THE SOURCES OF VARIABILITY IN THE SEARCH PROCESS

K. Cooke
BAe SRC
FPC267
Filton, Bristol
BS34 7QW
England.
E-mail: kevin.cooke@src.bae.co.uk

1. SUMMARY

Modelling of camouflage concealment and detection, needs to consider the terrain in which a target will appear. A variety of capabilities for evaluating target signatures through to the human response now exists. The modelling approach adopted at British Aerospace research centre for many years has been a statistical one. Although image analysis techniques have been explored it has been cost effective for our purposes to stay with the statistical model ORACLE.\(^1\) A complex problem in modelling human visual performance is to find an adequate relationship between recognition thresholds across the visual field and simple target descriptions. The ORACLE model represents recognition as the resolution of a fraction of the target perimeter. The fractional perimeter concept has been applied further to representing average observer performance in structured scenes. Modelling of the SEARCH 2 data\(^2\) was found to need a similar distribution of fractional perimeter values to a previous UK field trial. The use of a statistical lobe model for analysing search and recognition can be supported for generic information.

**Keywords** ORACLE, visual lobes, recognition, identification, search variables, field data, peripheral, search time.

2. INTRODUCTION

The issues raised when considering target acquisition can be divided into the visual and the cognitive components. From a physical design stance the cognitive issues are of less interest than the visual issues because physical design can impose specific visual limits on system performance. The development of ORACLE in British Aerospace has been focused on the products of the company and how to optimise visual performance. Around 1983 a decision was made not to attempt to model performance in specific scenes. At the time we were studying eye movement patterns and we were synthetically altering clutter levels in scenes. Computer power at the time made the analysis of the data slow compared to today but we also realised that the scenes we were analysing took a long time to generate and our customer base was mostly interested in the relative performance of one system design against another. Our belief was that effort spent on trying to unravel the complex interactions of clutter, strategy, training philosophies, etc, for inclusion in the ORACLE vision modelling approach was not cost effective. We had at the time developed modelling primarily for ground to air acquisition tasks and relatively uncluttered or plain backgrounds were appropriate. We chose to build a search model based on the detection visual lobe and its relationship to search in plain fields of view. The search modelling has remained unchanged in ORACLE for the last 16 years. We rationalised cluttered scene search to the use of visual lobes based on recognition or identification criteria. Search is a combination of peripheral cueing and foveal interrogation. A target in a structured scene may be surrounded by confusable objects and the task of the observer is to use peripheral discrimination for guiding fixation. Our belief was that at any one time an observer is using what could be thought of as a multitude of visual lobes operating from pure energy detection through to discrimination and that the largest signal to the visual system provides the cue for the next point of fixation. A far peripheral target that is highly detectable may provide the same strength of cue for fixation as a near but confusable object. We have attempted to model the variance experienced in structured scene search by using a set of visual lobes to allow for observer, target, and background statistics. In this way we hope to have encompassed the range of statistical variance that has been measured in field or laboratory trials. This simplification has avoided detailed consideration of the influence of strategy and other cognitive elements of the search task for purely pragmatic reasons. Our approach has excluded specifying cognitive processes that occur when an observer is searching in a specific scene.

In this paper I would like to provide an indication of some areas we have investigated for lobe modelling and its relevance to sensor design. Also to provide a feel for the sensitivity of the model and why we have addressed some issues and not others. As long as models have the sensitivity to deal with the parameters that influence visual performance in sensor/display design then advances along the cognitive dimension of the search task will hopefully enable further improvements in total system performance prediction.

3. DEVELOPMENT STRATEGY

We have attempted to make the models 'user friendly' to non vision scientists and to engineers with access to standard computer systems. As the outside world poses system limitations the early work was devoted to providing measures of target and atmospheric characteristics. Our philosophy was to calculate separately parameters of the displayed information reaching the cornea and then apply a generic model of human vision which would be applicable to any sensor/display system. I believe that the processes contained within models for representing human vision are less important than the model output. If we had test data sets agreed between modellers the method of modelling them could be independently derived, and I applaud the aim of this meeting to work with a common data set. Some models have aimed at predicting the threshold surface for size, contrast, and retinal position and several early data sets from Blackwell\(^3\) have been well used for test purposes. We also know that tunnel vision makes search performance poorer, thus emphasising the importance of peripheral vision. We have spent

