The roles of cortical image separation and size in active visual search performance

Brad C. Motter
Veterans Affairs Medical Center and SUNY Upstate Medical University, Syracuse, NY, USA

Diglio A. Simoni
Veterans Affairs Medical Center and Syracuse University, Syracuse, NY, USA

Our previous research examined the effects of target eccentricity and global stimulus density on target detection during active visual search in monkey. Here, eye movement data collected from three human subjects on a standard single-color Ts and Ls task with varying set sizes were used to analyze the probability of target detection as a function of local stimulus density. Search performance was found to exhibit a systematic dependence on local stimulus density around the target and as a function of target eccentricity when density is calculated with respect to cortical space, in accordance with a model of the retinocortical geometrical transformation of image data onto the surface of V1. Density as measured by nearest neighbor separation and target image size as calculated from target eccentricity were found to contribute independently to search performance when measured with respect to cortical space but not with standard visual space. Density relationships to performance did not differ when target and nearest neighbor were on opposite sides of the vertical meridian, underscoring the hypothesis that such interactions were occurring within higher visual areas. The cortical separation of items appears to be the major determinant of array set size effects in active visual search.

Keywords: visual search, cortical magnification, V1, crowding, eye movement, object recognition, object density


Introduction

Visual search is accomplished through a cycle of fixations and visual scene analysis interrupted by saccades. A saccade produces a rapid shift of gaze, redirecting the fovea onto a new point in the visual scene. As the visual system reacquires the image data, the visual scene is remapped onto primary visual cortex governed by the physical limits imposed by the retinal photoreceptor layout and the cortical magnification factor. These limits constrain the representational power of cortex and, therefore, also constrain the computational capabilities of the visual system. Given that the results of these visual computations lead to the behaviors that we observe, it is important to understand how these and other physical constraints affect performance. Previous search studies have generated many insights into this question using various visual search tasks (Cameron, Tai, Eckstein, & Carrasco, 2004; Duncan & Humphreys, 1989; Eckstein, Thomas, Palmer, & Shimozaki, 2000; Palmer, Verghese, & Pavel, 2000; Strasburger, Harvey, & Rentschler, 1991; Treisman, 1988; Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989; Wolfe, O’Neill, & Bennet, 1998). Although some of these studies permitted eye movements, the focus of the studies has been to characterize processes that occur during a fixation. A second group of more recent studies have specifically addressed what processes characterize visual search when the eye is allowed to move freely about the image (Findlay, Brown, & Gilchrist, 2001; Findlay & Gilchrist, 1998; Geisler & Chou, 1995; Hooge & Erkelenz, 1998; Maioli, Benaglio, Siri, Sosta, & Cappa, 2001; Motter & Belky, 1998a, 1998b; Najemnik & Geisler, 2005; Shen, Reingold, Pomplun, & Williams, 2003; Zelinsky, Rao, Hayhoe, & Ballard, 1997).

Indirect measures of the functional consequences of the retinocortical mapping of visual space on search performance have focused on the use of M-scaling manipulations of stimulus size to account for a variety of observations (Strasburger, Rentschler, & Harvey, 1994). In particular, studies using central fixation targets that analyzed the effects of target location on search performance have shown that reaction times and error rates increase with target eccentricity and that the extent of this eccentricity effect is more pronounced at more peripheral target locations for larger set sizes (Carrasco, Ever, Chang, & Katz, 1995). The reduction in visual acuity because of the decrease of spatial resolution, coupled with lateral inhibition factors due to the increase in receptive field size with increasing eccentricity, has been cited to account for these results (Carrasco & Frieder, 1997; Levi, Klein, & Aitsebaomo, 1985; Toet & Levi, 1992). In this study, we attempt to separate those factors by representing the visual stimuli on the surface of a model of V1 cortex and measuring both stimulus separations and stimulus size.

With each new fixation, the amount of cortical machinery associated with the various stimuli, as well as the linear
distances between their cortical representations, varies as the image information is mapped in a nonlinear fashion onto primary visual cortex to reflect the new foveal direction in the visual scene. Previous studies of active visual search in the monkey have shown that target detection probability is invariant with respect to set size after applying a proper scaling for stimulus density. This has been shown by appropriately normalizing the probability functions for different set sizes using a metric constructed from an average measure of local stimulus density obtained from each set size (Motter & Belky, 1998b). This observation, together with those obtained from lateral masking or crowding studies that describe the severe degradation in identification performance that results from the introduction of flanking distractors around a given target (Bouma, 1970; Pelli, Palomares, & Majaj, 2004; Toet & Levi, 1992), suggests the importance of local stimulus density upon target detection.

Indeed, the results reported here favor the hypothesis that cortical image density in the immediate vicinity of the target is a major constraint, limiting target detection during active visual search in humans. While corroborating previous similar findings in the macaque, our results provide further support for the idea that a passive constraint of cortical magnification in combination with an active selection of the fixation sites work together to set the spatial framework for target detection during active visual search.

