditions (a) and for thermal conditions (b). Separate curves are displayed depending on window size (small S and large L) and transition (first or last).

Hypothesis 1 stems from the idea that t_0 is likely to be dominated by the reaction time and therefore independent of condition type. Hypothesis 2 states that the rate at which targets are found is independent of the transition; this would mean that it is as difficult to find the target in the second transition as in the first transition. The 3rd hypothesis states that eventually the same fraction of targets is found, independent of window size and zoom option.

TABLE 1. Fit parameters of free fit.

free fit	deviation χ	t_0 (ms)	τ (ms)	P_∞
visL_last	0.0086	574	479	0.904
visL_1st	0.0068	576	469	0.716
visS_last	0.0070	539	1462	0.882
visS_1st	0.0035	591	1503	0.489
thermL_last	0.0118	464	758	0.837
thermL_1st	0.0061	507	694	0.558
thermS_last	0.0077	406	1274	0.801
thermS_1st	0.0032	477	1289	0.327

The results of the fitting procedure are shown in Table 1. The values of t_0 differ for the visual and the thermal conditions, which is probably related to the fact that in the thermal conditions target contrast was much higher than in the visual conditions. The second hypothesis appears to hold: the τ 's for the first and last transitions are similar. The third hypothesis means that the P_∞ of the corresponding large and small window conditions are the same, e.g. the P_∞ of visL_last and visS_last should match. This hypothesis holds reasonably well for the last transitions, but not for the first transitions. We tested these hypotheses further by i) forcing the same t_0 on all conditions (but different ones for the thermal and visual conditions), ii) forcing the same τ on the *last* and *first* transitions and iii) forcing the same P_∞ for the large window-conditions as for the small window-conditions and observing how the goodness of fit (χ) changes under

these restrictions. This analysis reconfirms the above. It shows that all hypotheses are true (leading to a small increase in the goodness of fit) except for hypothesis 3, which holds for the last, but not for the first transitions.

3.3.2 Time in view and zoom factor

We determined how the probability of detection depends on i) the time a target remains in view and ii) the zoom factor. To this end we sorted all data of the first transitions according to the time a target remained in view. Figure 5a shows (as an example) for condition visL the overall distribution of targets together with the distribution of detected and undetected targets. The time in view ranges from 0.1 to 10 s. Note that the shape of the detected curve is similar to the slope of the curve displayed in Figure 4.

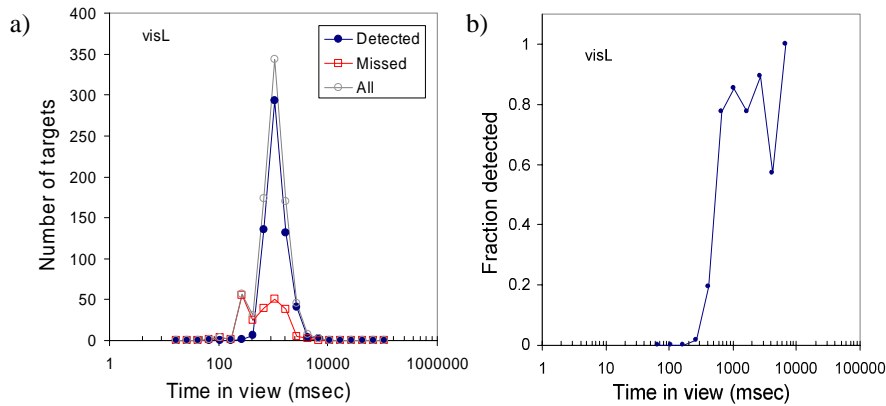


Figure 5. a) The distribution of targets as a function of the time in view of the first transitions for condition visL (visual imager, large window). Shown are the distributions of all targets (open circles), the detected targets (closed circles) and missed targets (open squares). b) The fraction of detected targets as a function of duration.