---

the bulk of our time trying to get the local retinal performance modeling accurate by concentrating on the physical attributes of a stimulus such as size, contrast, colour, motion, and image quality across the whole visual field. These are relevant to lobe evaluation. It is only when we have the lobe well represented that we can begin to look at the nature of successive fixations in further depth. Hence the still very rudimentary nature of our search model\(^2\). Our search model evaluates the target against its immediate background but does not analyze the visual information contained in the rest of the FoV and does not include specific observer strategies in the equations. The need to account for scene interactions becomes a necessary part of specific scene analysis but is less crucial for the generic scenario. The variance is great as is highlighted in the SEARCH 2 dataset as well as many others.

There are now a variety of direct image processing models which can process the whole scene rather than just the local target to background contrasts. As an industrial group we are still debating the cost effectiveness of attempting to do this using the ORACLE concepts. We still ask the question whether by adding several orders of complexity would we now meet customer requirements to a significantly better degree? We have continued to devote our efforts to the more easily quantifiable processes leaving the cognitive variances of stress, training, learning, etc as poorly quantified components of our observers.

The calibration of the sensor from which digital data are recorded should be well understood if full image analysis is to be correctly utilised. The range of intensity colour levels achievable when displaying images is restricted compared to the real world and some effort would need to be devoted to creating equivalent visibility images. Calculation of visual performance in extreme viewing scenarios has been a crucial point in the application of our model. Atmospheric effects on a field evaluation or when simulating a future event are not very predictable and are not controllable, and the resulting effects on contrast and image sharpness can be large.

The rest of this paper is a review of issues to do with applying lobes to search tasks concentrating on some of the fundamental components of our vision model and the associated data sources that have proved useful for calculating visual lobes for application to acquisition.

There are still many questions for which there are insufficient data to establish model validation. For example can we recognize objects with rod vision or do we rely solely on cone vision. The interface between rods and cones is not easily bridged and pure rod or pure cone data are sparse.

4. RECOGNITION

Representing recognition in a statistical model involves simplification of the target signature. The approach in ORACLE is similar to others such as the Johnson criterion which is essentially an equivalent resolution, and to the equivalent disc concept proposed by van Meerten\(^3\). Immediate limitations of the above are that the Johnson resolution\(^4\) bar patterns operate only in one dimension and the equivalent disc fails at resolution limits. Our approach is to define the resolution requirements to be a fraction of the vertical and horizontal dimensions of the target. This alone does not provide the necessary behaviour to deal with all size-contrast regimes and the second component of the model, more recently introduced, requires a minimum number of elements of the target object to be separately resolved for recognition to be achieved.

This aspect of our modelling has often required the greatest explanation to potential users of the model as the implementation of the equation to represent recognition reduces essentially to resolving features in a statistical fashion but does not specify exactly which features of a target are represented.

The support for our approach is derived from a series of experiments which provide the calibration and test data for establishing the fractional dimensions of the target perimeters that correlated to recognition and identification performance. The tasks, from energy detection through to detailed object discrimination, are viewed as a continuum with the resolution required for the detail of the target object decreasing in absolute size as the task shifts from detection to identification. Our earliest study\(^7\) was to measure size thresholds for recognition and identification of a set of images of tanks, trucks, buses and buildings. The observer's task was firstly recognition and then identification.

The rest of this paper is a review of issues to do with applying alphabetic characters. Experimental thresholds for recognising numbers and letters were measured (Figure 2). Prior experiments by Bowler\(^8\) had found that the alphabet had threshold sizes with a distribution which centred around the Landolt C.

The Landolt C is characterised by a gap one fifth of the height and a stroke width one fifth of the height. The model uses a median fractional perimeter value of 0.2 for modelling recognition of alphabetic characters. Limiting the relationship to just a fraction of the perimeter fails with sampled stimuli as the target object can be large and the detailed information may be lost. Studies of sampling on recognition\(^7\) reveal a linear relationship between required target size for recognition and sample size. The number of samples over the linear range was a constant for a given task.