**Methods**

**Observers**

One female and two male university students with normal or corrected-to-normal vision and without any observed oculomotor abnormalities participated in this study. S.M. and M.M. were naive with respect to the purpose of the experiment; D.S. is one of the authors. The subjects gave informed written consent. The study was conducted under protocols approved by the Institutional Review Board.

**Stimuli**

Stimuli were randomly rotated (in 60 deg increments); Ts or Ls formed from appropriate perpendicular arrangements of two small (1.25 × 0.25 deg), high-contrast bars. Ts and Ls were chosen because in our previous work with monkeys, Ts and Ls search never became automatic or “parallelized” despite hundreds of thousands of trials. Search arrays of 6, 12, 24, or 48 identically colored stimuli included a single target and a number of distractors of the opposite form chosen randomly from the six orientations available. Random permutation sequences controlled the composition of search arrays, selecting from 1 of 12 possible target types (two letters, six orientations) and four set sizes. All stimuli in a given array have the same color; the color was randomly chosen to be either red or green for every trial. Stimuli were presented binocularly on a gray (18 cd/m²) background at random locations on a 24.5 × 36.5 deg video display field using a minimal center-to-center separation of 2 deg. This procedure produced displays with variable local stimulus density but prevented confounds due to stimulus overlap. Custom software running on a Windows-based PC drove stimulus presentation on a 21-in. Sony display monitor (GDM-F520) at a resolution of 800 × 600 pixels (70 Hz update rate) and a viewing distance of 57 cm, which yielded a resolution of 22 pixels/deg.

**Procedure**

A trial was initiated with the presentation of a dot (0.5 deg) in the center of the screen. Subjects were required to fixate the dot and press a button to indicate their readiness to proceed. The display coordinate system was adjusted using the eye position at button press time via a drift correction procedure. If the required corrective deviation was less than 0.5 deg, the trial proceeded by withdrawal of the fixation dot and presentation of the target for 750–1,000 ms after which the target was removed and the stimulus array was presented. If the corrective deviation was greater than 0.5 deg or if the subjects failed to maintain fixation within 1.13 deg (25 pixels) of the target’s center, the trial was immediately aborted and another one was started using a new set of display parameters. Once the stimulus array appeared, the subjects were allowed to move their eyes freely about the entire display area for up to 7,000 ms in search of the target. The target was declared acquired if subjects fixated the target (within 1.13 deg) and remained near it (within 2.8 deg) for 600 ms. The trial was allowed to continue if eye position failed to stay within this spatiotemporal window. If subjects were unable to capture the target within the time allocated, the display field was blanked and the trial was stopped and excluded from analysis. Target acquisition within the final 600 ms prolonged the array presentation as required to determine the outcome of the fixation. We defined search time as the period between search array onset and the time of target acquisition and did not include the 600 ms period used to establish the behavioral response.

**Eye movement recording and calibration procedure**

A chin rest was used to minimize head motion and maintain a constant viewing distance. The position of one eye (left or right, always the same for each subject) was measured using an Eyelink I eye tracker (SMI, Inc.) under link/peripheral control with a temporal resolution of 250 Hz and a spatial resolution of 0.1 deg. An eye movement was classified as a saccade when its distance exceeded 0.1 deg; its velocity reached 30 deg/s, and its acceleration reached 8,000 deg/s². Otherwise, eye position samples were
averaged and the centroid thus obtained was regarded as
the point of fixation. A calibration/validation sequence
was executed at the beginning of every session. The
calibration step consisted of fixating a small dot presented
sequentially for 1 s at each of nine known locations that
together spanned the entire display area. A similarly
constructed validation step followed. The target presenta-
tion sequences for both calibration and validation steps
were generated at random every time. The average total
setup time, including head mount positioning and at least
one calibration/validation sequence, was 7.5 min.

Data

Many of the measurements to be made were of distances
between fixation and stimuli in randomly arranged arrays.
The position fixated is not controlled, and instances of
particular distances accumulate unevenly. Our approach
was to gather a large set of data and then require a minimum
of 50 observations per data point for analysis and
presentation.

We analyzed midtrial fixations only: All initial fixations
(which were always at the center of the screen) and final
fixations (the ones used to determine target acquisition)
were excluded. Subjects were instructed to be fast and
accurate. A total of about 6,000 trials per subject, yielding
about 29,000 fixations per subject, were collected. Subjects
completed between two and five experimental sessions per
day (240 trials/session) within a 30-day period. Four
percent of the trials or 7% of the fixations (a bias for long
trials) were excluded due to eye blinks. The trial exclusions
were not systematically biased with regard to stimulus or
analysis conditions, and given the large number of trials
collected, no attempt was made to replace them. In
addition, the Eyelink head mount could shift on the
subject’s head (especially during long sessions). Because
target acquisition used a tightly prescribed spatiotemporal
window, noisy samples generated fluctuations that inter-
fered with the correct execution of the algorithm that
determined target acquisition, even when subjects claimed
to be fixating the target. This resulted in trial timeouts or in
data records containing patterns of very long fixations
(>350 ms) and short saccades (<1 deg) near the target
location. Sessions (n = 3) were truncated when this pattern
occurred.

