

Research



Cite this article: Amor TA, Luković M, Herrmann HJ, Andrade Jr JS. 2017 Influence of scene structure and content on visual search strategies. *J. R. Soc. Interface* **14**: 20170406. <http://dx.doi.org/10.1098/rsif.2017.0406>

Received: 1 June 2017

Accepted: 30 June 2017

Subject Category:

Life Sciences – Physics interface

Subject Areas:

biocomplexity, computational biology

Keywords:

visual search, eye movement, search strategies

Author for correspondence:

José S. Andrade Jr

e-mail: soares@fisica.ufc.br

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3825496>.

Influence of scene structure and content on visual search strategies

Tatiana A. Amor^{1,2}, Mirko Luković¹, Hans J. Herrmann^{1,2}
and José S. Andrade Jr²

¹Computational Physics IffB, ETH Zurich, Stefano-Franscini-Platz 3, 8093, Zurich, Switzerland

²Departamento de Física, Universidade Federal do Ceará, 60451-970, Fortaleza, Ceará, Brazil

TAA, 0000-0002-6433-390X

When searching for a target within an image, our brain can adopt different strategies, but which one does it choose? This question can be answered by tracking the motion of the eye while it executes the task. Following many individuals performing various search tasks, we distinguish between two competing strategies. Motivated by these findings, we introduce a model that captures the interplay of the search strategies and allows us to create artificial eye-tracking trajectories, which could be compared with the experimental ones. Identifying the model parameters allows us to quantify the strategy employed in terms of ensemble averages, characterizing each experimental cohort. In this way, we can discern with high sensitivity the relation between the visual landscape and the average strategy, disclosing how small variations in the image induce changes in the strategy.

1. Introduction

What visual strategy do we employ when searching for a familiar face in a crowd and how would it change if we had to find a friend in a well-organized choir? For instance, do we explore each face sequentially or do we pick them randomly until reaching our target. Along these lines, it is not at all obvious whether there exists a characteristic strategy that is related to the scene content. To shed light into this problem, it is necessary to find a method that allows us to identify and quantify particular features associated to the strategy adopted while searching for a hidden target.

Different models have been developed with the goal of understanding what guides eye movement during visual search (see [1] for a review). One family of these models is based on the construction of *saliency maps*, which define regions of interest derived from properties of scene objects (such as luminescence, colour and orientation) [2–7]. By definition, these salient regions stand out from other parts of the scene and are therefore more susceptible to frequent eye fixations. In the context of visual search, these models prove to be more suitable for tasks involving a small number of equally relevant *distractors*, which are essentially all items in the scene that are not targets of the search [8,9]. On the other hand, in more complex visual search tasks, the salient regions might not necessarily be relevant. In order to address this issue, other implementations of the saliency model consider the relative information of an object with respect to the global information of the scene [10–12].

In the absence of salient elements, it is not possible to use these models as there are no *a priori* privileged regions within the scene. As a consequence, another family of visual searching models have been proposed, namely, the *saccadic targeting models* [13–17]. In these studies, the main hypothesis is that saccadic eye movements are directed to locations within the scene that contain elements similar to the target. This similarity can be due to the image content and to a neurobiological filter. Within this framework, Najemnik & Geisler [18,19] propose a model where each point in space has a certain probability of being explored and the saccadic movement is directed to the most probable

regions. These probabilities are then updated over time so that regions that were already explored are less likely to be revisited, introducing in a natural way the notion of persistence while searching. Moreover, other implementations of this model have taken into account the proximity between consecutive saccadic movements by adding a cost function that punishes longer saccades [20].

Scanpaths produced by eye movements, in general, reveal different forms of persistence, at the level of saccades [21] and also within the fixations [22,23]. During most forms of random search, persistence of the jittering movements (i.e. saccades in visual search) plays a crucial role because it unveils the strategy involved [24,25]. In fact, while looking for a hidden target in a field of distractors, a variety of patterns have been documented [21,26–28], ranging from systematic or completely persistent to random. However, while some of the models described above do make use of some form of persistence, they do not relate it to the overall strategy of the search. Besides some exceptions [29,30], the saliency models do not focus on understanding the sequence in which the fixations are performed but instead determine the regions where they are more likely to occur. On the other hand, the saccadic targeting models, although dealing with the saccadic sequence, do not account for different strategies.

Here, we propose a visual search model (VSM) that quantifies the global persistent behaviour and the overall strategy employed while looking for a hidden target. The parameters of the VSM define the saccadic orientation distribution, which has been experimentally proved to provide information regarding the strategy [21] and the scene content [27]. By studying this distribution obtained from experimental data, we were able to identify different search strategies that emerge from exploring the same scene and quantify them through the VSM parameters. We analysed the search strategies adopted in three different visual search scenes that differ in the features and arrangement of the scene items. We found that the average strategy changes with these scenes suggesting that scene content and structure influence the way subjects execute their search.

In what follows, we first describe the VSM and how to extract the model parameters from the experimental scanpaths during a search episode, which we will refer to as the ‘cloud number (CN)’ experiment. In this visual search task, the participants are requested to find a unique number ‘5’ embedded in a field of numbers ranging from ‘1’ to ‘9’ serving as distractors. By means of an efficiency measure about the eye paths, we validate our model with respect to the experimental data. Then, we identify the location of the experimental trajectories for three different visual search tasks in the model parameter space and compare the average strategies applied for each task. Given that the average position within the parameter space differs for the three cohorts, the VSM can serve as a tool to predict the average strategy of similar experiments. Finally, we discuss the implications of our results in the frame of visual search and the possibilities of exploiting the model to be used in other fields.