Figure 5 shows that for short durations the detection probability is low. The fraction of detected targets increases steeply from zero to a constant value (see Figure 5b) around durations of 500 msec (≈ 40 frames). This is due to the fact that for times shorter than the reaction time hardly any targets will be detected. The other conditions (not shown here) show a similar pattern. We divided the data into durations smaller and durations larger than 40 frames. For each transition we also calculated the average zoom factor (this is the geometric average over the period the target was in the window), and divided the data into a set with a zoom factor equal or smaller than 0.29 and a set in which it was larger than 0.29 (threshold value based on the distribution). Table 2 shows the fraction of detected targets for various selections.

TABLE 2. Fraction of detected targets of the first transitions for all conditions and various data selections: all the data (all), taking only small ($dur < 500$ ms) and long ($dur \geq 500$ ms) durations into account, small zoom ($z < 0.29$) or large zoom ($z \geq 0.29$) or a cross section of the two.

	visL	visS	thermL	thermS
all	0.73	0.50	0.61	0.36
$dur < 500$ ms	0.07	0.01	0.12	0.06
$dur \geq 500$ ms	0.82	0.55	0.69	0.39
$z \leq 0.29$		0.42		0.29
$z > 0.29$		0.80		0.70
$z \leq 0.29$ & $dur \geq 500$ ms		0.45		0.32
$z > 0.29$ & $dur \geq 500$ ms		0.89		0.79

Table 2 shows that the fraction of detection (in the first transitions) was larger in the visual conditions than in the thermal conditions. As expected, the fraction detected is smaller in the conditions with the smaller windows. The detected fraction is low for durations shorter than the reaction time (duration < 500 ms). Consequently, the fraction detected is somewhat higher when only transitions with targets longer than 500 ms in view are considered than when all transitions are considered.

In conditions with a small window the zoom option could be used. The detected fraction is much lower in transitions in which a large zoom factor was used than when a small zoom factor was used.

The performance in the small window conditions with large zoom factor can be compared directly with the large window conditions, since the resolution is the same. That the detected fraction is higher in the small window condition can be explained from the fact that in this case a smaller area has to be inspected.

The results indicate a reaction time of about 0.5 s. When the targets appear into view for longer detection performance does not improve much further (see e.g. Figure 5b, although more data is required to draw a firm conclusion). Secondly, the results show that performance is highly dependent on the zoom factor.

3.3.3 Scan speed

In a similar way we analyzed the speeds of the targets during the first transitions, taking only targets into account that remained into view for longer than the reaction time (duration = 500 ms). Figure 6a shows as an example the speed distribution of condition visL. Shown is the number of targets within a given range of speeds as a function of the average speed (the geometric average over the viewing duration). Figure 6b shows the fraction of detected targets as a function of the speed. Performance decreases steeply above a certain threshold speed.

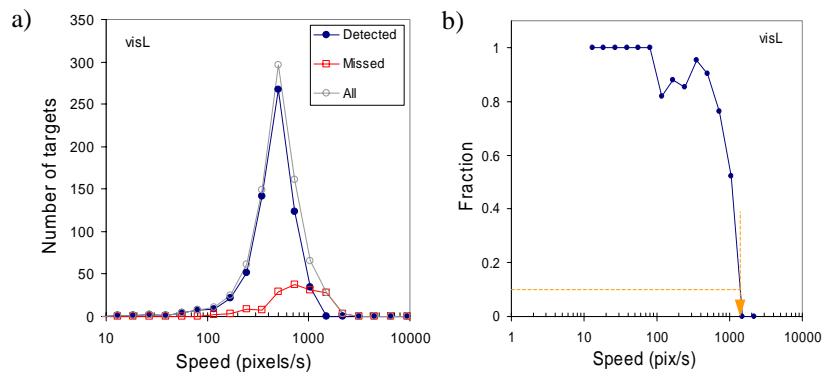


Figure 6. Figure 6a shows the distribution of targets as a function of the mean scan speed of the first transitions for condition visL (visual image, large window). Shown are the distributions of all targets (open circles), detected targets (closed circles) and the missed targets (open squares). Figure 6b shows the fraction of detected targets as a function of speed. The arrow indicates the threshold speed corresponding to a fraction of detected targets of 0.1.