Density measures and retinocortical
transformation

We use and discuss several measures of stimulus density.
The overall spatial density of stimuli in the display area can
be described simply as the number of stimuli in the display
divided by the area of the display. A rough estimate of the
distance between stimuli can be obtained by taking the
square root of the display area divided by the number of
items (although this estimate ignores the display contour).

We can achieve a somewhat better metric by measuring the
average nearest neighbor distance (ANND) directly and
using this value as a metric of item separation and global
stimulus density in visual space (Motter & Belky, 1998b).
In rectangular displays, this measure is less than, but
proportionate to, the rougher estimate above. Here, we
wanted to obtain a measure of density in a local area
around a given stimulus. Local density can be charac-
terized by the number of stimuli within a radius r divided
by the area contained within that radius. Previously, we
determined a weighted density measure that accounted for
the proximity of each of three stimuli flanking the central
one. Here, to simplify this metric, we used the distance
from the reference object to its nearest neighboring item
as an estimate of the local density around the reference
stimulus. This distance was measured in degrees of visual
angle when local stimulus density was calculated with
respect to visual space and in millimeters of cortex when
calculated with respect to cortical space.

Given an estimate of the cortical magnification factor as a
function of eccentricity $M(w)$, we measured the separations
between stimulus representations in primary visual cortex
using a 3D model of the surface of primary visual cortex,
area V1. We associate each point in visual space $v(\theta, w)$
described by azimuth $\theta$ and eccentricity $w$ with a
corresponding point in cortical space $c(r, z, \phi)$ described
by the radius $r$ of the 3D model surface at that
eccentricity, the distance $z$ along the axis of rotation (both
given in millimeters), and a rotational angle $\phi$ (identical
to the azimuthal angle $\theta$), according to the following set of
equations (Rovamo & Virsu, 1984):

$$r = M(w) \sin w$$

$$z = \int_0^w \left[ M(w)^2 - (dr/dw)^2 \right]^{0.5} dw$$

$$\phi = \theta.$$

We have used an estimate of the average human cortical
magnification factor from a report by Duncan and
Boynton (2003), expressed there as $1/M(w) = 0.065w +
0.054$ (deg/mm). Using the above equations, we con-
structed 3D model views of the surface of primary visual
cortex (see Figure 1). These views portray the surface of
primary visual cortex in an accurate manner, assuming
only local isotropy and the magnification factor. The same
surface is found by implementing the Daniel and
Whitteridge (1961) equations, and a similar surface is
produced by using the magnification equations used by
them, differing in size naturally because their magnifica-
tion factors were for an average nonhuman primate. There
are several notable departures from the orthographic
projections depicted in the study of Daniel and
Whitteridge, particularly in that the peripheral extents of
the meridians do not meet at a point. We address those
issues elsewhere. For the purposes of display, we have
seamlessly joined the two hemifield representations into a single 3D view. Onto this surface, we can project the images of stimuli to illustrate the differences in size and spacing resulting from the cortical magnification. This illustration simply sets the basic constraints on the areal representations of the stimuli. One should bear in mind that there is some debate as to whether different meridia have somewhat different magnification factors (Blasdel & Campbell, 2001; Schiessl & McLoughlin, 2003). Those factors as well as variation in the magnification factor (every individual has a unique magnification factor) produce locally minor differences in the overall surface picture.

Results

Basic search performance

Figure 2 shows the basic search performance averaged over the three subjects. Total search time can be characterized by a quasilinear monotonically increasing function of array set size, as expected from previous studies of conjunction search (Treisman & Gelade, 1980; Wolfe, 1998). In active search, the number of fixations executed during the target discovery process can also describe search performance (Maioli et al., 2001; Motter & Belky, 1998b). Both measures of performance exhibit the same functional form up to some constant scaling factor. These results are in agreement with those found under similar experimental conditions in the macaque (Motter & Belky, 1998b) and in humans (Maioli et al., 2001). That the search time and fixation count functions are parallel with each other reflects not only that fixation duration is approximately invariant with respect to changes in set size but also that most of the time available for target detection occurs during the periods of fixation; saccades account for only 20% of the search time. Neither total search time nor fixation count functions support a simple linear model of item-by-item search: The slope of the fixation count function directly implies that more than one stimulus is processed in parallel during each fixation. This observation leads to an analysis of just where targets are when they are detected during active search. They were found to be clustered in an area around the point of fixation, the conspicuity zone, within which stimulus information can be extracted effectively (Motter & Belky,
This zone can be described mathematically as a sensitivity function. With a quantitative characterization of focal attention thus constructed, one can further examine its properties in response to systematic manipulations of stimulus conditions, particularly to changes in stimulus density.

**Estimating the conspicuity zone**

We estimated the size of the conspicuity zone in a manner similar to that in Motter and Belky (1998a). For every midtrial fixation, we measured the eccentricity of the target from the point of fixation and noted whether the ensuing saccade captured the target or not. Using these data, we calculated the probability of detecting a target as a function of its distance from the fixation point. The results are shown in Figure 3. Consistent with the previous results in monkeys, the zone of focal attention during active visual search in humans can be described in terms of the probability of target detection as a smooth monotonically decreasing function of target eccentricity. The four panels in Figure 3, one for each array set size, illustrate the remarkable similarity that exists between the data gathered from all three subjects.