The model therefore required a second component that limited recognition probabilities for conditions with large stimuli at high luminance contrast and with very coarse sampling. For a given task level we have included a function that models a required number of elements of the target object that must be resolvable before recognition can be achieved. If the sampling of the displayed information is not the limiting parameter then the limit is imposed by the optical spread function and receptor sampling of the eye. Data from legibility experiments from Ginsburg\(^5\), Tomoszek\(^6\) and Bowler\(^8\) (Figures 2 and 3) are shown, which provide evidence of the limiting sizes below which observers are unable to discriminate characters. Figure 4 shows the data from the sampling experiment with the trend for increased target size with sample size.

Having introduced the above concepts to the model and using the standard search model we could then establish a distribution of fractional perimeter values for search tasks. Figure 5 shows the fractional perimeter distribution for laboratory measured foveal recognition and identification as shown earlier in Figure 1. Figure 6 results from an analysis of the SEARCH 2 data set and shows the distribution of fractional perimeter values which were arrived at by best fit between modelled median time and experimental median times for individual scenes and sets the required task level for modelling average performance of
The lobe will not model much more than the pure visual 6.5. Temporal variation
provides a scaling of detection with eccentricity, whether the process is as sensitive with small high contrast
aperture in combination with the optical effects on the MTF to across the retina. A similar problem applies to target size and
receptor density and the receptive field or cone sampling whether signal to noise threshold criteria are maintained equally
peripheral visual performance by using a simple relationship to component but we have very limited evidence to suggest
provides data related to acuity studies. Modelling can represent pronounced. The decision process may have an attentional
of 0.25 seconds. The scaling of thresholds for small targets between forced and free choice on an observer's threshold is
0
Taylor
thresholds. Detection has been measured peripherally by There is a need to consider experimental methods and their
consistency between cone receptor density and small target
The threshold contrast trends for peripheral viewing show a consistency between cone receptor density and small target
thresholds. Detection has been measured peripherally by Taylor for single glimpse viewing using an exposure duration
of 0.25 seconds. The scaling of thresholds for small targets
provides data related to acuity studies. Modelling can represent peripheral visual performance by using a simple relationship to
receptor density and the receptive field or cone sampling
aperture in combination with the optical effects on the MTF to provide a scaling of detection with eccentricity.

6. ISSUES IN MODELLING SEARCH FROM LOCAL VISUAL PERFORMANCE

The lobe will not model much more than the pure visual components of the search process although we have found that
next fixation can be calculated with a good accuracy for moderate clutter scenes where glimpse to glimpse patterns are
governed by peripheral signal strength for a single confusable item. Where strategic choices are needed then the lobe model is
not sufficient. We know that using technical training manuals as guides to how observers search in order to develop strategy in
models is not ideal, as we have found that some military observers will recite the manual when we ask how they search,
but when measured using eye movement analysis the patterns used were different. Foreground, midground, far ground priority
did not feature strongly in their eye movement patterns.

6.1. Closing Range
For some tasks simple modelling is very appropriate. Our early studies into the detection of closing range aircraft found that the
rate of target growth was the dominant component. Where the scene is relatively uncluttered and the task is the detection of an
energy difference then the growth of the intensity of signal for a rapidly closing aircraft is dominated by the increase in signal
intensity, and simple lobe models are very effective.

6.2. Probability or signal space
Does it make sense to model search in probability space? We have had some success in predicting the next fixation by
modelling the level of signal above threshold in combination with an aspect ratio contrast and size difference contrast. These
were for contrived scenes where objects were stretched or brightened to make signal strength in one dimension noticeably
different, and a high success rate was achieved in predicting the next fixation.

6.3. Search time statistics
We assume the observer is dedicated to the search task and that
time sharing is not a factor. There are many ways of analysing
search data but the method chosen must be consistent with the
modelling approach to be a fair test. The statistics of search
times are not normally distributed unless transformed into log or
alternative space, therefore mean times and 50% cumulative
probability times provide different values. Reaction time,
particularly in an easy task, needs to be removed from the
analysis unless it is included as part of the modelling
consideration.