We calculated the speed at which the fraction of correct detections crosses a value of 0.1 (see Figure 6b). These (thus defined) threshold speeds were found to be 1404 (visL), 674 (thermL), 483 (visS/small zoom), 233 (thermS/small zoom) and 2728 (thermS/large zoom) pixels per second. No threshold speed was found for the visual condition with small window and large zoom factor (visS/large zoom). One can use the following model assumptions. First, one can assume that the threshold speed scales in proportion to the window size in case the displayed area scales proportionally. This means that the threshold speed in terms of visual angle (deg/s) in the outdoor scene is independent of the image size (as long as the same area is displayed). Second, one can assume that the threshold speed increases when the displayed area decreases, since it takes shorter to inspect a smaller area. One may for instance assume that the threshold speed is proportional to the inverse of the area that needs inspection. Note that only the area in which targets are expected needs inspection (e.g. for ground targets the sky can be ignored).

Figure 7 shows the speed threshold expressed in terms of visual angle (in deg/s in the outside world) for all conditions. Whereas the threshold speed in pixels/s varies largely from condition to condition, it is relatively constant when expressed in deg/s: 6.4 (visL), 7.5 (thermL), 8.8 (visS/small zoom), 10.4 (thermS/small zoom) and 30.5 (thermS/large zoom) deg/s.

The threshold speed in condition thermS/large zoom is however considerably larger than in the other conditions. This is consistent with the 2nd assumption. The threshold speed is about 4 times larger in condition thermS/large zoom than in condition thermL while the displayed area (of the outdoor scene) is 4 times smaller. This might also explain why no threshold speed was found in condition visS/large zoom. In this case one predicts the threshold speed to be 4x larger than in condition visL (i.e. 5616 pixels/s), and this falls outside the (observed) speed distribution.

An interesting comparison can be made between the small-window/small-zoom and the large-window conditions. The window size is smaller in the first than in the second condition, but the displayed outdoor area is larger. The second model assumption predicts the threshold speeds in the small-window/small-zoom conditions to be (4x) larger than in the large-window conditions. In correspondence with the prediction, the threshold speeds are larger, but the difference is smaller than predicted (about 1.4x). This is likely related to the fact that the area in which targets appear is smaller than the area displayed in the small-window/small-zoom condition (see Figure 1). The sky can for instance be ignored.

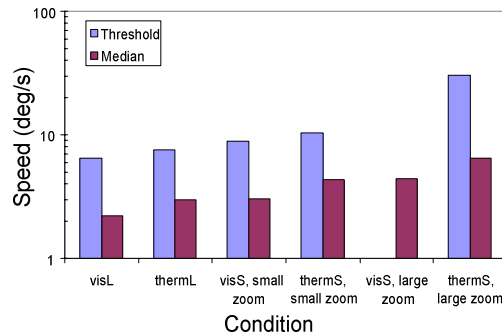


Figure 7. Threshold speed expressed in visual angle (deg/s) in the outdoor scene for all conditions, along with the median values of the speed distribution. The data of the small window conditions are separated into transitions in which the small and the large zoom factor were used. No threshold speed was found for condition (visS/large zoom).

Also shown in Figure 7 are the median speeds. They approximately scale in proportion with the threshold speed. This suggests that the observers used a strategy in which the threshold speeds are in the tail of the speed distribution, thereby stretching detection performance to its limits. Only when the threshold speeds become too large the observer uses lower speeds. This suggests that the observers are aware of their own limitations and try to use an optimal strategy.

3.4 Individual observer performance

Figure 8a shows for each observer separately average search time (geometric average over all detected targets) versus fraction of detected targets for all conditions. Average search time and fraction of detection are largely uncorrelated. Only in condition thermL average search time is *positively* correlated with the fraction of detected targets ($p=0.02$).