These results show a steeper rise in probability with smaller target eccentricities than did the monkey data. We were concerned that this might reflect less accurate targeting by the humans, followed by a smaller corrective saccade, thus artificially inflating the near-target probability of detection measure. In the human data, 75% of the eye
movements placed fixation within 2 deg of an item during midtrial fixations. For monkeys, about 80% of eye movements ended within 1 deg of an item (Motter & Belky, 1998a). In a control experiment (data not shown), subjects were asked to fixate a single target drawn from the same stimulus set and presented sequentially for 1 s at each of 25 known locations. The results of this control experiment indicate that 80% of the targeting fixations landed within 1 deg of the target regardless of the length of the previous saccade. This demonstrates that our subjects were capable of higher accuracy in the targeting of stimuli in the displays. This observation suggests that the results shown in Figure 3 may portray a slight difference in search strategies between monkeys and humans; for example, they may sometimes land saccades between stimuli, a less efficient strategy when the goal is to fixate the target. Alternatively, humans, despite the many trials involved in these experiments, nevertheless had far less experience in this task than the monkeys due to the extensive training required for the monkeys. The eye positioning differences may reflect this difference in experience.

The role of stimulus density in controlling the conspicuity zone

The suggestion that stimulus density plays an important role in limiting target detection arises initially from the observation that, for any given eccentricity, the probability of target detection decreases with increasing set size, as has been previously observed under centrally maintained fixation search conditions (Carrasco et al., 1995; Carrasco & Frieder, 1997). This can be seen clearly in Figure 4A, which depicts the sensitivity curves for the different set sizes averaged over the three subjects. These results suggest that overall stimulus density as defined by set size plays a major role in controlling the size of the zone surrounding the point of fixation over which targets can be detected with high probability. If correct, then, after normalization for stimulus density, target detection probability should exhibit invariance with respect to set size. The data were normalized for the separation between stimuli by expressing interstimulus distances in ANND units (see Methods section). Empirical ANND values for each set size were calculated from the entire data set (see inset in Figure 4B) and used to normalize the target detection probability curves for array set size by expressing target eccentricity in ANND units. For example, an array set size of 12 items had an ANND of 5.2 deg. Targets at 10.4 deg eccentricity (inverted triangles) have a probability of detection of 0.25. When spatially scaled into ANND units, that same detection probability occurs at an eccentricity of 2.0 in ANND units. The sensitivity curves for the different set sizes superimpose one another after this transformation, as shown in Figure 4B, providing support for the idea that global stimulus density is a principal factor controlling target detection in the area surrounding the point of fixation. Note that the curves did not have to collapse because of this scaling; the scaling simply normalizes the overall spacing between items. Note also that scaling does not alter any actual lateral interactions between stimuli; the scaling is just a different representation of the data. The fact that the curves collapse indicates a very clear linkage between stimulus separation (density) and detection.

Notice also that the sensitivity curves tend to reach asymptotic base values with increasing eccentricity. This is most evident in the topmost curve in Figure 4A for the array set size of six items. In this case, the asymptotic value of the probability of target detection on the next saccade is approximately 0.2, which is equivalent to chance performance (one out of five) under the assumption

Figure 4. Detection sensitivity curves. (A) Probability of target detection as a function of target eccentricity from current point of fixation for all four set sizes averaged across all three subjects. The curves form an ordered set of monotonically decreasing functions, where the probability of target detection decreases with increasing set size for any given target eccentricity. The asymptotic values for larger target eccentricities approach chance performance levels for each set size. (B) When stimulus density is normalized by expressing target eccentricity in terms of ANND units (see text), the target detection curves superimpose, indicating that stimulus density plays an important role in target detection during active visual search.
that the fixation is currently examining one out of the six stimuli. Probability of detection should not fall below this chance level. The superimposition of the curves in Figure 4B should be broken once this chance level is reached, and indeed, this can be seen for the data of array set size of 6. A similar argument can be used to explain the asymptotic behavior of the sensitivity curves for larger set sizes.

**Target detection as a function of local cortical density around the target**

Although normalizing for the distances between stimuli across array set sizes collapses the curves in Figure 4B, detection remains a clear function of eccentricity. It makes little sense that the average density of the display should affect target detection. Why should the density of stimuli in locations far away from the target affect target detection performance? Instead, it seems that the local density of items surrounding the target might be a more reasonable measure. However, for a given array set size across many trials, the ANND values are the local density gradients everywhere in the display. We recognized that the eccentricity dependence and the density dependence could both be the same issue if they were addressed at the level of the cortical representation of the visual field. The M-scaling work of others is in line with this idea, although usually approached through a scaling of stimulus size rather than spacing. Indeed, our previous work in the macaque has shown that the decline in sensitivity with eccentricity is rooted in the density of items represented within a unit area of cortex (Motter & Holsapple, 2000). Our preliminary results suggested that this explanation generalized to human subjects under similar search task conditions (Simoni & Motter, 2003). Therefore, the effects due to variations in the density of the cortical representations of stimuli surrounding the target were examined. As an estimate of the local density about the target, we used the linear distance (in millimeters) along the curved surface of the cortex in the model shown in Figure 1 from the center of the cortical representation of any given target to the center of the cortical representation of its nearest neighboring item. The functions that relate this estimate to the probability of target detection are shown in Figure 5 for the different set sizes. The four panels in Figure 5, one for each subject and one for the subject average, show that target detection can be expressed as a simple function of the cortical separation between the target and its nearest neighbor.

