



Turning visual search time on its head

S.P. Arun

Centre for Neuroscience, Indian Institute of Science, Bangalore 560 012, India

ARTICLE INFO

Article history:

Received 15 January 2012

Received in revised form 6 April 2012

Available online 25 April 2012

Keywords:

Object representations

Object recognition

Perceptual similarity

Psychophysics

Visual search

ABSTRACT

Our everyday visual experience frequently involves searching for objects in clutter. Why are some searches easy and others hard? It is generally believed that the time taken to find a target increases as it becomes similar to its surrounding distractors. Here, I show that while this is qualitatively true, the exact relationship is in fact not linear. In a simple search experiment, when subjects searched for a bar differing in orientation from its distractors, search time was inversely proportional to the angular difference in orientation. Thus, rather than taking search reaction time (RT) to be a measure of target–distractor similarity, we can literally turn search time on its head (i.e. take its reciprocal $1/RT$) to obtain a measure of search *dissimilarity* that varies linearly over a large range of target–distractor differences. I show that this dissimilarity measure has the properties of a distance metric, and report two interesting insights come from this measure: First, for a large number of searches, search asymmetries are relatively rare and when they do occur, differ by a fixed distance. Second, search distances can be used to elucidate object representations that underlie search – for example, these representations are roughly invariant to three-dimensional view. Finally, search distance has a straightforward interpretation in the context of accumulator models of search, where it is proportional to the discriminative signal that is integrated to produce a response. This is consistent with recent studies that have linked this distance to neuronal discriminability in visual cortex. Thus, while search time remains the more direct measure of visual search, its reciprocal also has the potential for interesting and novel insights.

© 2012 Elsevier Ltd. Open access under [CC BY-NC-ND license](#).

1. Introduction

We frequently engage in searching for an object among clutter. Some searches, such as finding a red fruit among green leaves, are easy whereas others, such as finding a face in a crowd, are hard. What makes search difficult or easy? It is generally believed that search is hard when the target is similar to its distractors and easy otherwise (Alexander & Zelinsky, 2012; Duncan & Humphreys, 1989; Verghese, 2001; Wolfe, Cave, & Franzel, 1989). The existing literature consists of a rich list of features that determine similarity: for instance, differences along features such as brightness, orientation or color make search easy (Wolfe & Horowitz, 2004). Although search time increases with target–distractor similarity, several fundamental questions regarding its nature remain unanswered. For instance, does search time vary linearly with target–distractor similarity? Can search time be formally thought of as a distance measure? These questions are important because similarity measurements can yield insights into the underlying object representation (Basri et al., 1998; Cortese & Dyre, 1996; Desmarais & Dixon, 2005; Edelman, 1998).

To address these questions, I performed two visual search experiments, one with oriented bars and the other with natural

objects. I found that while search reaction times (RTs) qualitatively do increase with target–distractor similarity, the relationship is in fact non-linear. The inverse relationship between RT and target–distractor differences naturally suggests that the reciprocal of search reaction time ($1/RT$) can be used as a measure of *dissimilarity*. I show that this measure has several desirable properties. First, it varies linearly with the difference in orientation between the target and distractor over a large range, making it a suitable measure for situations in which the underlying feature dimensions are unknown. Second, it has the properties of a mathematical distance metric: in other words, it is always positive, approaches zero as the target becomes increasingly similar to the distractors, it satisfies the triangle inequality and is roughly symmetric. Thus, search distance can be thought of literally as distance in search space. Although symmetry in distance is not always satisfied since there are asymmetries in visual search (Wolfe, 2001), I show that asymmetries are relatively rare on a set of natural objects, occurring only 7% of the time. Interestingly, when these asymmetries do occur, the asymmetry in $1/RT$ is a small and fixed quantum of dissimilarity across all asymmetries. This regularity would never have been observed using search times alone. I then show that this dissimilarity measure can be used to yield insights into the object representations underlying visual search. Finally, search distance has a straightforward mechanistic interpretation: it is proportional

E-mail address: sparun@cns.iisc.ernet.in

to a discriminative signal that accumulates over time until it reaches a decision threshold. Thus, while search times remain the most obvious and direct measure of search performance, search distance can yield novel insights particularly in the study of similarity relations between objects.

2. Materials and methods

2.1. Observers

Experiments 1 and 2 were performed on two independent groups of six human subjects each. Subjects were aged 20–30 years, had normal or corrected-to-normal vision and were naïve to the purpose of the experiments. Subjects gave written consent to a protocol approved by the Institutional Human Ethics Committee of the Indian Institute of Science.

2.2. Apparatus

Subjects were seated approximately 50 cm from a computer monitor that was under control of custom Matlab programs based on PsychToolbox (Brainard, 1997), running on a Dell workstation.

2.3. General procedure

In both Experiments 1 and 2, subjects were instructed to perform an oddball visual search task in which they had to detect the location of an oddball item among multiple identical distractors. Subjects were given no instruction as to the nature of the target or distractor items. On each trial of the task, a fixation cross appeared for 500 ms, followed by a search display measuring $22^\circ \times 22^\circ$ in visual angle, consisting of one oddball target among 15 identical distractors. The target could appear either on the left or the right of the display and subjects had to hit a key (“M” for right, “Z” for left) to indicate the side on which the target was located. Subjects were instructed to respond as quickly and accurately as possible. To facilitate their judgments, the display also included a single red vertical line that separated the screen into two halves. Trials were repeated later in the block if subjects made an incorrect response or if they failed to respond within 5 s of display onset. Trials involving target on the left and right were interleaved randomly.