6.4. Signal to noise
There is a need to consider experimental methods and their
effects on visual performance. For example the difference
between forced and free choice on an observer's threshold is
pronounced. The decision process may have an attentional
component but we have very limited evidence to suggest
whether signal to noise threshold criteria are maintained equally
across the retina. A similar problem applies to target size and
whether the process is as sensitive with small high contrast
targets compared to large low contrast targets. Factors of 2.5 to
6 above Blackwell's 1946 threshold have been suggested.
Typical minimum contrast thresholds in a practical task are
rarely less than 1%.

6.5. Temporal variation
Every threshold measurement is subject to a variance and
specific studies have examined the change of threshold with
time. The inclusion of these variances is part of the search
process. We allow for a small variance between glimpses but a
larger one over minutes.

6.6. Area under visual lobe to search
The use of plain fields of view has provided control over target
signature, and results from experiments have generated both
lobes and search data, from which a basic search model has been
developed. We can establish a high correlation between area
under the lobe and the rate of accumulation. We have also
needed to allow for within and between observer variability to
represent the search process.

6.7. Foveal Scotoma in a search task
The loss of equal volumes of lobe probabilities during a search
task do not lead to equal search performance if the fovea is
involved. This may be due to a foveal dependence of
accommodation mechanisms, or due to fixation times taking
longer due to loss of high acuity vision. The relationship of the
lobe to search is not simple because the loss of the fovea
considerably limits the speed of target acquisition. Measured eye
movements do produce an increase in the mean fixation duration
but not significant enough to account for the change of search
probability as shown in Figure 8. The measures of saccadic
amplitude are not significantly different under the influence of a
simulated scotoma.
6.1. Colour

Figure above table shows non targets were generally of longer purposes. The following conclusions were drawn by Bell 14 from falsely designated as targets provided useful information on the detection of targets.

An additional analysis of the confusable objects that were varying target density and recognition of targets taken from 515 scenes containing targets.

9 degrees. Figure 9 shows the resultant average visual lobe for probability of finding the first target was significantly higher. Centrally. Targets appeared in the scene at eccentricities of up to the final probability of finding all targets decreased although the to the observer for 0.33 seconds while the observer fixated of 0.75, 4. 8 and 20 vehicles per square kilometre showed that vision using military vehicles as targets were conducted using Experiments to establish discrimination capability of peripheral object and so were of greater signal strength.

An analysis of the objects in structured terrain involves discrimination of the vehicles from estimate of duration of exposure on a retinal location, will obey position.

Calculate target signal strength parameters related to glimpse from performance lobe for carrying out search. Ideally we could estimate foveally, recognition near periphery and detection important and the loss of efficiency at tracking with increasing

The useful visual lobe in a search task cannot be easily measured and may not exist. Fixation cuesing may be identification foveally, recognition near periphery and detection in far periphery. The modelled lobe is in concept the average performance lobe for carrying out search. Ideally we could calculate target signal strength parameters related to glimpse position. Measurement of recognition lobes for military targets in structured terrain involves discrimination of the vehicles from equivalent clutter. An analysis of the objects falsely designated as targets showed that they had near constant aspect ratio, larger perimeter length and higher contrast than the average target object and so were of greater signal strength.

Experiments to establish discrimination capability of peripheral vision using military vehicles as targets were conducted using images captured by video camera from a terrain model board. The model was 300:1 scale and contained villages and rural scenery. Static images, some containing targets were displayed to the observer for 0.33 seconds while the observer fixated centrally. Targets appeared in the scene at eccentricities of up to 9 degrees. Figure 9 shows the resultant average visual lobe for recognition of targets taken from 515 scenes containing targets.

An additional analysis of the confusable objects that were falsely designated as targets provided useful information on the likely cues that observers were using in designating an object as a target. Table 1 is a list of the aspect ratio, perimeter length and luminance contrast of the objects. The table includes both the non-targets and the targets.

Sizes and contrasts of the non-target objects meant they were more detectable than the targets contained in the same scenes. The above table shows non targets were generally of longer perimeter length and greater luminance contrast than the targets, but with smaller aspect ratios.