![Figure 5](image_url)
target eccentricity. Essentially what this means is that targets are detectable when their neural representations in cortex are isolated from “cortical crowding” effects from nearby distractors.

**Target detection as a function of cortical image separation and target size**

But what about eccentricity? Nearly every previous report considers eccentricity to be the major factor in the peripheral decline of target detection, yet Figure 5 collapses across eccentricity, showing an invariance to eccentricity. This suggests that eccentricity and cortical separation (as a measure of density) are independent factors in determining detection. The 3D bar chart of Figure 6 depicts the probability of target detection as a function of stimulus density around the target and target eccentricity when both are expressed as distances on the surface of the cortical model. Data were collapsed across array set size and subjects. Detection probability decreases as eccentricity (distance from the fovea in millimeters) increases, and detection probability increases as the cortical separation between target and nearest neighboring stimulus increases. A multiple logistic regression was conducted using these factors. The regression equation determined was Logit \( P = 0.63 - 0.073 \times \text{ecc\_mm} + 0.092 \times \text{sep\_mm} \). Standard errors for the fit constants are 0.056, 0.001, and 0.002, respectively. The factor \( \text{ecc\_mm} \) is the target eccentricity in millimeters along the surface of the model, and the factor \( \text{sep\_mm} \) is the separation in millimeters between target and nearest neighbor along the surface of the model. Both factors were found to contribute (Wald statistic, \( p < .001 \)) to the logistic equation independently and with no evidence for multicollinearity (variance inflation factor <2). We assume that the eccentricity factor is related to acuity and that target size, not eccentricity per se, governs target detection. Therefore, in Figure 6, the data are replotted in terms of target size in square millimeters rather than in terms of eccentricity. The estimate is based on the area of a circle enclosing the target and having a diameter of 1.42 deg. The calculation is based on the radius length of the circle projected along an equal eccentricity line at the eccentricity of the target and does not correct for the distortion along a meridian. Those differences are relatively minor once beyond a few degrees from the fovea as can be seen in the letter “T” depictions in Figure 1. As target size increases above 2–3 mm\(^2\), target detection rises rapidly as long as the separation from the nearest interfering stimulus is also relatively large. On the other hand, at the largest target sizes—targets located close to the fovea—the separation distance does become less relevant to target detection. As depicted in Figure 6, eccentricity, in terms of target size, and stimulus density, in terms of nearest neighbor distance, are equally important in determining target detection.

**Bouma’s observations and cortical mapping**

Bouma (1970) noted that crowding effects were stronger for flanker positions with greater eccentricity relative to the target than for flanker positions between the target and point of fixation. This relationship (also noted by others, e.g., Toet & Levi, 1992) is exactly what is predicted from the cortical mapping of visual space where, because of the magnification factor, angular distances in visual space are compressed in their placements on the cortical surface. Thus, in regard to Bouma’s observation, although the angular separations for near and far flankers are the same in visual space, the far flanker is actually closer to the target than the near flanker after mapping to cortical space.

Bouma (1970) also provided a crowding rule of thumb, noting that crowding occurs when the separation between a target and a flanking object is within one half of the eccentricity of the target. Bouma recognized that the interactions were not symmetric and referred to the area of
interaction as more egg shaped. The observations in Figure 5 suggest that Bouma’s rule might result from the cortical magnification factors’ remapping of spatial relationships. To examine this issue, we determined the cortical separation of two objects obeying the 1/2 eccentricity rule based on the model shown in Figure 1. One stimulus, the “target,” was moved along a meridian from near fovea to 40 deg in eccentricity. A second stimulus, the “flanker,” was positioned at a distance of one half of the eccentricity of the target at each of three positions, nearer to the fovea or farther than the target along the meridian and to one side along the equal eccentricity arc. The results are shown in Figure 7. From about 10 deg of eccentricity to 40 deg, each of these three target–flanker separations is represented by a constant cortical separation. The differences in the separations for different flankers represent the asymmetry of the flanker effects as noted above, varying from a separation of 10 mm, representing the 1/2 eccentricity distance on the fovea side of the target, to a separation of 6 mm, representing the 1/2 eccentricity distance on the peripheral side of the target. The rule itself clearly requires that the lines converge as the target approaches the fovea, and indeed, the constant cortical spacing is not maintained within 10 deg of the fovea. Most of Bouma’s observations were made within 10 deg of the fovea, and most of our target detections were also made in this range. Thus, within the range of particular interest, the 1/2 eccentricity rule does not map out a constant area in cortex. We
suggest in the discussion that a cortical-based rule might actually be the basis of Bouma’s observation.