2.4. Experiment 1: Visual search for orientation

In this experiment, the target on each trial was a bar measuring $0.5^\circ \times 2.5^\circ$ in visual angle, with an orientation chosen uniformly at random from 0° to 180° . The inter-item spacing measured 6° in visual angle. The distractor items were bars with the same dimension as target, but whose orientation differed from that of the target by a fixed amount. The orientation differences took values of 7° , 10° , 13° , 15° , 17° , 20° , 25° , 30° , 45° or 60° . The target and distractor items were positioned randomly with a jitter of $\pm 0.5^\circ$ to minimize perceptual grouping. Subjects performed 20 trials of each of the 10 orientation differences.

To assess whether the ability of the subjects to discriminate a particular orientation difference (say 10°) would depend on the absolute orientations of the target and distractor (say 0 and 10 degrees versus 60 and 70 degrees), I plotted search times for each orientation difference across many target–distractor orientations as a function of the average orientation of the target and distractor items. This revealed no significant correlation (data not shown). I therefore report only the analyses assessing the dependence of search time on the difference in orientation between the target and distractors.

I fit several functions to the data consisting of search times at each orientation difference – these included exponential, linear and sigmoid functions. The sigmoid function was chosen because it is widely used in fitting psychometric functions and because it would capture the saturation observed in search distance for large orientation differences. The sigmoid function was defined as the integral of a Gaussian function, and had three parameters: the amplitude A , which specifies the maximum value the sigmoid can attain, the mean μ which specifies the value at which the sigmoid reaches half-maximum and the standard deviation σ , which controls the rate of rise. The equation for the sigmoid function was given by:

$$d(\theta) = A \int_{-\infty}^{\theta} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx,$$

where $d(\theta)$ is the distance (1/RT) value at the orientation difference θ . This function was fit using standard optimization functions in MATLAB (*lsqcurvefit*) and the best-fitting parameters for the data in Fig. 1B were $A = 1.08 \text{ s}^{-1}$, $\mu = 12.7^\circ$ and $\sigma = 13.9^\circ$.

2.5. Experiment 2: Visual search on natural objects

The stimuli consisted of 48 images – these included 4 views each of 12 unique objects (6 animals and non-animals). The four views of each object were either profile or oblique views, each of which could be pointing left or right. The oblique view was chosen to be a three-dimensional rotation of the object in its profile view, by an angle of roughly 45° out of the image plane. Objects were chosen from a standard image database or from the internet, and were equated for brightness. I also equated image size across all profile views such that their longer dimension measured 3.84° . To ensure that oblique views appeared to be plausible three-dimensional rotations of their corresponding profile views, the oblique views were scaled such that their vertical dimension matched that of their corresponding profile view.

In the experiment, subjects performed oddball search for each of the 1128 possible pairs of the 48 images, in which each image of a pair was a target in one trial and a distractor in one trial. Thus there were a total of $1128 \times 2 = 2256$ correct trials performed by each subject. To prevent low-level visual cues such as size from contributing to search times, the distractors in the array varied in size: of the 15 distractors in the array, seven distractors had their longer dimension set to 3.84° (i.e. same as the target), and four distractors each measured 75% and 125% of this size.

2.6. Measurement of motor reaction times

To estimate the contribution of motor preparation time to visual search times, subjects were asked to perform a simple motor task before the visual search task. On each trial of the task, subjects saw a white disk appear on the left or right of the screen (with a vertical red bar down the middle of the screen), and were asked to press a key (M or Z, as before) to indicate the side on which the disk appeared. The reaction times on this task were on average 384 ms (standard deviation = 92 ms). In comparison, the average reaction times in the search tasks were: 1516 ms in Experiment 1 (standard deviation = 971 ms) and 1063 ms in Experiment 2 (standard deviation = 579 ms).

3. Results

A total of 12 subjects were recruited for two visual search experiments. In both experiments, on each trial, subjects were instructed to find an oddball item among multiple identical