2. Visual search model

While searching for a hidden target during a visual search episode, participants perform different search strategies that

vary from a very systematic to a seemingly random, as depicted in figure 1. These strategies are identified by studying the relative orientation between saccadic movements performed while searching. The saccadic relative angle distribution provides information on the type of strategy employed [21] and may as well provide insight regarding the scene an observer is looking at [27]. Therefore, the aim of the VSM is to emulate visual search trajectories via the study of the distribution of the inter-saccadic angles as main ingredient.

We sample the relative angle, θ , between two consecutive saccadic movements from two distributions, as exemplified in figure 2*a,b*. The first distribution corresponds to a uniform distribution of all possible angles between 0 and 2π (figure 2*a*), and the second one is a uniform distribution with angles restricted to the interval $[-\delta\theta, \delta\theta]$ (figure 2*b*). Angles are chosen from the second distribution with probability λ , and from the first distribution with probability $1 - \lambda$. Therefore, only two parameters, λ and $\delta\theta$, govern the statistical properties of the relative angle distribution. The combination of these parameters gives rise to a wide range of different trajectories. In the case $\lambda = 0$, angles can be sampled only from the first distribution, corresponding to a purely random strategy (figure 2*c*). On the other hand, if $\lambda = 1$, the angle is only sampled from the step distribution, so that only angles in the range $[-\delta\theta, \delta\theta]$ are allowed. If in the latter case, we further assume that $\delta\theta$ is very small, so that each saccadic movement persists in the direction of the previous one, then the emergent trajectory will correspond to a systematic type of search (figure 2*e*).

Between the two extreme cases, we find a repertoire of trajectories that result from the interplay between the parameters λ and $\delta\theta$ (figure 2*d*). The parameter λ is related to the strategy employed during the search and $\delta\theta$ to the persistent behaviour of the search. For instance, an intermediate value of λ combined with a small value of $\delta\theta$ results in a trajectory that can be identified as a random search with a strong persistence imprint.

Besides the relative angle distribution, another important ingredient in our model is the length of the saccadic movements, i.e. the length of the jumps. We consider them to be sampled from a lognormal distribution, as reported in [26].

2.1. Estimation of λ and $\delta\theta$ from the experimental data

From the VSM, we propose the inter-saccadic relative angles, θ , to be sampled from a combination of a uniform distribution over all angles and one in the interval $[-\delta\theta, \delta\theta]$ (figure 2). A conventional method to acquire the parameters from experimental data would be to study the histogram of θ . However, one issue of this approach is that the parameter $\delta\theta$ is bin-size dependent, and thus is not a robust method for the parameter estimation. Moreover, the average number of θ values obtained in each trajectory is around 250, thus making it difficult to determine with precision the distribution from each of the scanpaths. One way around this issue is to extract the model parameters directly from the experimental data. When the angles are ranked in increasing order from $-\pi$ to π , they present a characteristic shape of three linear segments, as shown in figure 3*a* with green dots. This shape is similar for all the trajectories (see electronic supplementary material, figure S1). In the extreme case that θ is sampled from the uniform distribution between

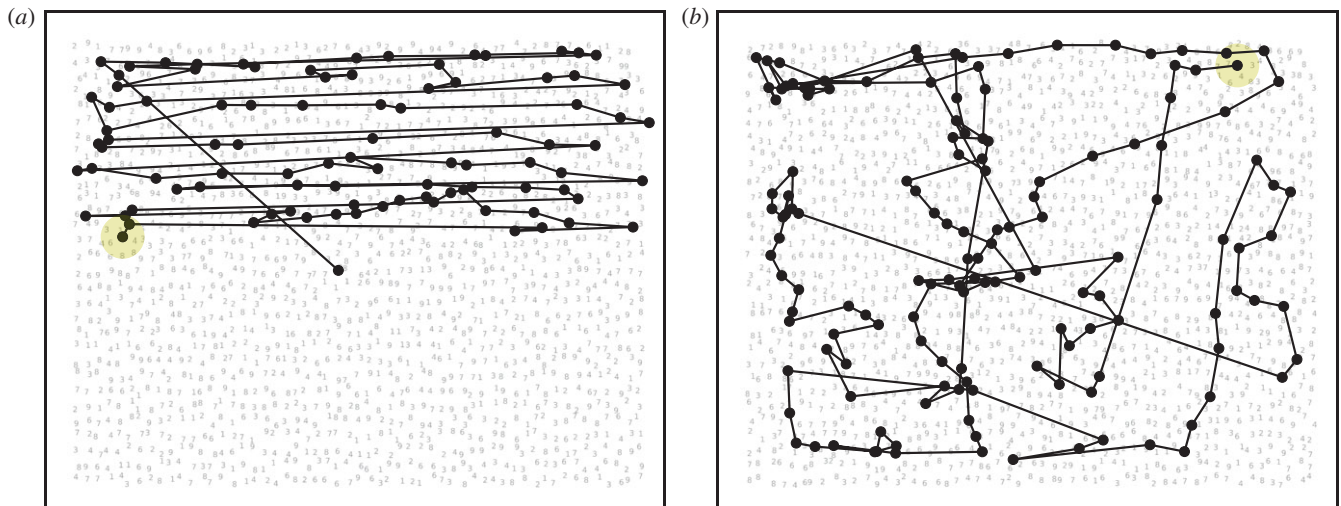


Figure 1. Different types of visual trajectories are found in the ‘Cloud number (CN)’ experiment. During a visual searching task, where participants have to find a unique number ‘5’ embedded in a cloud of numbers between ‘1’ and ‘9’ (namely, the CN experiment), participants perform very different trajectories. These trajectories range from a systematic search (a) to a trajectory resembling more a random search with a strong persistence imprint (b). (Online version in colour.)