Target detection is unaffected by stimulus density around the fixation point

To verify the results above, we examined whether search is affected by local stimulus density variations around the fixation point. As expected, we found no relationship between this density measure and target detection. Using the same intuition that guided the use of the cortical interstimulus distance as an estimate of local cortical stimulus density, we then used the distance from any given fixation to the center of its second nearest neighboring item in degrees of visual angle as an estimate of its local stimulus density. This estimate assumes that the current fixation is positioned near a stimulus; therefore, the second nearest stimulus provides a better estimate of local stimulus density. The functions that relate the probability of target detection to changes in stimulus density local to the fixation point for the different set sizes are shown in Figure 8. The four panels in Figure 8, one for each subject and one for the subject average, provide a strong indication that target detection does not depend in any systematic way on the local stimulus density around the point of fixation. There is no decrease in detection probability at higher densities around fixation (smaller second nearest neighbor distances); in fact, there is a slight decrease with increasing second nearest neighbor distances. However, the slight decrease is a result of the sampling. If the second nearest neighbor is at sequentially further eccentric locations, then, by definition, the target is also at sequentially further eccentric locations; thus, a sampling bias for target locations at increasing distances underlies the slight decline in target detection. Likewise, the fact that detection appears to be based on array set size is simply a reflection that stimulus density around the target determines detection probability, which the results previously laid out.

These results imply that across a wide range of stimulus density variations surrounding the fixation point, the probability of target detection is approximately the same. The same results fall out of a consideration of the cortical scaling about the fixation point because with regard to distances calculated from the fixation point itself, those distances are simply eccentricity scaled versions of the data in Figure 8.

The correct framework for target detection is cortical density and not spatial density around the target

To demonstrate that it is cortical density rather than simple visual space density, we examined whether search
performance is affected by local stimulus density variations around the target calculated with respect to visual angle. The estimate of the local stimulus density around the target on the display screen was measured as the distance from the center of the target to the center of its nearest neighbor item in degrees of visual angle. The resulting functions are shown in Figure 9. The four panels in Figure 9, one for each subject and one for the subject average, show that target detection is not a function of the spatial density of items surrounding the target in normal visual space. If density in normal space were a factor, we would expect to see a low probability of detection for small separations between target and nearest distractor that gradually increases to a higher probability of detection for larger separations. This is not the case; for each array set size, the probability of detection is essentially flat and thus independent of the separation between target and its nearest neighbor. The slight decrease in detection probability in Figure 9 reflects a sampling bias in a fashion somewhat similar to Figure 8. If we require the nearest neighbor to the target to be at increasing larger distances, then we are probabilistically biasing the sample for target locations toward the edges of the display. That bias also makes the target probabilistically farther from the fixation point given any random pairing. The slight decrease in detection with increasing eccentricity in Figure 9 reflects this bias.

The result pictured in Figure 9 is a very different result from the cortical density result shown in Figure 5, where detection is a clear function of separation between the target and its nearest neighbor. The difference between figures resides in the way space is mapped by cortical magnification. When cortical magnification is not factored in, then spatial density around the target averaged across all eccentricities is not related to performance as seen in Figure 9. On the other hand, the separation of array set size curves in Figure 9 remains as evidence that density is in fact a key factor in regulating target detection. The key in this case is the recognition of the correct frame of reference, cortical space rather than visual space.

**Intra- and inter-hemifield density-invariant target detection**

We have shown that the area within which a target is discovered is a function of the density of items surrounding it. In particular, search performance can be fully accounted for by simply considering the proximity of the nearest distractor to the target on an accurately scaled retinotopic transformation of visual space onto the surface of V1.

![Figure 10](image-url) Probability of target detection as a function of cortical separation between stimulus representations of the target and the target's nearest neighbor in the V1 model. Each panel shows subject averages for each of the four set sizes along with standard error bars. The data were analyzed separately according to whether the target and its nearest stimulus were in the same hemifield (Same) or in different hemifields (Diff) for each fixation. The similarity between the two conditions indicates that search performance cannot be based exclusively on neural processing occurring at the level of V1.
However, this is not a claim that the behavioral response is based entirely on the neural computations performed in V1. In fact, we can provide a quick check of this using the known anatomical segregation of visual hemifield information for primary visual cortex. If V1 itself were responsible, then a differential performance should be observed between conditions where the target and its nearest distractor are located in the same hemifield compared with conditions where the target and its nearest distractor are located in different hemifields. This analysis provides evidence to the contrary. In Figure 10, we have separated the data into two groups, corresponding to ipsilateral and contralateral hemifield divisions, and calculated the probability of target detection as a function of the cortical separation between the target and its nearest stimulus. To avoid contamination associated with stimulus overlap at the vertical meridian and small differences in eye position, we excluded target–nearest neighbor pairs where either of the two stimuli was within 1.5 deg of the vertical meridian. Each of the four panels in Figure 10, one for each array set size, show that the two curves corresponding to the two conditions are essentially the same. Although there appears to be a trend for same-side separations having greater detection probability, a repeated measures ANOVA on each array set size data set did not find significance ($p > .10$) for hemifield differences. Overall, Figure 10 is consistent with the idea that behavioral performance does not depend solely on processing at the level of V1.