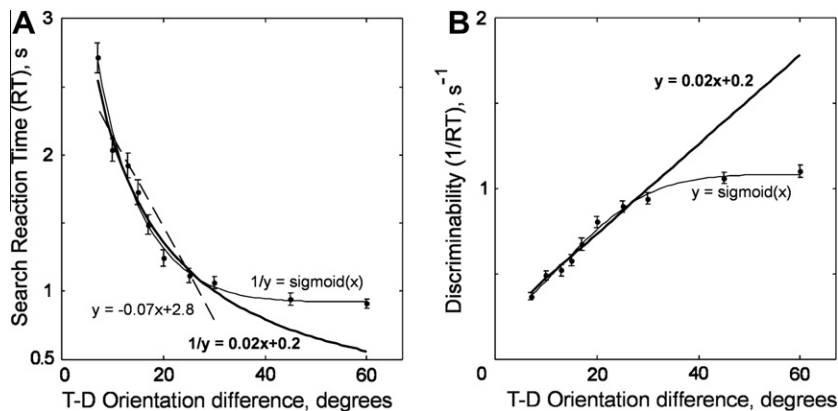


Fig. 1. Subjects performed oddball visual search for a bar that differed in orientation from multiple identical distractors (Experiment 1). (A) Visual search reaction time (RT) plotted against difference in orientation between the target and distractors. Observed reaction times (dots) are shown with error bars representing standard error of the mean (s.e.m.). These data were fit using a linear fit (dashed lines), an inverse fit (i.e. a linear fit on $1/\text{RT}$ versus orientation difference; thick line) and a sigmoid fit to $1/\text{RT}$ (thin line). (B) Reciprocal of search time, $1/\text{RT}$, plotted against the orientation difference (dots with error bars representing s.e.m.). These data were fit using a linear fit on orientation differences below 30° (thick line) and using a sigmoid function (thin line).

distractors, and report the location of the item using a key press. In Experiment 1, six subjects searched for an oddball target that differed only in orientation from the distractors. In Experiment 2, a separate set of six subjects performed searches involving all possible pairs of 48 natural images (12 natural objects in four different three-dimensional views each).

3.1. How does search time vary with target–distractor differences?

I investigated the relationship between search times and target–distractor similarity, when the target and distractors differed only in orientation (Experiment 1). To this end, I plotted the search reaction time (RT) against the difference in orientation between target and distractors (Fig. 1A). The plot shows that RT varies non-linearly with the orientation difference. To characterize the shape of this non-linearity, I fit the data using three different functions. Search times decreased faster than a simple linear decrease, as evidenced by clear deviations from a straight line fit to the data for orientation differences less than 30° ($r = 0.93$, $p = 0.0009$; Fig. 1A, dashed line). For these same orientation differences, an inverse relationship (i.e. a linear fit between $1/\text{RT}$ and orientation difference) yielded an excellent fit to the data ($r = 0.99$, $p = 5.6 \times 10^{-6}$; Fig. 1A, thick line). There was no appreciable drop in reaction times when the orientation differences increased beyond 30° . Given the inverse relationship between RT and orientation difference, I plotted $1/\text{RT}$ against the orientation difference (Fig. 1B). This $1/\text{RT}$ data was reasonably fit by a straight line for differences in orientations up to 30° ($r = 0.98$, $p = 2 \times 10^{-5}$; Fig. 1B, thick line), but for larger orientation differences the data was fit better by a sigmoid function ($r = 0.99$, $p = 5.0 \times 10^{-9}$; Fig. 1B, thin line). These sigmoid predictions, when converted back into reaction times, yielded a good fit to the reaction times as well ($r = 0.99$, $p = 1.7 \times 10^{-9}$; Fig. 1A, thin line). However, the quality of fit between the sigmoid predictions and the linear fit did not differ for orientation differences less than 30° , even though the sigmoid had an extra free parameter compared to the linear fit ($r = 0.98$ for the linear fit versus $r = 0.99$ for the sigmoid).

The above analyses show that, at least in the range where search times vary with target–distractor orientation differences, they are fit reasonably well by a linear relationship between $1/\text{RT}$ and target–distractor differences. This linear relationship suggests that $1/\text{RT}$ can be used as a measure of target–distractor dissimilarity even when the underlying features are not known.

3.2. Does $1/\text{RT}$ have the properties of a distance metric?

Before proceeding to use $1/\text{RT}$ to characterize perceptual distances, I sought to verify whether it satisfies the properties of a mathematical distance metric (Tversky, 1977). The first two properties are trivial to establish. First, that the distance must be positive is immediately satisfied because reaction times are positive. The second property requires that distance is zero only if target and distractor are identical. Although this cannot be verified in practice, this property is consistent with the fact that reaction times become very large as the target–distractor difference decreases, making $1/\text{RT}$ very small (Fig. 1A).

We now turn to the third property which is the triangle inequality. In the context of visual search distances, given any three objects A, B and C and their pair-wise distances d_{AB} , d_{BC} and d_{AC} , the triangle inequality requires that the sum of the two sides of a triangle ($d_{AB} + d_{BC}$) be greater than the third side (d_{AC}). This property must hold for each of the three distances. Verifying this property requires measuring these three distances for a large number of object triads. To this end, I performed an additional visual search experiment (Experiment 2) in which the stimuli were 48 images of 12 natural objects each in four possible three-dimensional views. To maximize the number of available triads, subjects were required to perform a total of 1128 visual searches in which every possible pair of these 48 images were shown as target and distractor ($48 \text{ choose } 2 = 1128$). This yielded a total of 17,296 triads ($48 \text{ choose } 3$) on which the triangle inequality could be tested, with each triad giving rise to three possible comparisons (one for each side of the triangle against the sum of the other two sides). For each pair of images, the search time was taken as the average of the search times for either image in the pair as target (i.e. ignoring search asymmetry for the moment – see below for a detailed analysis of asymmetry). In the triangle ABC formed by each triad of images, the third side (d_{AB} , d_{BC} or d_{AC}) was plotted against the sum of the other two sides ($d_{AC} + d_{BC}$, $d_{AB} + d_{AC}$ or $d_{AB} + d_{BC}$ respectively). The resulting plot (Fig. 2) shows that only 143 of the 51,288 points – a mere 0.2% of all possible triplets – violate the triangle inequality (i.e. fall above the $y = x$ line).