– π and π , the ranked angles form a straight line with a slope equal to 2π . On the other hand, if θ is sampled from the uniform distribution between $-\delta\theta$ and $\delta\theta$, the ranked angles form a straight line where $\theta_{\max} - \theta_{\min} = 2\delta\theta$. Considering the evident threefold behaviour of the ranked angles obtained from the experiments and its origin as a combination of these three extreme cases, we fit the data using three consecutive linear segments (see figure 3a, solid orange lines). The middle region depends on $\delta\theta$, and the first and third regions are related to the uniform distribution between $-\pi$ and π . The estimation of $\delta\theta$ is straightforwardly performed as the difference between the angle values defining the middle section equals $2\delta\theta$. The parameter λ is the ratio between the number of angles originated by the oriented distribution, N_λ , and the total number of angles, N . Within this framework, N_λ corresponds to the difference between the total number of angles and n_{step} , defined as the angle index value for which the extension of the first linear segment is equal to π . This allows us to estimate the parameters of the VSM using the experimental data without the need of any intermediate manipulation.

3. Results

In order to compare the VSM with the experimental data, we study the efficiency, ε , defined in terms of the space-filling attributes of the trajectories investigated (similar to the method applied in [31]). The underlying hypothesis is that if the trajectory is able to fill the image space, then the chances of finding the target increase, as more locations are examined. The efficiency is measured in the following way: given a trajectory of a finite amount of steps, N , we divide the figure being explored with a grid of $m \times m$ cells. Next, we determine the percentage of cells that were explored by the trajectory, as shown in figure 3b. When a fixation point falls into a cell, this cell is considered to be explored, because it is during fixational events that most visual information is gathered [32]. On the other hand, if a saccadic movement passes over a cell but there are no fixations inside it, the cell is considered unexplored.

For a fixed length N of the trajectory, we have different mean ε values within the parameter space. An efficient trajectory is the one that visits many locations, i.e. has a large value for ε . The efficiency values obtained in the parameter space are depicted in figure 3c. As expected, the values corresponding to $\lambda = 0$ and any $\delta\theta$ are the same as the ones obtained for $\delta\theta = \pi$ and any λ value, because these trajectories emerge from the same distribution. The optimal values are those having a large λ value and a small $\delta\theta$. These are typical of a systematic type of search, in which very few locations are inspected again. Therefore, the final trajectory explores more locations when compared with the random strategy.

From the study of the relative inter-saccadic angle distribution of the CN cohort, we estimate the corresponding parameters λ and $\delta\theta$. When placed in the parameter space, we see that the experimental data present a dispersion in the parameter λ ranging from 0.1 to 0.8 (figure 3c). Interestingly, the estimated experimental values for $\delta\theta$ are confined within a range related to small angles, therefore associated to a persistent movement. We find that the centre of mass of these data points, i.e. the mean values, is close to 0.4 for λ , and to $\pi/8$ for $\delta\theta$.

For the estimated values of λ and $\delta\theta$, we run the VSM with the corresponding values of N for each trajectory and compare the obtained ε values. The VSM matches 78% of the experimental data, as shown in figure 3d. For some points, however, the model fails to predict the expected ε . For instance, one participant initiates with a systematic search and later breaks away from that strategy to go for a systematic search truncated in space, i.e. without reaching the right end of the image (see electronic supplementary material, figure S2). Our model is not able to reproduce this behaviour and, therefore, for the same set of parameters, λ and $\delta\theta$, it generates instead a more efficient search. However, the large number of trials that have matching ε values to the ones obtained by means of the VSM suggest that the information carried on the inter-saccadic angle distribution is enough to reproduce statistical similar scanpaths.

Interestingly, for the CN cohort we do not find an average behaviour that is purely systematic, namely a λ value close to unity. As a consequence, one could ask whether the average

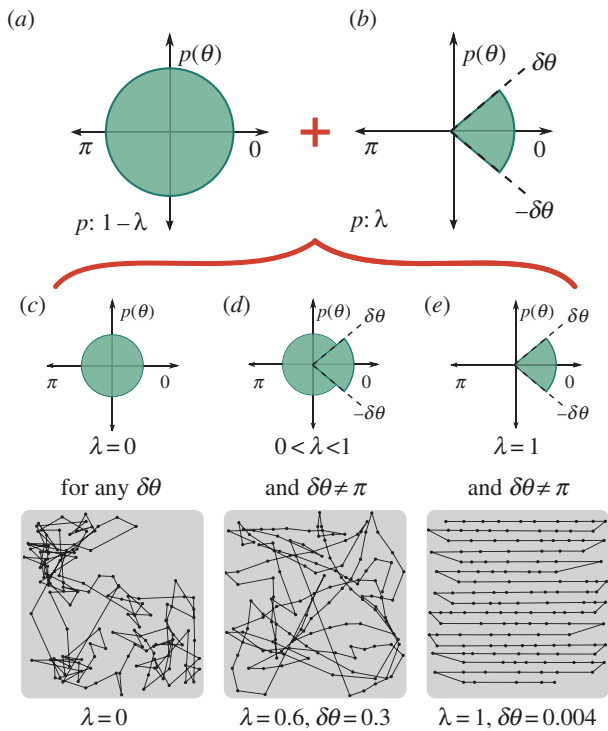


Figure 2. Schematic definition of the VSM. Two parameters, λ and $\delta\theta$, govern the way to select the relative angles between two consecutive saccadic movements. We consider two uniform distributions: one over all possible angles (a) and one in the range $[-\delta\theta, \delta\theta]$ (b). With probability $1 - \lambda$, angles are chosen from the first distribution and with probability λ from the second one. Different combinations of these parameters result in a variety of distributions (d) responsible for different visual search strategies, ranging from a pure random walk (c) to a systematic search (e).