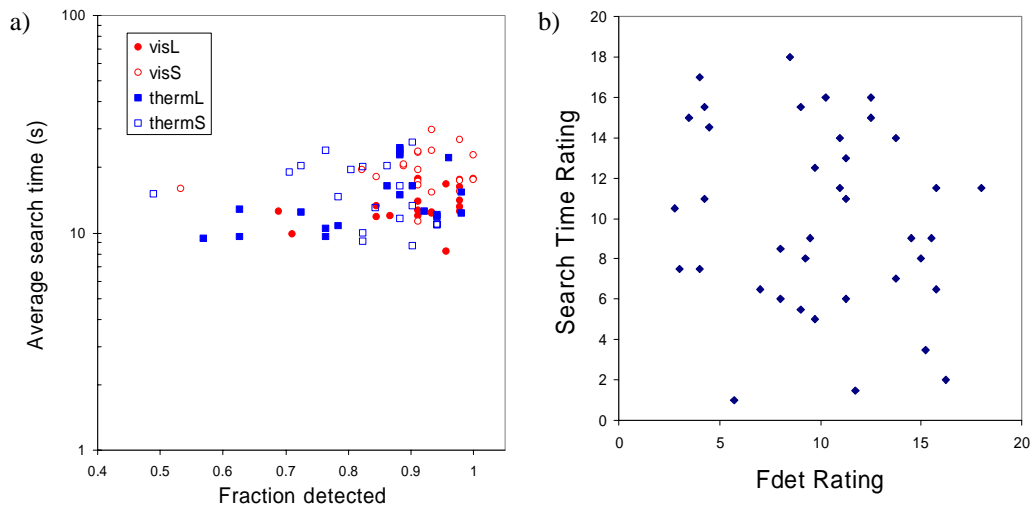


Figure 8. a) Average search time (average over all detected targets) versus Fraction of detected targets for each of the observers for all conditions. b) Performance ratings for average search time vs. fraction detection for each observer. Small values correspond to good performance.

Each individual can choose between performing an accurate, but slow search, or a fast search with a higher probability of losses (speed-accuracy trade-off). Differences in performance between observers cannot be explained from

differences in this optimization criterion since this would lead to a negative correlation. As expected, average search times are higher and detection rates are lower for small window conditions than for large window conditions.

For each observer, performance ranking scores were derived for average detection speed (search time) and fraction of detection as follows. For each condition (visL, thermL, visS, thermS) points were awarded to each observer on the basis of their performance ranking. The observer with the lowest fraction of detection was awarded 1 point, the next one 2 points etc. When performance was the same, the average rating was awarded to both. Finally, the rating of the two conditions performed by the observer was averaged, resulting in a rating score for the observer with respect to search time and fraction of detection. Figure 8b shows search time rating versus detection rating for all observers. Again, no correlation between these ratings is found.

Inspection of the scan paths of the individual observers (e.g. overlap) did not show which characteristics of the scan strategy could account for the differences in performance.

4. CONSPICUITY

TNO Human Factors developed a psychophysical method to quantify the visual conspicuity of a target in a complex scene⁵. This measure can be obtained quickly and only a few observers are required to achieve a reasonable accuracy. The measure is directly related to the conspicuity area defined as the area around the point of fixation within which the target can be distinguished from its background⁶⁻⁸. It has been shown that the conspicuity measure developed at TNO correlates well with search performance (mean search time, detection probability) in situations with a static FOV⁹⁻¹⁰. Also, in a similar experiment with predefined scan paths detection performance was found to correlate with conspicuity²⁻³.

Here we determine the extent to which detection performance can be predicted by target conspicuity under free scanning conditions. While previous experiments were performed with a single sensor, we are now able to compare performance for a visual and a thermal sensor. We will determine the extent to which performance can be predicted on the basis of conspicuity and whether this depends of the sensor.