There are potentially two ways in which this result may have been obtained as a trivial consequence of nature of the data. First, the triangle inequality may be trivially satisfied if search times are dominated by a motor preparation time: If RT_{AB} , RT_{BC} and RT_{AC} were all equal to a motor preparation time M , then $1/\text{RT}_{AB} + 1/\text{RT}_{BC}$ would be $2/M$, which is greater than $\text{RT}_{AC} = 1/M$. However,

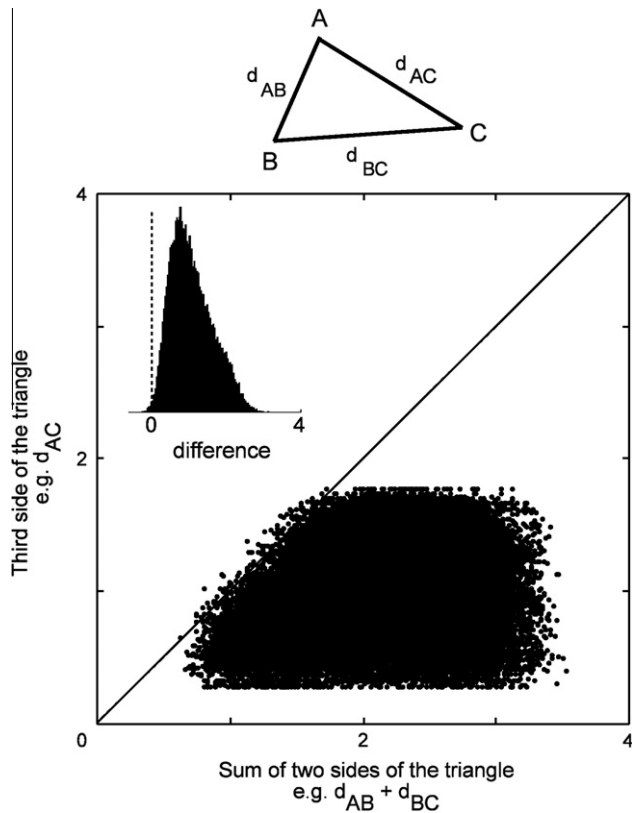


Fig. 2. Subjects performed multiple visual searches involving all possible pairs of 48 images with each image in a pair as target or distractor (Experiment 2). For each triplet of images (A,B,C) – depicted schematically above the plot – the pairwise dissimilarity measurements (1/RT) were denoted by d_{AB} , d_{BC} and d_{AC} . If 1/RT satisfies the triangle inequality, each of these distances (plotted on the y-axis, depicted as d_{AC}) must be smaller than the sum of the other two (plotted on the x-axis, depicted as $d_{AB} + d_{BC}$). The inset shows the histogram of the difference ($d_{AB} + d_{BC} - d_{AC}$) across all triplets. Triplets with a difference less than zero (in the inset plot) or triplets that fall below the $y = x$ line (in the scatter plot) violate the triangle inequality. Such violations occurred only in 143 of 51,888 triplets, i.e. in 0.2% of all triplets.

according to my estimates, the baseline motor reaction time (mean = 384 ms; see Section 2.6) is roughly only one third of the search reaction times in these experiments (mean = 1289 ms). To investigate this further, I repeated this analysis on search times of each subject after subtracting his/her mean motor reaction times. Although in this case the triangle inequality is violated more frequently than before (3560 of 51,288 triplets violated the inequality, i.e. 7% of all cases; data not shown), the predominant trend still favored the triangle inequality since it was satisfied in 93% of all triplets.

Second, the triangle inequality may be trivially satisfied by any three randomly chosen distances: if d_{AB} , d_{BC} and d_{AC} are all random positive numbers from the same distribution, then $d_{AB} + d_{BC}$ may frequently be larger than d_{AC} . To investigate this further, this analysis was repeated by choosing three random distances for each triplet, so that they no longer came from triplets of objects (A–C). If the triangle inequality arises trivially from the observed distribution of distances in the data, then the proportion of triplets that violate the triangle inequality should be approximately the same as that observed in the data. In this shuffle control analysis, 1860 of the 51,288 triplets (3.5%) of the triplets violated the triangle inequality – a significantly greater proportion of the triplets than the 0.2% observed in the data ($p = 0$, χ^2 test).

To allay the concern that the above two effects (motor reaction time and random sampling) might have together contributed to the observed effect, I repeated the analysis by first subtracting

the motor reaction times as before, and then randomly choosing distances. In this case, roughly 6850 of 51,888 triplets (13%) of all triplets violated the triangle inequality, again a significantly greater proportion than the 7% observed in the data when motor reaction times were factored out ($p = 0$, χ^2 test). Thus, the observed number of violations of the triangle inequality is far smaller than expected from the combined effect of motor reaction times and the distribution of distances in the data. Thus, search dissimilarity (1/RT) does indeed satisfy the triangle inequality.