**Discussion**

Target detection during search is a central issue in our understanding of natural vision. In the last few decades, many studies have probed the limits of what can be detected during maintained fixation, with increasing attention being paid to the role of attention in visual search. These studies have identified a number of important factors underlying detection performance such as array set size, shared target and distractor properties, scene segmentation, and grouping and crowding conditions, to name a few. Our previous work, wherein we used monkeys, suggested that at least during active search, the spatial area within which targets can be discovered is limited by stimulus density to an area surrounding the point of fixation (Motter & Belky, 1998b). This conclusion was based on the observation that a global normalization for stimulus density based on the ANND removed the differences in target detection probability that were observed between array set sizes. The global density normalization is shown for the human observer data. Simple visual angle measures of object density around the target do not replicate the global normalization (Figure 9). Later, we found that if we scaled stimulus distances by the cortical magnification factor, the density of objects around the target itself determined detection performance (Motter & Holsapple, 2000). Stimulus identification was constrained by the relative isolation of the target from nearby distractors in terms of the locations of their neural representations within primary visual cortex. In the discussion to follow, we try to unify these two observations.

In this study, we have investigated these issues for human visual search and we have illustrated the cortical spacing hypothesis using a 3D model of the surface of human primary visual cortex. This 3D model is produced by mapping the visual field onto a surface as specified by a cortical magnification factor and assuming azimuthal symmetry. It represents an unfolded view of primary visual cortex. This is essentially what Daniel and Whitteridge (1961) did in their original description of the cortical magnification factor in primates. Our implementation of the model uses equations published by Rovamo and Virsu (1984). The model illustrates the expansion of the central representation of the visual field. About half of the surface is devoted to the central 10 deg of the visual field. We found that mapping the stimulus display onto the 3D surface model made it easier to appreciate the impact that the cortical magnification factor makes on the topography of stimulus representation. In this study, we have used the distance measured along the surface of the model between the target and its nearest neighbor as an index of density. We have shown that detection performance is a simple function of this measure (Figure 5). Control analyses demonstrated that detection performance was unrelated to the density of stimuli about the fixation point. In summary, the results suggest that stimulus identification is a function, perhaps a threshold one, of the separation between the cortical image representations of the target and its nearest relevant distractor.

In Figures 3 and 4, we also show that the detection of the target is a function of eccentricity. How does this result fit with the cortical separation result? The geometry of the situation produces both results. Because cortical magnification decreases as a function of increasing eccentricity, a target and distractor separated by x degrees are represented in cortex by ever decreasing distances. Because detection probability decreases with decreasing separation, the probability of detection therefore decreases with increasing eccentricity. This relationship generates the curves in Figure 3, where what is being depicted is the decrease in detection probability with increasing eccentricity for the average target to distractor distance set by the number of items in the array. The density of the items around the target determines how close to the point of fixation a target must be to be identified correctly. This simple principle accounts for both the array set size and the eccentricity effects that constrain target detection during active visual search. Previously, we interpreted the decreasing detection probability curves in figures like Figures 3 and 4 as depicting the limits of the zone of focal attention, or conspicuity, within which targets could be
detected (Motter & Belky, 1998b). Indeed, this still may be the case. However, the changing size of this zone with different array set sizes does not reflect an active dynamic zoom but instead represents the passive spatial range associated with the average stimulus density of each array set size.

Stimulus identification also varies as a threshold function of the stimulus size or eccentricity. In our studies, individual stimuli were of constant size and above identification threshold to at least an eccentricity of 21 deg (furthest position tested for single items presented during maintained fixation conditions). Studies have shown that increasing stimulus size according to a cortical magnification factor can compensate for the array set size effects in the near periphery (Carrasco & Frieder, 1997). However, how to interpret those results within our framework is not entirely clear because the size but not the spacing was magnified in those studies. If size and spacing were magnified, then the increase could be just a spacing issue. If size but not spacing were magnified, then we would expect what? It could be that larger stimuli might be less crowded, but there is some question as to whether or to what extent stimulus size is that critical in crowding (Pelli et al., 2004; Strasburger et al., 1991, 1994). The use of stimuli of different sizes brings up the issue of whether center-to-center spacing is the best index of density; perhaps it is the gap between contour edges or perhaps it needs to be defined in an areal context, even including more than just the nearest distractor. We have tried several variants of both contour edge separation and areal definitions of density but found no clear difference with the simple center-to-center spacing measure. However, we may not see a difference because the size of our cortical stimulus image was small, relative to the crowding range.