4.1 Method

The conspicuity area of a target is defined as the region around the centre of the visual field where the target is capable of attracting visual attention, because it is perceived as significantly distinct from its local background. The conspicuity measurement procedure introduced here is as follows. First, the observer visually inspects (foveates) the target. Next, the observer fixates a point in the scene that is at a large distance from the target location. When the observer fixates this point, the target should be positioned so far in his peripheral visual field that it can not be distinguished. The observer then successively fixates locations in the scene that are progressively closer to the target location, until he can perceive the target in his peripheral FOV. The successive fixation points are along a line through the initial fixation point and the centre of the target. The distance between the fixation location at which the target is first noted and the centre of the target is then recorded (expressed here in deg of visual angle in the outdoor scene). Previously²⁻³, we obtained conspicuity estimates from 3 experts (MH, PB and JV) for all 45 visual targets using the large panoramic image (zoom factor of 1). Here, we also obtained conspicuity estimates from 3 experts (MH, LT and JV) for all thermal targets.

4.2 Results

The conspicuity estimates given by the various experts are correlated, with correlation coefficients for visual targets of 0.66 (JV vs. MH), 0.73 (JV vs. PB) and 0.76 (MH vs. PB), and correlations for thermal targets of 0.70 (JV vs. MH), 0.71 (JV vs. LT) and 0.78 (LT vs. MH). The fact that the correlation coefficients are relatively low is likely due to the fact that for each target only one estimate was obtained. We will use the geometric average of the conspicuity estimates of the 3 experts as our best estimate of the conspicuity of the target.

We also determined whether detection performance is related to the vertical position of the target. When the vertical position is larger this means that it is closer to the horizon, and is displayed smaller. As in previous studies²⁻³ we found a negative correlation between target conspicuity and vertical position: $R = -0.67$ for visual targets and $R = -0.77$ for thermal targets.

For each target we calculated the fraction of observers that detected the target in the first and last transitions, as well as the (geometric) averages of search time and detection time over observers (taking only the detected trials into account). Search time refers to the total time required for detection (see Figure 3) and detection time refers to the time a target is in view before it is detected (see Figure 4). Table 3 shows the correlations of these performance measures for each target with target conspicuity and the vertical position of the target. The significance level is indicated by the font style.

TABLE 3. Correlations between various (average) detection performance measures for each target and target properties (conspicuity and vertical position Y) for all conditions, in which Fdet refers to the fraction of detected targets, search time refers to the (average) total time required to detect the target and detection time refers to the average time a target is visible before it is detected, for first and last transitions. Normal font indicates that the correlation is not significant, italic a significance level of $p < 0.05$, and italic bold a significance level of $p < 0.01$.

x	y	visL_last	visL_first	visS_last	visS_first	thermL_last	thermL_first	thermS_last	thermS_first
Conspicuity	Fdet	0.31	0.46	0.56	0.71	0.47	0.48	0.67	0.74
Conspicuity	Search time	-0.06		-0.58		-0.45		-0.50	
Conspicuity	Detection time	-0.52	-0.40	-0.72	-0.63	-0.71	-0.68	-0.46	-0.28
Y	Fdet	0.12	-0.01	-0.29	-0.56	-0.32	-0.38	-0.36	-0.60
Y	Search time	-0.27				0.39		0.28	
Y	Detection time	0.44	0.42	0.47	0.38	0.52	0.53	0.35	0.19

Generally speaking, when target conspicuity increases (or the vertical position decreases), fraction detected increases and average search time and detection time decrease. Conspicuity gives a somewhat better prediction of the fraction of detections during the first transition (e.g. visL_first) than in the last transition (e.g. visL). The same holds for the vertical position. On the other hand, detection time is better predicted by conspicuity (or vertical position) for the last transitions than for the first transitions. In the visL_last condition the correlation between fraction detected and conspicuity is relatively low, since in this condition most of the targets were found eventually. Therefore for many targets the detection fraction was (close to) one leading to a relatively low correlation coefficient.