6.8. Cluttered scenes require recognition lobes

The useful visual lobe in a search task cannot be easily measured and may not exist. Fixation cuesing may be identification foveally, recognition near periphery and detection in far periphery. The modelled lobe is in concept the average performance lobe for carrying out search. Ideally we could calculate target signal strength parameters related to glimpse position. Measurement of recognition lobes for military targets in structured terrain involves discrimination of the vehicles from equivalent clutter. An analysis of the objects falsely designated as targets showed that they had near constant aspect ratio, larger perimeter length and higher contrast than the average target object and so were of greater signal strength.

Experiments to establish discrimination capability of peripheral vision using military vehicles as targets were conducted using images captured by video camera from a terrain model board. The model was 300:1 scale and contained villages and rural scenery. Static images, some containing targets were displayed to the observer for 0.33 seconds while the observer fixated centrally. Targets appeared in the scene at eccentricities of up to 9 degrees. Figure 9 shows the resultant average visual lobe for recognition of targets taken from 515 scenes containing targets.

An additional analysis of the confusable objects that were falsely designated as targets provided useful information on the likely cues that observers were using in designating an object as a target. Table 1 is a list of the aspect ratio, perimeter length and luminance contrast of the objects. The table includes both the non-targets and the targets.

Sizes and contrasts of the non-target objects meant they were more detectable than the targets contained in the same scenes. The above table shows non targets were generally of longer perimeter length and greater luminance contrast than the targets, but with smaller aspect ratios.

6.9. Glimpse time

Do we agree that the average glimpse, which consists of fixation and saccade has a duration of 0.33 seconds. We know the process is saccadic. Several data sources support the average glimpse time of 0.33 seconds. The glimpse patterns of individuals are different and a random glimpse statistic becomes indistinguishable from the glimpse fixations of multiple observers.

6.10. Colour

We have not successfully developed a colour model which directly uses the cone-integrated spectral signals in a manner that can be related to proposed theoretical processes in the visual system. Our most recent attempt at implementing the model of DeValois and DeValois 11 was unsuccessful at modelling threshold and suprathreshold search processes. We could not model a visual signal level that was adequate across a variety of luminance conditions. Our own tests on the combination of colour and luminance data have led us to use the CIE u'v' space for calculation of colour difference. Our calculation of colour contrast is

\[ C_c = \sqrt{(u'_T - u'_B)^2 + (v'_T - v'_B)^2} \]

where \( C_c \) is the colour contrast, and \( u', v' \) are CIE co-ordinates for the Target (T) and background (B). This is substituted into the luminance equation and a constant included for relating colour signal to measured performance.

To account for target motion effects on the highest acuity channel, performance is modelled as a degradation with increasing velocity. The ability to track a moving target is important and the loss of efficiency at tracking with increasing velocity provides our modelling approach. The modelling is based on the change of threshold acuity with velocity using data from Ludvigh and Miller 15, and Barber 16. A function representing efficiency of foveal fixation and therefore an estimate of duration of exposure on a retinal location, will obey Bloch's law. No effect is apparent below 0.125 degs/sec which is the average eye motion due to tremor and drift components of eye movement. Degradation is not large at velocities below 30 degs/sec where smooth eye tracking is achieved.

6.11. Target number

A study of multiple targets in a scene showed that overall detection probability decreased with a large number of targets. Target numbers of 1, 6, 12, and 30 per scene and with densities of 0.75, 4, 8 and 20 vehicles per square kilometre showed that the final probability of finding all targets decreased although the probability of finding the first target was significantly higher.

Figure 11 shows target acquisition probability against time for varying target density and Figure 12 for the progressive detection of targets.

6.12. Eye movements in search studies with plain Fields of View (FoV).

Eye movements of observers show a full coverage of field of view including the sky which we believe was for orientation purposes. The following conclusions were drawn by Bell 14 from eye movements studies during a search task.

1. Observers tended to neglect some areas of the FoV, most consistently the 1 degree at the edge.
2. The number of wasted glimpses decreases with increasing field of view size. (16.8%) in a 5 degree FoV, 12.2 in 10 degree FoV, and 8.2% in a 20 degree FoV) most were within 1 degree of the FoV edge.