Bouma (1970), Toet and Levi (1992), and others have found that crowding effects increase with eccentricity. That is consistent with our results; the effective distance between cortical stimulus representations decreases with greater eccentricity. Given the hypothesis of a threshold function based on cortical separation, one would expect stronger set size effects at large eccentricities, with all things being equal. This is also consistent with the observations reported in Carrasco et al. (1995) who found that the set size effect was stronger at far target eccentricities as compared with near eccentricities. We examined Bouma’s observation that crowding seems to take place when target and flanker distances are less than 1/2 the target eccentricity and found that it corresponds to a fairly constant separation of cortical images on the surface model of V1 that breaks down when the target is within 10 deg of the fovea. We suggest that Bouma’s observation is better captured in cortical space, and based on Figure 5, an updated version of Bouma’s rule states that crowding occurs when target and flanker are separated in cortical space by less than a threshold distance. From Figure 5, that distance based on a 50% threshold is about 14 mm.

Our results are based on considering the effects of stimulus interactions in terms of the cortical magnification factor as expressed in V1, primary visual cortex. Certainly, we do not infer that the neural locus upon which performance is based is in V1 but rather that the scale of interactions is set by the V1 cortical magnification factor. Our evidence in monkeys has shown that, at least up to the level of V4, there is no further eccentricity-dependent gain between V1 and V4 receptive fields; there is just a simple gain factor (Motter, 2003). This is significant because it is at the level of V4 processing that receptive fields first attain the areal size that individually could account for the stimulus interactions that underlie crowding phenomena. With this in mind, it is interesting that, across eccentricities, the effects of stimulus separation seem to occur at a cortical distance of about 14 mm and that this value is about the size of V4 receptive fields (after scaling for species difference) as measured on the model surface.

In this report, we have confined experimental tests and discussion to issues of density related to stimulus displays composed of Ts and Ls of a single color to limit potential factors other than density. As we have just described it, this seems a passive mechanism. However, we have described evidence from monkey studies that active, feature-selective, attentive mechanisms can adjust the effective stimulus density. Under conditions where the display can be segmented by color differences, the nearest neighbor distances and/or cortical separations that account for performance are determined by stimuli containing the target color—essentially discounting the remaining stimuli from consideration (Motter & Belky, 1998b; Motter & Holsapple, 2000). Presumably, other feature differences that support target and distractor segregation also result in a dynamic determination of effective density. It is possible that segregation simply grades smoothly into a measure of stimulus discriminability. Search studies that have examined the relationship between stimulus discriminability and the number of array items (e.g., Verghese & Nakayama, 1994) have reported a reciprocal relation between small array set sizes and more difficult stimulus discriminations in maintaining a fixed level of performance. In a general sense, as target–distractor differences decrease, the cortical separation between them must increase to maintain a given level of performance. Many visual search phenomena have been linked to the array set size effect; what we have demonstrated here is that another visual phenomena, crowding, when measured in cortical space, accounts for the array set size effect under our experimental conditions. These results support and explain many visual lobe models of active search that probabilistically account for search performance by repeated application of a spatial filter to the visual scene (Engel, 1977; Geisler & Chou, 1995; Motter & Holsapple, 2001).

A significant portion of the psychophysics of visual search has used a maintained fixation paradigm with briefly presented stimuli that both prevent effective eye
movements and essentially mimic the fixation duration of active visual search. There is little doubt that directed focal attention plays a significant role under conditions of actively maintained fixation; in many cases, it is directed at the targets or assumed to be directed at the targets. Cued covert shifts of attention have been shown to alter performance in a manner consistent with improved processing at the target’s location (Engel, 1977; Yeshurun & Carrasco, 1998). Under these conditions and with acknowledgement and control of stimulus spacing, signal detection theory has been shown to account for some limited capacity or set size effects (Cameron et al., 2004; Palmer et al., 2000). These studies provide a bridge between standard spatial discrimination models and visual search. However, the role of covert directed focal attention in active eye movement search remains uncertain, primarily because it is not clear whether there is normally sufficient time for focal attention to be directed at objects other than those fixated during rapid search.

Summary

This report corroborates and extends previous findings in the macaque, suggesting that the cortical magnification factor accounts for both set size and eccentricity effects on the probability of target detection during active visual search in humans. That the same principle can be used to account for results obtained from both monkeys and humans suggests that by using simple random displays, we have been able to identify an important low-level constraint on search performance that is shared by both species. Although we need to explore further how grouping or search strategies contribute to target detection, the simple search task used here has revealed a basic computational constraint: it is the separation of items immediately surrounding the target in cortical space rather than visual space that limits our ability to identify isolated forms at a distance under crowding conditions. In our limited stimulus conditions, the separation of items appears to be the major determinant of array set size effects. The independence of the separation distance as a function of eccentricity suggests that the anatomical sampling of V1 and convergence beyond V1 are similarly independent of eccentricity, at least with respect to crowding phenomena.

Acknowledgments

This research was supported by the Department of Veterans Affairs Medical Research Program.

Address: Veterans Affairs Medical Center, Research Service, 800 Irving Ave, Syracuse, NY 13210, USA.

References