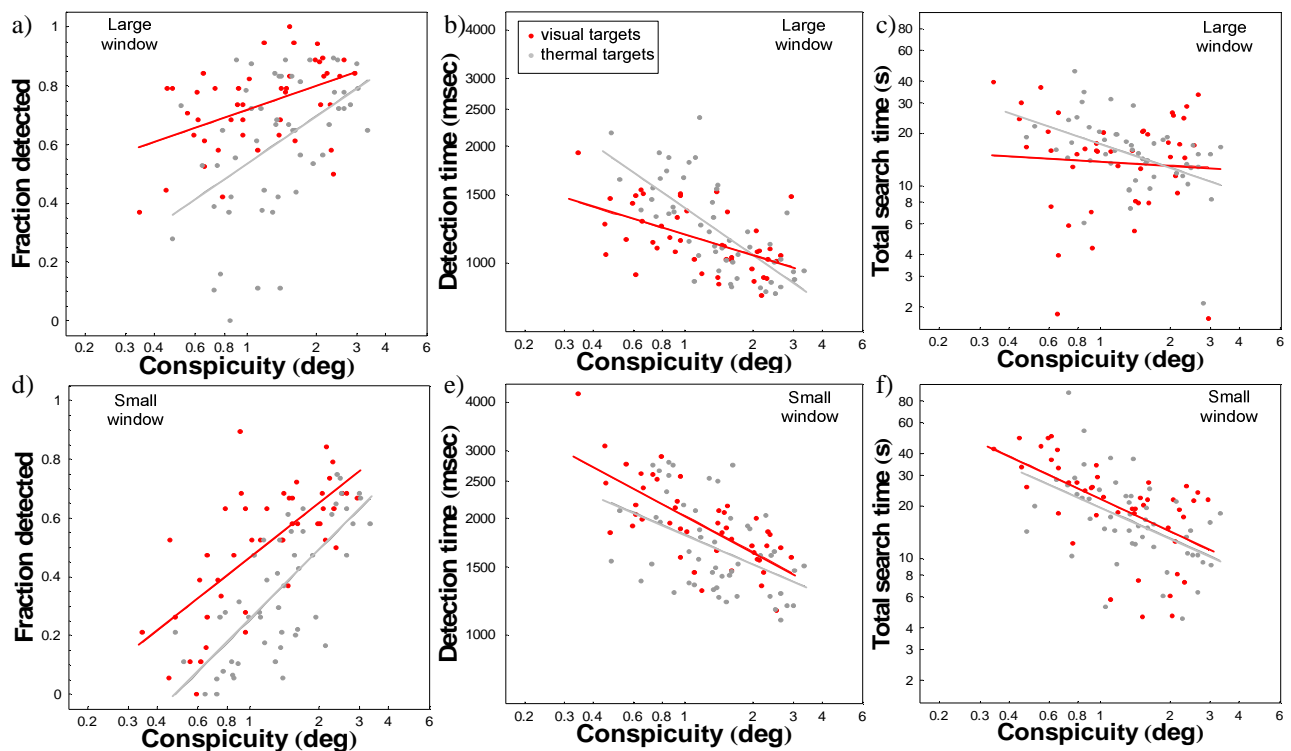


Figure 9. Several performance measures (fraction detected: a, d; average detection time: b, e; average total search time: c, f) versus target conspicuity (expressed in the visual angle in the outdoor scene, i.e. in deg) of all targets for visual (dark, red) and thermal targets (light, gray). The top row shows data for the conditions with a large window (visL & thermL) and the bottom row shows the conditions with a small window (visS & thermS). The drawn lines represent linear regression lines to the data.

Figure 9 shows the relationships between various performance measures (fraction detected in first transition, detection time and total search time) and target conspicuity for visual and thermal targets. The top row shows the relationships for large window conditions and the bottom for the small window conditions. The relationship between the search performance measures and target conspicuity is highly similar for visual and thermal targets. Note that conspicuity is (only) one of the factors determining search performance. This explains the considerable spread around the fitted lines.