searching behaviour changes for different searching tasks. We analysed two other experimental cohorts, namely, the 5-2 and *Where's Wally?* (WW) sets, as reported in [26] (see Material and methods). The 5-2 experiment consists in finding a target, namely, a single number '5', in a regular array of distractors, namely, numbers '2', that can be in red or green. The WW experiment corresponds to scanning images from the famous children's book where the goal is to find Wally, a character wearing a striped shirt in a crowd of people [33]. A visual search task as the 5-2 experiment, where the target and distractors are placed on a regular lattice, might evoke predominantly systematic search strategies as the regularity of the lattice over which the items are placed serves as a guide to the eye. If indeed this is the case, this cohort should be characterized in the parameter space with large λ values and small $\delta\theta$. On the other hand, an experiment as WW, which is composed of a very complex scene with no apparent regularities, might evoke a random strategy, as there is no preferential searching direction. In the parameter space, this would be achieved with a small λ value and, in principle, any value of $\delta\theta$. Within this hypothesis, the CN experiment can be viewed as an intermediate configuration between the aforementioned experiments, as the items are positioned on a lattice but with random displacements.

We compute the parameters from the 5-2 and WW cohorts (see electronic supplementary material, figure S3) and validate the VSM to the experimental data (see electronic supplementary material, figure S4). The locations of all the trajectories composing the experimental cohorts in parameter

space are depicted in figure 4a. The large variability across the λ parameter appears within subjects as well as within trials performed by the same subject in a searching episode (see electronic supplementary material, figure S5). However, even though there is variability present in each of these visual tasks, it is possible to identify a mean strategy employed by each experimental cohort. The 5-2 experiment presents a median λ value of about 0.6, whereas for the WW experiment it becomes about 0.2, as shown in figure 4b. The values for $\delta\theta$ are always confined within $[0, \pi/4]$, as shown in figure 4c.

By combining the efficiency results from the VSM with the location of the experimental data in parameter space, we realize that the 5-2 cohort is the most efficient one. The other trajectories are relatively less efficient. This raises the question whether the efficiency analysed here, as well as the criteria reviewed in the literature, such as the overall time to find a target, is adequate when it comes to compare different visual search tasks.

Through the study of the inter-saccadic relative angle distribution, we were able to quantify the average strategies employed in three visual searching experiments and catalogue them according to their λ and $\delta\theta$ values. The average strategy changes along the different cohorts but always evoking persistent searches (small $\delta\theta$ values), figure 4c. One could argue that the discrepancy within the average strategy employed appears as a consequence of the amount of distractors present in the image. However, the 5-2 experiment was performed with different amounts of distractors and significant differences between the average λ values have not been observed (see electronic supplementary material, figure S6). This suggests that the difference between the average strategies appears as a consequence of the scene content and structure.

4. Discussion

Our work focuses on the study of ocular patterns that emerge during a visual search task. Through a simple model containing two parameters, we show that the knowledge of the saccadic relative orientation distribution gives the relevant information on the visual trajectories found in experiments. The validation of the model with experimental scanpaths through an indirect efficiency measurement confirms this fact. Moreover, the distribution of the experimental data points over the parameter space shows that indeed there is a large variability between the different visual strategies employed while searching, even for the same subject performing the same experiment. Nevertheless, we are able to discriminate changes in the average strategy across different scenes. This is shown through the analysis of the model parameters λ and $\delta\theta$, which suggests that the visual strategy employed is linked to the underlying structure of the scene, even when the cognitive task remains the same, i.e. visual search.

The experimental cohorts presented in this work correspond to participants performing visual searches in different scenarios. Even though these scenarios are easily distinguishable from one another, they have the following common properties: (i) the scenes are static, no moving items were analysed in these experiments; (ii) the scenes have no salient regions, therefore there are no *a priori* privileged regions to explore in the image; (iii) the task always involves finding a unique target hidden in a set of distractors;