The fourth and final property of a distance metric is that distances should be symmetric. In the context of visual search, this implies that for any two objects A and B, $d_{AB} = d_{BA}$. This implies in turn that search times for A among Bs should be identical to search times for B among As, which is clearly violated in the case of asymmetries (Wolfe, 2001). However, the relative frequency or magnitude of search asymmetries for natural objects is unknown. I therefore set out to establish the relative frequency and magnitude of search asymmetries on the objects used in Experiment 2. For each of the 1128 pairs of images of the form (A and B), I took the average search times for A among Bs and that of B among As across subjects, and assessed whether they were significantly different using a paired t -test ($\alpha = 0.05$). This revealed a total of 69 of 1128 image pairs (only 6% of pairs – no greater than the 5% expected by chance alone) with asymmetric search times. Thus asymmetries are relatively rare among a set of natural objects. However, these asymmetries, although rare, may be substantial in magnitude when they occur. To investigate their magnitude, for each image pair, the dissimilarity of the easy target (i.e. the shorter RT in AB versus BA) was plotted against the dissimilarity of the hard target (i.e. the longer RT). The resulting plot (Fig. 3A) reveals a surprising regularity in the magnitude of the asymmetries – they all fall along a straight line whose slope is nearly equal to 1 (best-fit slope = 0.99), with an intercept of 0.44. Thus, the magnitude of the asymmetry ($d_{AB} - d_{BA}$, where A is the easy target) is fixed at 0.44 for this set, and is independent of the dissimilarity between the target and distractor. This difference is relatively small compared to the variation in the dissimilarity across different image pairs (min = 0.3 s^{-1} , mean = 1.15 s^{-1} , max = 1.71 s^{-1}). In contrast, an analogous plot of the asymmetric search times (Fig. 3B) reveals a linear dependence between the search times for the easy target versus the hard target in each pair with a slope equal to 0.56. In other words, the magnitude of the asymmetry ($RT_{BA} - RT_{AB}$ where A is the easy target) increases with the mean search time. Thus, although search asymmetries can and do occur for natural objects, they are relatively rare and have a surprising regularity when characterized by the search distance (1/RT).

Why is it that asymmetry in search distance is constant whereas asymmetry in search time is dependent on the search time? This discrepancy can be understood by considering the transformation between RT to 1/RT (Fig. 3C). Consider two image pairs (A and B) and (C and D) which have asymmetries that differ in dissimilarity by the same amount Δd (i.e. $d_{AB} - d_{BA} = d_{CD} - d_{DC} = \Delta d$), except that in one image pair, the average search time is small (labeled as ΔRT_1 in Fig. 3C) and in the other pair, search time is large (labeled as ΔRT_2 in Fig. 3C). Because the slope of the function 1/RT decreases with increasing RT, it can be readily seen that $\Delta RT_1 < \Delta RT_2$. As a result, the magnitude of the asymmetry will depend on the mean reaction time, as observed in Fig. 3B. This analysis reveals an unexpected insight into search asymmetries using search distance (1/RT) that could not have been obtained using search times alone, namely that at least on a set of natural objects varying in their three-dimensional view, all asymmetries differ by a fixed distance in visual search space.

To summarize, the dissimilarity measure 1/RT is positive, approaches zero as the target and distractor become identical, generally satisfies the triangle inequality and is approximately

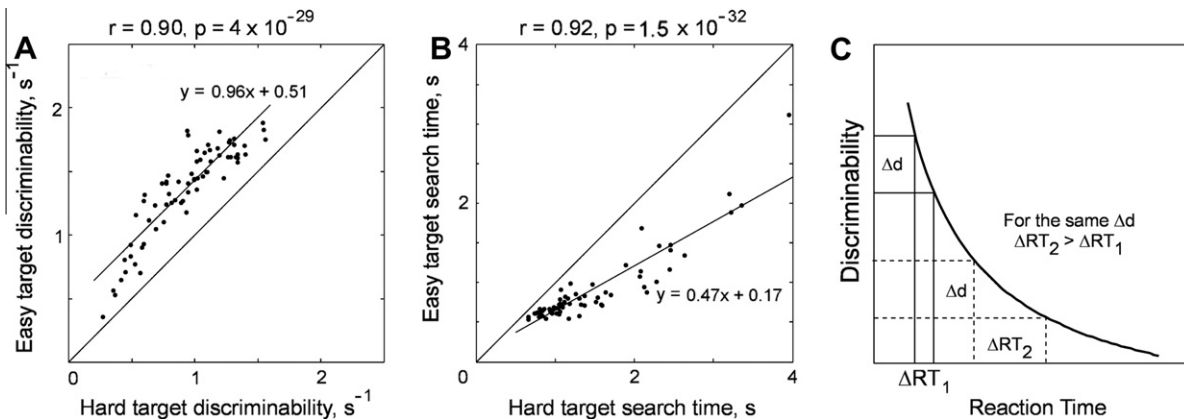


Fig. 3. Asymmetry analysis. (A) A total of 69 of 1128 image pairs in Experiment 2 showed a statistically significant asymmetry. For each of these image pairs, the discriminability (1/RT) of the search involving the easy target is plotted against the discriminability of the hard target. The data was fit using a straight line with slope of nearly 1. (B) The same data represented using search times (RT). For the same 69 image pairs, the search times for the easy target are plotted against the search times for the hard target. The data suggest that the size of the asymmetry increases in proportion to the average search time. (C) Illustration of how asymmetries differing by a constant amount in Δd (as shown in A) might give rise to RT differences that increase with the mean (as shown in B) – see text for details.

symmetric as revealed by relatively few asymmetries. We conclude that, to a first approximation, the dissimilarity measure 1/RT can be treated as a valid distance in visual search space.