The fact that the fraction of detections is slightly lower for thermal targets may be due to the fact that some (small) targets are difficult to discriminate from false targets. This would explain why the difference decreases with increasing conspicuity. Analysis of the fraction of detection of the last transition (not shown here) gives a similar result.

Summarizing, the results show that the relationship between search performance and conspicuity is largely independent of the type of sensor.

5. DISCUSSION

In previous studies²⁻³ we investigated how search performance depends on scanning parameters and target properties by using predefined scan paths. In the current study observers (army scouts) performed a free search to find the targets. This study gives insight into (properties of) scanning strategies and in how they affect search performance. We found that scanning speed should not exceed a maximum, which depends on window size and FOV size. We found that the speed distribution was adapted to the circumstances, such that only in a few (but some) instances the speed exceeded the threshold maximum. These findings suggest that human observers are aware of their limitations and are well capable of adapting their strategy to a (near) optimal one. In agreement with the optimal strategy idea is that in previous studies²⁻³ we obtained evidence that eye movement patterns are (to a large extent) optimized by taking expected target conspicuity into account. These results suggest that search strategies of (trained) observers can be modelled by assuming an optimal scanning strategy. Such a model also gives insight into the highest search performance that can be reached in a certain situation given the limitations of the visual system.

The results obtained here may be partly biased by the search strategy. For instance, the fact that the average speed of detected targets is smaller than that of missed targets (see Figure 6) may be partly due to a (potential) tendency to slow down before indicating detection. This shows that studies with and without predefined scan paths are both necessary to get insight in search strategies and the way search performance is related to properties of the scan path.

On the one hand, one might expect that observers that are good at the search task have good scores for both performance measures. On the other hand, when differences in performance originate from individual differences in the choice of what is the best balance between speed and accuracy, one would predict a negative correlation. Instead, we found that these measures were uncorrelated. This means they are independent measures of search performance. In some situations (with no time limit) the emphasis will lie on finding *all* targets. In other situations there may be a limit to the search time. In practice the circumstances will define which behaviour is most desirable. This can be formalised by introducing a cost function associated with certain outcomes.

As a model of search performance equation (1) can be used. For our situation the model parameters were derived for the various conditions. This model describes the accumulative fraction of targets that will be detected a certain time after appearing in view. Another useful model can be derived from Figure 5. This model is shown in Figure 10 for condition visL for the first transition. In the model the distribution of durations is modelled by a Gaussian distribution (in the logarithmic domain). Furthermore, the model assumption is used that the probability of detection is zero for durations smaller than 0.5 s (the reaction time) and constant for larger durations. Figure 10 shows that the model gives a reasonable prediction of the cumulative fraction of detected targets as a function of duration (time in view) in this case (visL, first transitions), although time constant τ is somewhat smaller than in reality. In a similar way, Figure 6 can be used for modelling, using a Gaussian speed distribution and a threshold speed. Note that the probability of detection decreases more gradually with speed, such that a step function is less appropriate here. Useful properties (for modelling) are that the threshold speed and the speed distribution (in deg/s) are largely determined by the size of the outside area that is displayed. A mere magnification of the displayed image has little effect. When the displayed area is diminished

the threshold speed (and the speed distribution) increases (in proportion with the inverse of the area when the resolution remains the same).