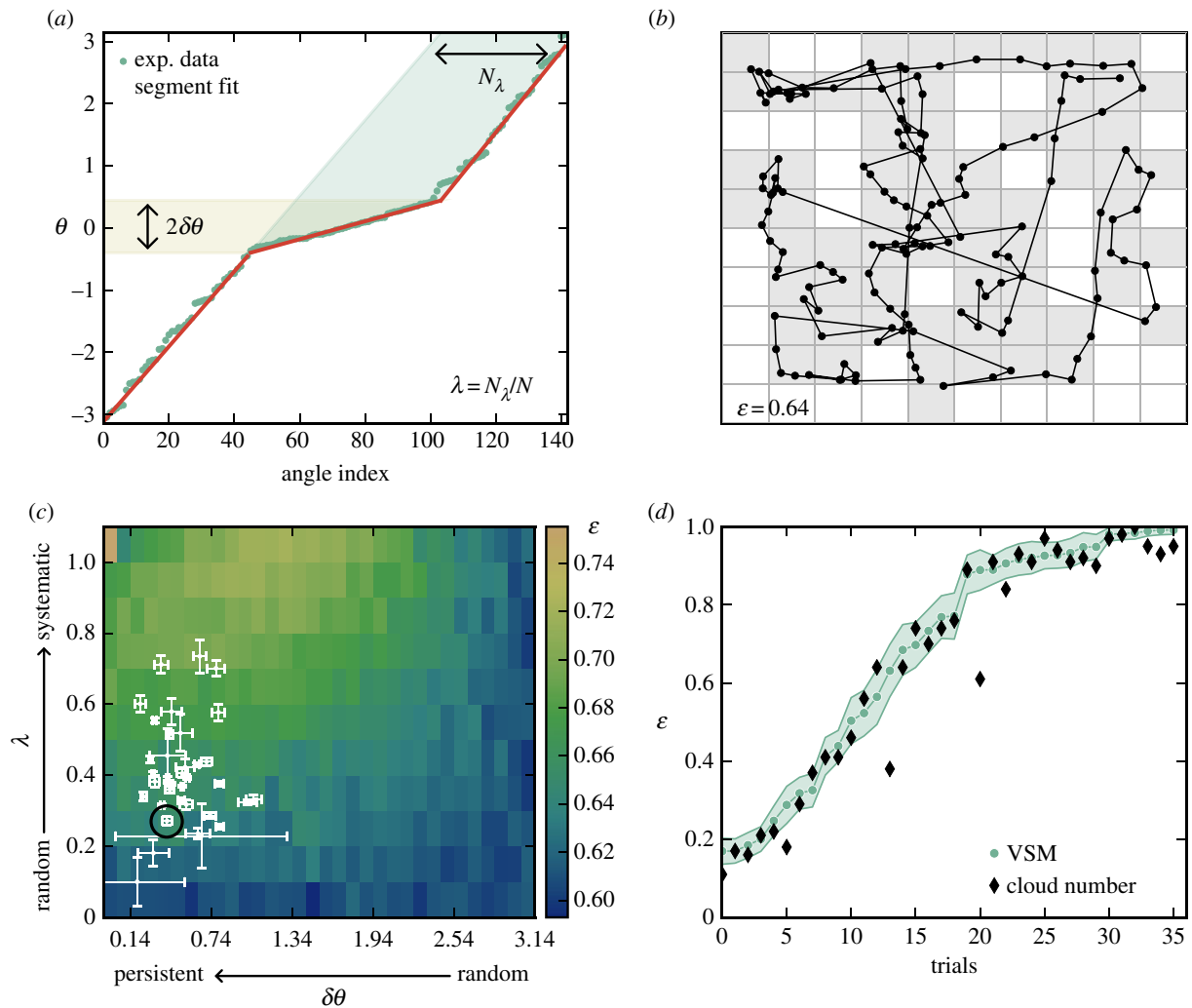


Figure 3. Comparison of the VSM with the experimental eye paths. (a) Estimation of λ and $\delta\theta$ from the experimental trajectories. We rank θ (circles) for each experimental trajectory and perform a fit of three linear segments (solid lines). The limits of the middle region (yellow shaded region) define $\delta\theta$. The difference between the extended line of the first region and the end of the third region (green shaded region) defines the amount of angles, N_λ , originated by the distribution restricted by $\delta\theta$. The parameter λ is defined as the ratio between N_λ and the total number of angles N . (b) Definition of space-filling efficiency, ϵ . ϵ corresponds to the percentage of visited locations (grey cells) after parcelling the image into $m \times m$ cells. In this case, $m = 10$. (c) Mean efficiency values obtained from the VSM parameters space for 100 realizations of the model. The experimental data from the CN cohort (white dots) are placed on their corresponding λ and $\delta\theta$ values. (d) Comparison between ϵ values obtained with the VSM and the experimental ones. For each trajectory (i.e. black points), we run the model with the same number of jumps and the estimated λ and $\delta\theta$ values. The shaded area corresponds to the standard deviation of 100 realizations of the model. The data presented in (a) and the data point highlighted in (c) correspond to the trajectory depicted in (b) with length approximately 140. (Online version in colour.)

(iv) the distractors have the same size as the target; and (v) all distractors are equal, in principle, there are no distractors less relevant than others. The layout of the items (from a regular lattice to a crowded space) and the features associated to the distractors are the main sources of differences between the experiments. In the 5-2 experiment, the items can either be a number 5 or a number 2. On the other hand, for the WW experiment, the items presented in the scene, although they always involve people, present a variety of shapes and colours that make them richer in visual content, thus making the scene more complex.

The difficulty of the task in which the participant is engaged while performing visual search has a strong effect on the number of items that can be processed during a fixation [34,35]. Following this idea, it has been suggested that what determines a visual task to be easy or difficult is linked to the discriminability of the target [36,37]. Accordingly, we can arrange our set of experiments in the following order of increasing difficulty: 5-2, CN and WW. In this way, our results indicate that a difficult task is more

likely to be executed in a random fashion with a relative degree of persistence, whereas an easy task evokes a more systematic search. An intermediate difficulty task, such as the CN experiment, leads to a mixed type of strategy within the aforementioned extremes.