3.3. Object representations as revealed using search distances

Having established that the dissimilarity measure is a valid distance, I set out to investigate whether dissimilarity measurements on a large object set can elucidate the underlying object representation. As an example, consider all possible 276 pair-wise searches performed on 24 images in Experiment 2 (six objects in four different three dimensional views). Since the reciprocal of these search times are distances in visual search space, I used multi-dimensional scaling to embed these images in a two-dimensional space (Fig. 4). In multidimensional scaling, the coordinates of each image are chosen such that pair-wise distances between images in two-dimensional space are as close as possible to the distances observed during visual search. These two-dimensional distances were indeed a close approximation to the observed distances, as

evidenced by a highly significant correlation between them ($r = 0.78, p = 3 \times 10^{-57}$). The resulting plot shows that objects short of vertical mirror reflection are close together (Fig. 4) – this is concordant with the mirror confusion observed at both behavioral and neural levels (Gross, 1978; Rollenhagen & Olson, 2005). In addition, the plot also shows that the three-dimensional views of each object are clustered together in visual search space: in other words, searches involving different views of an object are hard whereas searches involving multiple objects are easy. Note that if the underlying object representation were completely invariant, multiple views of an object would be indistinguishable in visual search, which was not the case (these distances are non-zero in Fig. 4). Thus, object representations underlying visual search are roughly invariant to three-dimensional view.

4. Discussion

Here, I have proposed a distance measure for similarity relations in visual search, namely the reciprocal of search reaction time

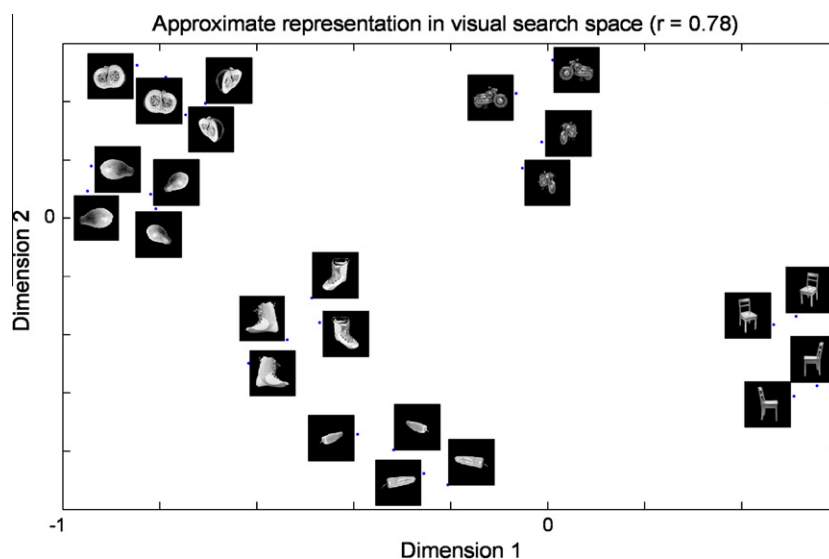


Fig. 4. Visualization of search distances using multi-dimensional scaling (Experiment 2). Multi-dimensional scaling was performed on all 276 pair-wise search distances (1/RT) between 24 images (6 objects × 4 views each), to find the best-fitting two-dimensional coordinates such that their distances match the observed pair-wise distances. The resulting plot shows clustering of objects by view as well as mirror confusion between left and right views.

(1/RT). This measure yields several novel insights. First, unlike search time, 1/RT varies linearly over a large range of target–distractor differences for variations along a single known feature dimension (orientation), making it a potentially useful tool to discover novel features that guide search. Second, 1/RT can be treated literally as a distance in visual search space, because it satisfies the conditions required for a mathematical distance metric. Third, search asymmetries, as measured using 1/RT, differ from each other by a fixed quantum of dissimilarity. Fourth, pair-wise distances measured in this manner reveal that object representations underlying visual search are roughly view-invariant. Below I review the relevance of these findings in the context of the existing literature.

4.1. 1/RT as a distance measure

I have shown that search distance calculated as 1/RT has the properties required of a metric distance: in particular it satisfies the triangle inequality and is roughly symmetric. This investigation of symmetry properties has revealed two novel insights into the nature of search asymmetries: First, at least for a set of natural objects varying in their three-dimensional view, asymmetries are relatively rare, occurring only 6% of the time. The relative frequency with which search asymmetries do occur in naturalistic searches has not been reported before, but the generality of this result remains to be established. Second, the asymmetries that do occur differ from each other by a fixed distance in visual search space. This regularity in the object representation would never have been observed using search times alone (Fig. 3). Although this result does not explain why search asymmetry occurs in the first place, the constant difference between many asymmetric searches implies that the underlying mechanism must be independent of the processes that generate the discriminative signal itself.