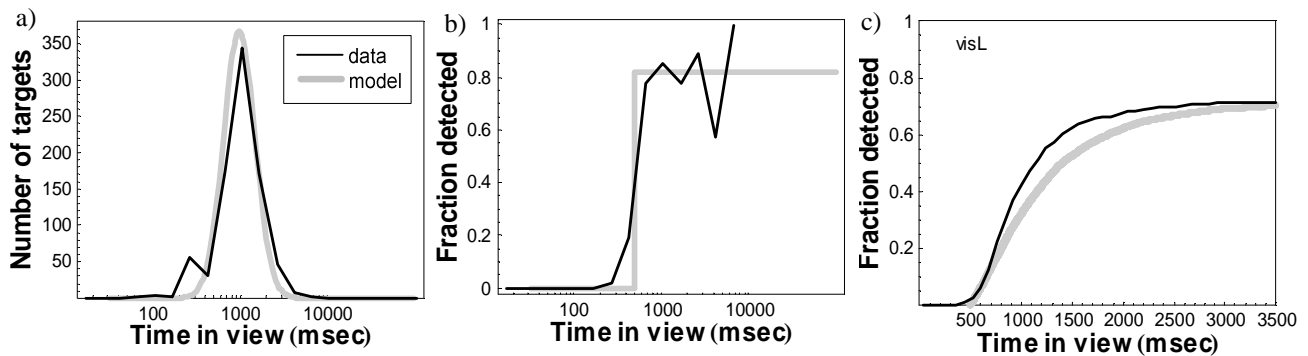


Figure 10. Model of detection probability (grey) vs. the measurements (black). This example shows the data of condition visL during the first transition. The distribution of durations is modelled by a Gaussian distribution in logarithmic units (Figure 10a), and the probability of detection is modelled by the stepfunction displayed in Figure 10b. Monte Carlo simulation with this model predicts the accumulated fraction of detections (Figure 10c).

The model described above captures some general characteristics of the scanning process. A more elaborate model incorporates (also) conspicuity, since search performance is largely dependent on target conspicuity. The main finding of this study is that the relationship between search performance and conspicuity is largely independent of the sensor type. One may even draw the tentative conclusion that the relationship between search performance and target conspicuity is largely independent of image content, which would mean that for a given scan path search performance can be predicted on the basis of the target conspicuity regardless of other image properties.

On the basis of this and previous studies²⁻³ a human observer model can be constructed predicting search performance with a limited FOV. This consists of two parts: i) a model that predicts detection probability for a given scan path and target conspicuity, and ii) a statistical model of search strategy based on the idea that a trained observer will optimize his/her scan path properties by taking expected conspicuity into account as well as the limitations to detection (for given scan path properties).

6. REFERENCES

1. DISSTAF field trials: organized by NVEDS (Night Vision & Electronic Sensors Directorate, Ft. Belvoir, Virginia) in June 1995 in Ft. Hunter Liggett, California, USA
2. Hogervorst, M.A., Toet, A., Bijl, P. (2002) What is the optimal scan path for search with a limited field of view? *Perception Supplement*, **31**, 173b.
3. Hogervorst, M.A., Bijl, P., Toet, A., Valeton, J.M. (2002) Human search with a limited Field of View: the effect of scanning parameters and scene content. *SPIE Proceedings*, **4718**, 83-94.
4. Toet, A., Bijl, P. (2003) Visual search. *Encyclopedia of Optical Engineering* ed. by Marcel Dekker, Marcel Dekker Inc., New York, 2949-2954.
5. Toet, A., Kooi, F.L. (1999). Conspicuity: An efficient alternative for search time. *Vision in vehicles – VII* ed. by A.G.Gale, Elsevier, Oxford, p451-462.
6. Engel, F. L. (1971) Visual conspicuity. Directed attention and retinal locus. *Vision Research*, **11**, 563-575.
7. Engel, F. L. (1974) Visual conspicuity and selective background interference in eccentric vision. *Vision Research*, **14**, 459-471.
8. Engel, F. L. (1977) Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*, **17**, 95-100.
9. Toet, A., Kooi, F.L., Bijl, P., Valeton, J.M. (1998a). Visual conspicuity determines human target acquisition performance. *Optical Engineering*, **37** (7), 1969 - 1975.
10. Toet, A., Bijl, P., Kooi, F.L., and Valeton, J.M. (1998b). Quantifying target distinctness through visual conspicuity. *Proceedings of the SPIE Conference on Targets and Backgrounds, vol. SPIE-3375*, **8**, 152-163.