Our work focuses on the study of images with no saliencies such that the target and distractors are, in principle, equally relevant to the participant while performing the task at hand. We believe that these types of scenes are ideal for the study of visual search tasks involving hidden targets. The presence of a few salient regions creates a bias in the scene so that they become the most probable locations to be fixated by a viewer [2]. If we were to think of a visual search in the presence of saliencies [38–40], the viewer would first search on the salient object for the presence of the target, as these are the most probable regions to be fixated. If the target corresponds to one of the salient images, then the search is over and the strategy could be studied by means of a saliency model, where each object is associated to a probability to be fixated. On the other hand, if the target does not correspond to a

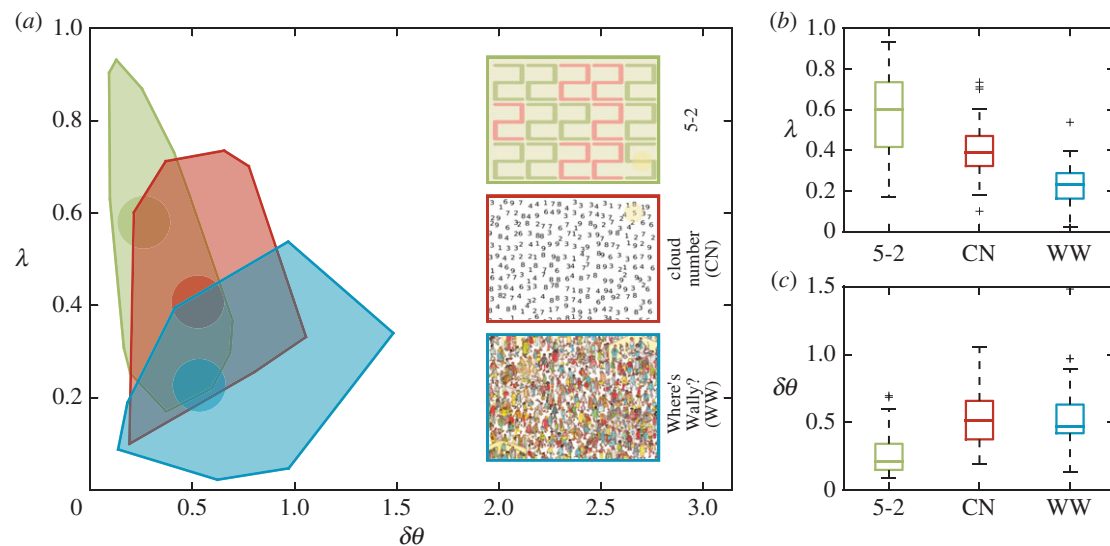


Figure 4. Different tasks are placed in distinctive positions in the parameter space. (a) Three visual search experimental cohorts, going from numbers placed in a regular lattice to a random position, exhibit distinctive λ values. The shaded area shows where the experimental data fall and the circle represents the centre of mass of each group. (b) The 5-2 experiment presents a median λ value 0.6 (green), the CN experiment, 0.4 (red) and the WW experiment, 0.2 (blue). (c) For the median value of $\delta\theta$, there is no clear distinction and all experiments are confined to the range $[0, \pi/4]$. (Online version in colour.)

salient object, we believe the participant would then ‘filter out’ these objects and proceed to do a similar type of search as the ones depicted in this work over the unsalient remaining objects, which could then be analysed with the VSM.

It has been the aim of several previous studies to elucidate how the eye movement relates to the structure of the scene [41–43]. For instance, this has been performed studying how the scanpaths change while free viewing abstract or realistic paintings [43]. However, the images presented in this study do not present a clear difference in the structure of the scene, rather a difference in the content. Our experiments clearly differ in structure; in the 5-2 experiment, the items are located in a regular lattice, whereas in the WW experiment, they are randomly distributed. Moreover, the studies presented in [41–43] are concerned with free viewing; therefore, it is not clear whether there is any strategy planning taking place at all. Interestingly, in [41], the authors claim that participants tend to follow a specific scan-path for a well-defined learned pattern. They conclude that this behaviour cannot be due to low-level features, such as the peripheral visual field, because this characteristic should be constant across subjects, resulting in the same visual scanpaths. However, different participants performed different visual patterns, similar to our experiments. By means of our model, we are able to discriminate between the different patterns across participants and find, at the same time, an average search strategy linked to the underlying scene. Furthermore, the authors in [41] suggest that the scanpaths could be a consequence of eye movement habits that are independent of the pattern viewed. We have found that different average strategies emerge when searching in different landscapes, which suggests that the strategy planning process could go beyond the habits of eye movement.

The idea of *inhibition of return* as a foraging facilitator has been already discussed in the literature [44–47]. In our model, it is achieved by sampling the inter-saccadic angles mainly from the restricted angle distribution. When $\lambda = 1$, we are in the presence of a systematic search where it is very unlikely to return to already explored regions of the scene. Here, we introduced ε which quantifies the efficiency of the

different strategies in terms of how space filling they are. Therefore, if a better exploration of the space increases the chances of finding the target, the presence of an inhibitory behaviour facilitates visual search. This is depicted by a larger ε for the systematic search in comparison to the random one. However, as we have found in our experiments that participants, depending on the scene in front of them, do perform average search strategies that differ from the optimal one, the question remains whether participants choose poorly the search strategy or if there is some other quantity that is being optimized. Moreover, research in the field has shown that visual search could be a memoryless phenomenon [48]. For the 5-2 experiment, we found that most participants tend to perform a systematic type of search. We believe this is mainly due to the strong regular structure of the image which guides the eye in this particular way. Hence, even though people use an inhibitory mechanism to search, from our results we cannot conclude directly whether the participants engage in a memory process, in the sense that they can recall the locations of items observed during the course of the search.