Recent theories of search asymmetry have proposed that asymmetries occur when the variance in the representation of the easy target is larger than that of the hard target (Palmer, Verghese, & Pavel, 2000; Rosenholtz, 2001; Verghese, 2001; Vincent, 2011). In the framework of signal detection theory, a typical sample from a distribution with large variance (i.e. the easy target) is unlikely to have come from a distribution with small variance (the distractor), and is therefore easy to distinguish from the small-variance distribution. Conversely, a sample from a small-variance distribution is more likely to have come from the large variance distribution, and is therefore harder to distinguish from the large-variance distribution. As a result, it is easier to distinguish a large-variance sample from a small-variance sample than vice versa. Our finding can easily be incorporated into this framework if one interprets search distance as the discriminative signal that drives visual search (see below). A target representation with large variance will produce a large difference signal, in turn leading to a short reaction time – and vice versa. Thus, our finding that many asymmetries differ by a fixed search distance might arise from a bimodal distribution in the variance of the representation across objects. Importantly, while the asymmetry might arise from a difference in variance, object discriminability may arise primarily from a difference in means between the two distributions. It is thus possible that these two effects act independently to influence visual search. However, there is still an important gap in our understanding, namely the mechanistic link between the mean/variance of the object representations and the generation of the discriminative signal that guides visual search.

4.2. Search distances as a measure of dissimilarity

Although search is known to become hard when the target is similar to its distractors (Alexander & Zelinsky, 2012; Duncan &

Humphreys, 1989; Wolfe et al., 1992; Yang & Zelinsky, 2009) or when distractors become heterogeneous (Bauer, Jolicoeur, & Cowan, 1996; Duncan & Humphreys, 1989; Palmer, Verghese, & Pavel, 2000; Neider & Zelinsky, 2011), very few studies have quantified the relationship between search times and target–distractor differences. Our finding that search time is inversely proportional to target–distractor orientation difference is consonant with a similar report based on search slopes (Wolfe, Klempe, & Shulman, 1999). It is also qualitatively similar to the exponential decrease in search times observed with target–distractor size differences (Blough, 1988) and with color differences (Nagy & Cone, 1996; Nagy & Sanchez, 1990) and with complex shapes (von Grunau, Dube, & Galera, 1994). Although an inverse relationship is quantitatively different from an exponential decay, these have not been compared directly. To investigate this further, I fit the data in Fig. 1A using both an exponential RT model and the 1/RT model. The residual sum of squared error for the 1/RT model ($sse = 0.07$) was smaller than that of the exponential model ($sse = 0.24$) and this difference approached significance ($p = 0.06$, paired t -test). Thus at least for the data in Experiment 1, a linear relationship between 1/RT and target–distractor differences appears to be a better fit. To resolve this discrepancy between exponential and linear relations will require experiments that directly compare these fits using several types of feature differences.

4.3. Object representations underlying visual search

The results of this study motivate the use of visual search distance to elucidate the object representations that underlie visual search. In particular, I propose two ways in which this distance measure might be used: First, because it varies linearly with target–distractor differences (at least when RT changes), it can be used to discover features underlying search: a feature whose differences vary linearly with 1/RT may be deemed better than one that has a non-linear relationship with 1/RT. Thus, quantitative differences in the predictions based on different features can be distinguished even if they qualitatively co-vary. Second, pair-wise measurements of distance can be used to visualize the underlying object representation. Specifically, I have shown that the representation of natural objects in visual search space is roughly invariant to changes in three-dimensional view.

The finding that object representations in visual search space are roughly view invariant is consistent with evidence that visual search is sensitive to three-dimensional structure (Enns & Rensink, 1990; von Grunau & Dube, 1994). However I have gone further to demonstrate that different three-dimensional views of the same object form separate clusters in visual search space. This is a novel finding in visual search, but it is concordant with a growing body of evidence that even higher-level feature representations can guide visual search (Wolfe & Horowitz, 2004). This finding is also consistent with view invariance reported in other behavioral paradigms (Logothetis & Pauls, 1995) and in neuronal activity from high-level visual cortex (Freiwald & Tsao, 2010; Logothetis, Pauls, & Poggio, 1995). The most parsimonious explanation for these observations is that different views of an object share multiple low-level features, resulting in similar neuronal and behavioral responses including in visual search. Alternatively, object representations may be roughly view-invariant even after controlling for low-level feature differences. Distinguishing between these possibilities will require testing view invariant representations for objects varying in their perceived similarity.