3. The shape of FoV had no influence on fixation distributions.

4. Cross wires in the FoV had the effect of increasing the number of wasted fixations and altering the strategies that observers used to search.

5. Search strategies can be modelled as random although an individual directs his search pattern in an orderly manner.

As the size of the field of view increased observers tended to organise their search strategies in a more orderly manner. There was a consistent reduction in cumulative probability with time for increasing size FoVs. There is a considerable variation within performance of observers even in plain fields of view. Interfixation distance decreased significantly from a length of 3.3 degs for a 20 deg FoV, to 2.3 for 10 deg FoV, and 1.6 degs for a 5 deg FoV. These experiments involved the introduction of the target at an unknown time interval to the observer. The both reaching approximate answers. No vision model can support of ORACLE modelling, BAe Report JS12603, 1993.


The visual lobe is an essential building block for the search process and to use it requires the capability to model foveal and peripheral vision. We can show that next fixation is predictable in some scenes from a series of target related measures. We must also consider the observer’s adaptation state which requires knowing about the status of the observer before search commences; such data we have found in the past to be scarce. Therefore dependent on the task set to the modeller there is a choice of statistical versus image based modelling. Statistical models provide general performance and cannot be applied to anything other than a statistical selection of scenes. Use of a single real scene is an inadequate approach for camouflage assessment and a multitude of background types and target ranges and atmospheric conditions must be considered. With enough scenes specific image analysis will contain the variance which we have attempted to put into the statistical model. There is a place for both statistical modelling and image modelling, both rely on greater understanding of human visual processes and the choice is probably dependent on cost effectiveness with both reaching approximate answers. No vision model can provide the perfect answer as all our models are built to be an approximation of the truth.

Can we carry over any of the properties of an empty field search model to a structured scene?

1. Glimpse frequency is consistent.

2. There is randomness in a population of observers in a plain field and similar effects did occur in structured scenes, although this is less likely to be a major influence on search time.

6.13. Search slewing influences on fixation
Holman (1985) showed that observers put under a time constraint achieved a higher glimpse rate. If the field of view is slewing under the observer’s control there is a concentration of fixations on the centre of the display. If the sensor is moving under independent control then fixations were towards the leading edge, with confirmatory saccades tracking back to cross check confusables objects in the scene.

7. CONCLUSIONS
We have not yet explored sufficiently the cognitive components of search and target acquisition. Search tasks in a military scenario can involve vigilance, fear, stress, expectation, strategy, pre cueing, etc. It is easier to model those data which are simpler to measure as we can define them accurately. We need to measure our data test sets comprehensively so that as we progress from the simple to the complex cognitive task, we can test the intermediate stages. All laboratory studies by their very nature limit the variables to a fixed number, and impose control beyond a level experienced in natural practice. The use of models derived from such behaviour must be suspect for the practical environment, hence the necessity for a degree of field validation. Our experience of terrain model simulation is that short range targets are often missed. We have attributed this to observer expectation. The same seems apparent in field trials but whether it happens operationally we do not know. In some of the studies where short range targets were not consistently detected, eye movement studies showed that observers fixated on the target several times before eventually responding to a positive detection. The fixations prior to designation were substantially longer than others. Observers briefed to expect short range targets were more likely to detect them.


Figure 1. Detection Recognition and Identification probabilities against target size.

Figure 2. Alphanumeric character recognition experimental threshold data taken from Bowler\(^2\) and Tomoszek\(^3\).
Figure 3. Ginsburg\(^9\) character recognition data for variable size and contrast.

Figure 4. Recognition threshold sizes dependent on sample size.
Figure 5. The distribution of fractional perimeter values required to model the threshold sizes for BAE recognition and identification experiments.

Figure 6. Fractional perimeter values for modelling SEARCH 2 and a UK field study.
Figure 7. Peripheral contrast thresholds for varying size discs (Taylor, 1961).  

Figure 8. The effect of simulated 2 degree scotoma on search.
Figure 9. Peripheral recognition of targets in a structured scene for a single glimpse.

Figure 10. Motion Detection data from Barber and acuity with motion from Ludvigh and Miller Target Motion Data.
Figure 11. Search performance for varying target density (probability of finding all targets).

Figure 12. Search performance for varying target density.