Our method is simple enough to be used as a tool to further understand or diagnose mental disorders. For instance, a recent study investigated how the visual patterns, while performing different free viewing tasks, change for healthy control versus schizophrenic patients [49]. The authors presented a set of scenes, with different levels of image complexity, for the participants to freely explore and recorded the position of their eyes while doing so. They found that both groups reduced the area of exploration as the images become less complex, while schizophrenic participants tend to maintain the same type of scanpath for all the scenes. Therefore, it would be interesting to measure the parameters λ and $\delta\theta$ over these trajectories and study how they change from one group to the other. Based on the reported trajectories in [49], we believe that the average λ should change for the healthy control group, whereas for the schizophrenic group it should remain constant along the different levels of image complexity. Within this simple framework, we would be able to quantify behaviours that have been qualitatively observed.

Our findings and methods can be also applied to other types of search such as way-finding, where similar patterns appear for people looking for a target in a confined or open space [50]. Furthermore, within the framework of foraging and random walks in general, we propose a very simple model that exhibits persistent behaviour similar to the effect of 'wind', namely, a directional bias in a random walk search [51]. Additionally, our model comes with the advantage of being able to quantify the type of search by means of two parameters which can be easily determined from experimental data.

In summary, we were able to use the parameter space of the model to correlate a particular scene with a distinct average strategy. With this new relationship between visual strategies and scene content, we can predict the average strategy applied in a new experiment by comparing the new scene structure with the ones discussed here. In contrast with the other models that study visual search in the absence of salencies, such as the saccadic targeting models [13–18], in our experiment, we are able to characterize in a very simple and quantitative way the wide range of different strategies adopted by various observers. We conclude that the intersaccadic angle distribution provides enough information to characterize the majority of visual search strategies and to reproduce the search efficiencies found experimentally. As a prospective work, it would be also interesting to extend our findings to search tasks that involve moving objects, or perhaps to general visual tasks where the sequence of fixations is relevant. Furthermore, given that the scenes are shown on a two-dimensional display with well-defined boundaries, it would be worth investigating whether the systematic search strategies remain once the boundaries are removed.

5. Material and methods

Both experimental cohorts were conducted in the Universidade Federal do Ceará with participants who had normal or corrected to normal vision. The experimental data corresponding to the CN cohort are published in [21] and those corresponding to the 5-2 and WW cohort, in [26]. The study has been approved by the Ethics Research Committee of the Universidade Federal do Ceará (COMEPE) under the protocol number 056/11. All methods used in this study were carried out in accordance with the approved guidelines and all experimental protocols were approved by COMEPE. Informed consent was obtained from all subjects.

5.1. 'Cloud number' cohort

Experimental data were recorded with an EyeLink 1000 system (SR Research Ltd, Mississauga, Canada), with an acquisition frequency of 1 kHz on a monocular recording over ten subjects (mean age: 24). Each subject carried out a sequence of four trials, each one with a maximum duration of five minutes. In each trial, we presented the participant an image with numbers randomly distributed in a 1024×1280 px image, where the goal was to find a unique number '5' within 1499 distractors.

Between each trial, the subject had the possibility of relaxing and before starting the recording we performed a new calibration. At the beginning of each trial, the participant was asked to fixate his/her eyes on the centre of the screen, in case a drift correction needed to be performed.

The classification of the fixation and saccades was made using the EyeLink online filter [52, Section 4.3]. Fixations in the EyeLink system are identified using a saccade-pick algorithm. The system analyses the moment-to-moment velocity and acceleration of the eye using fixed thresholds for both eye velocity and acceleration. If the eye goes above either the velocity or acceleration threshold, the start of a saccade is marked. Analogously, when both the velocity and the acceleration drop back below their thresholds, the algorithm identifies the saccade end. By default, every movement which does not lie within this definition is considered as being part of a fixation. The saccade velocity threshold was set to be 30° s^{-1} , the saccade acceleration threshold, $8000^\circ \text{ s}^{-2}$ and the saccade motion threshold, 0.15° .

5.2. '5-2' and 'Where's Wally?' cohorts

Eye movements were recorded with a Tobii T120 eye-tracking system (Tobii Technology), over 11 subjects (mean age: 23). The stimuli were presented on a 17 TFT-LCD monitor with a resolution of 1024×1280 px and an acquisition frequency of 60 Hz. The 5-2 experiment consists in finding a number '5' within an array of numbers '2' serving as distractors. All numbers (target and distractors) are positioned on a square lattice and are randomly coloured red or green, hindering the visual detection of the target through the identification of patterns on the peripheral vision. The number of distractors present in each task is related to a degree of difficulty: DF0 (207 distractors), DF1 (857 distractors) and DF2 (1399 distractors). The maximum time given to search the target was 1 min, 1.5 min and 2 min, respectively.

The WW experiment consists in finding Wally, the famous character from the series of books with the same name [33], who is hidden within a very complex background of crowded characters, with a maximum searching time of 2 min.

As explained in [26], the identification of fixations and saccades was carried out with a modified version of the fixation filter developed by Olsson [53].

For each experimental cohort, namely CN, 5-2 and WW, we fitted the experimental saccade length distribution and used the obtained parameters, accordingly, to generate the simulated scanpaths in our model.

Data accessibility. Experimental data are available at <https://figshare.com/s/9e2b1d3545dd04e44d94>.

Author's contributions. T.A.A., M.L., H.J.H. and J.S.A. designed research; T.A.A. performed the experiments; T.A.A. and M.L. analysed data; T.A.A., M.L., H.J.H. and J.S.A. performed research; T.A.A., M.L., H.J.H. and J.S.A. wrote the paper.