4.4. Relation to models of visual search

The distance measure 1/RT has an obvious relationship to accumulator models of visual search in which a difference signal

is integrated to threshold in the presence of noise (Purcell et al., 2010; Schall et al., 2011). Although the precise time to threshold is difficult to calculate analytically under biophysically plausible conditions of leak and noise, the reciprocal of the reaction time ($1/RT$) is roughly proportional to the magnitude of the difference signal (Tuckwell, 1988). Where does this difference signal arise? In a recent study involving targets differing from distractors in global arrangement, we have shown that the reciprocal of search time is tightly correlated with neuronal discriminability in inferior temporal cortex and to differences in coarse image content (Sripati & Olson, 2010). However this result may only pertain to targets and distractors differing in complex aspects of shape where high-level areas are likely to contribute to search. In general it is possible that the difference signal that underlies search may be based on neuronal activity differences throughout the visual cortex.

5. Conclusions

Although search times (RT) and accuracy remain the most direct measures of search performance, search distance ($1/RT$) can be used to yield additional insights into the processes underlying visual search and has a straightforward interpretation as the difference signal that drives search. Search distance is advantageous particularly in the investigation of similarity relations between objects, because, unlike search time, it varies linearly with dissimilarity. Characterizing visual search in terms of search distance also brings up several interesting questions. For instance, how do search distances combine for multiple features? Can search distance be used to discover novel features that guide visual search? Can search distances be used to explain more complex search phenomena? Ultimately, the usefulness of this distance measure will lie in its ability to further elucidate search phenomena.

Acknowledgments

I thank Jeremy Wolfe for helpful references, and two anonymous reviewers for their comments on this work. I also thank Vigneshvel T. and Krithika Mohan for help with data collection, Supratim Ray for comments and other lab members for valuable discussions. This research was supported by a startup grant from the Indian Institute of Science and an Intermediate Fellowship from the Wellcome Trust-DBT India Alliance.

References

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The frankenbear experiment. *Vision Research*, 54C, 20–30.

Basri, R., Costa, L., Geiger, D., & Jacobs, D. (1998). Determining the similarity of deformable shapes. *Vision Research*, 38(15–16), 2365–2385.

Bauer, B., Jolicoeur, P., & Cowan, W. (1996). Distractor heterogeneity versus linear separability in visual search. *Perception*, 25, 1281–1293.

Blough, D. S. (1988). Quantitative relations between visual search speed and target-distractor similarity. *Perception & Psychophysics*, 43(1), 57–71.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.

Cortese, J. M., & Dyre, B. P. (1996). Perceptual similarity of shapes generated from Fourier descriptors. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1), 133–143.

Desmarais, G., & Dixon, M. J. (2005). Understanding the structural determinants of object confusion in memory: An assessment of psychophysical approaches to estimating visual similarity. *Perception & Psychophysics*, 67(6), 980–996.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433–458.

Edelman, S. (1998). Representation is representation of similarities. *Behavioral and Brain Sciences*, 21(4), 449–467 [discussion 467–498].

Enns, J. T., & Rensink, R. A. (1990). Influence of scene-based properties on visual search. *Science*, 247(4943), 721–723.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, 330(6005), 845–851.

Gross, C. G. (1978). Inferior temporal lesions do not impair discrimination of rotated patterns in monkeys. *Journal of Comparative and Physiological Psychology*, 92(6), 1095–1109.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cerebral Cortex*, 5(3), 270–288.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, 5(5), 552–563.

Nagy, A., & Cone, S. M. (1996). Asymmetries in simple feature searches for color. *Vision Research*, 36(18), 2837–2847.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 7(7), 1209–1217.

Neider, M. B., & Zelinsky, G. J. (2011). Cutting through the clutter: Searching for targets in evolving complex scenes. *Journal of Vision*, 11(14).

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40(10–12), 1227–1268.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, 117(4), 1113–1143.

Rollenhagen, J. E., & Olson, C. R. (2005). Low-frequency oscillations arising from competitive interactions between visual stimuli in macaque inferotemporal cortex. *Journal of Neurophysiology*, 94(5), 3368–3387.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics*, 63(3), 476–489.

Schall, J. D., Purcell, B. A., Heitz, R. P., Logan, G. D., & Palmeri, T. J. (2011). Neural mechanisms of saccade target selection: Gated accumulator model of the visual-motor cascade. *European Journal of Neuroscience*, 33(11), 1991–2002.

Sripati, A. P., & Olson, C. R. (2010). Global image dissimilarity in macaque inferotemporal cortex predicts human visual search efficiency. *Journal of Neuroscience*, 30(4), 1258–1269.

Tuckwell, H. (1988). *Introduction to theoretical neurobiology. Non-linear and stochastic theories* (Vol. 2). Cambridge University Press.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31(4), 523–535.

Vincent, B. T. (2011). Search asymmetries: Parallel processing of uncertain sensory information. *Vision Research*, 51(15), 1741–1750.

von Grunau, M., & Dube, S. (1994). Visual search asymmetry for viewing direction. *Perception & Psychophysics*, 56(2), 211–220.

von Grunau, M., Dube, S., & Galera, C. (1994). Local and global factors of similarity in visual search. *Perception & Psychophysics*, 55(5), 575–592.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, 63(3), 381–389.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 419–433.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. L., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 34–49.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501.

Wolfe, J. M., Klempen, N. L., & Shulman, E. P. (1999). Which end is up? Two representations of orientation in visual search. *Vision Research*, 39(12), 2075–2086.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, 49(16), 2095–2103.