Competing interests. The authors declare no conflict of interest.

Funding. We acknowledge financial support from the European Research Council (ERC) Advanced grant no. 319968-FlowCCS. This work was supported by the Brazilian agencies CNPq, CAPES, FUNCAP and National Institute of Science and Technology for Complex Systems in Brazil.

Acknowledgments. The authors thank H.F. Credidio for making available the experimental data corresponding to the 5-2 and WW cohorts.

References

1. Eckstein MP. 2011 Visual search: a retrospective. *J. Vis.* **11**, 14. (doi:10.1167/11.5.14)
2. Itti L, Koch C, Niebur E. 1998 A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 1254–1259. (doi:10.1109/34.730558)

3. Itti L, Koch C. 2000 A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* **40**, 1489–1506. (doi:10.1016/S0042-6989(99)00163-7)
4. Over E, Hooge I, Vlaskamp B, Erkelens C. 2007 Coarse-to-fine eye movement strategy in visual search. *Vis. Res.* **47**, 2272–2280. (doi:10.1016/j.visres.2007.05.002)
5. Foulsham T, Underwood G. 2008 What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *J. Vis.* **8**, 6. (doi:10.1167/8.2.6)
6. Ehinger KA, Hidalgo-Sotelo B, Torralba A, Oliva A. 2009 Modelling search for people in 900 scenes: a combined source model of eye guidance. *Vis. Cogn.* **17**, 945–978. (doi:10.1080/13506280902834720)
7. Nakayama K, Martini P. 2011 Situating visual search. *Vis. Res.* **51**, 1526–1537. (doi:10.1016/j.visres.2010.09.003)
8. Li Z. 2002 A saliency map in primary visual cortex. *Trends Cogn. Sci.* **6**, 9–16. (doi:10.1016/S1364-6613(00)01817-9)
9. Shiffrin RM, Gardner GT. 1972 Visual processing capacity and attentional control. *J. Exp. Psychol.* **93**, 72. (doi:10.1037/h0032453)
10. Zhang L, Tong MH, Marks TK, Shan H, Cottrell GW. 2008 SUN: a Bayesian framework for saliency using natural statistics. *J. Vis.* **8**, 32. (doi:10.1167/8.7.32)
11. Bruce ND, Tsotsos JK. 2009 Saliency, attention, and visual search: an information theoretic approach. *J. Vis.* **9**, 5. (doi:10.1167/9.3.5)
12. Gao D, Han S, Vasconcelos N. 2009 Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**, 989–1005. (doi:10.1109/tpami.2009.27)
13. Rao RP, Zelinsky GJ, Hayhoe MM, Ballard DH. 2002 Eye movements in iconic visual search. *Vis. Res.* **42**, 1447–1463. (doi:10.1016/S0042-6989(02)00040-8)
14. Beutner BR, Eckstein MP, Stone LS. 2003 Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *JOSA A* **20**, 1341–1355. (doi:10.1364/josaa.20.001341)
15. Eckstein MP, Drescher BA, Shimozaaki SS. 2006 Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychol. Sci.* **17**, 973–980. (doi:10.1111/j.1467-9280.2006.01815.x)
16. Pomplun M. 2006 Saccadic selectivity in complex visual search displays. *Vis. Res.* **46**, 1886–1900. (doi:10.1016/j.visres.2005.12.003)
17. Zelinsky GJ. 2008 A theory of eye movements during target acquisition. *Psychol. Rev.* **115**, 787. (doi:10.1037/a0013118)
18. Najemnik J, Geisler WS. 2005 Optimal eye movement strategies in visual search. *Nature* **434**, 387–391. (doi:10.1038/nature03390)
19. Najemnik J, Geisler WS. 2008 Eye movement statistics in humans are consistent with an optimal search strategy. *J. Vis.* **8**, 4. (doi:10.1167/8.3.4)
20. Araujo C, Kowler E, Pavel M. 2001 Eye movements during visual search: the costs of choosing the optimal path. *Vis. Res.* **41**, 3613–3625. (doi:10.1016/S0042-6989(01)00196-1)
21. Amor TA, Reis SD, Campos D, Herrmann HJ, Andrade JS. 2016 Persistence in eye movement during visual search. *Sci. Rep.* **6**, 20815. (doi:10.1038/srep20815)
22. Engbert R, Mergenthaler K, Sinn P, Piskovsky A. 2011 An integrated model of fixational eye movements and microsaccades. *Proc. Natl Acad. Sci. USA* **108**, E765–E770. (doi:10.1073/pnas.1102730108)
23. Marlow CA, Viskontas IV, Matlin A, Boydston C, Boxer A, Taylor RP. 2015 Temporal structure of human gaze dynamics is invariant during free viewing. *PLoS ONE* **10**, e0139379. (doi:10.1371/journal.pone.0139379)
24. Codling EA, Plank MJ, Benhamou S. 2008 Random walk models in biology. *J. R. Soc. Interface* **5**, 813–834. (doi:10.1098/rsif.2008.0014)
25. Tejedor V, Voituriez R, Bénichou O. 2012 Optimizing persistent random searches. *Phys. Rev. Lett.* **108**, 088103. (doi:10.1103/physrevlett.108.088103)
26. Credidio HF, Teixeira EN, Reis SD, Moreira AA, Andrade JS. 2012 Statistical patterns of visual search for hidden objects. *Sci. Rep.* **2**, 920. (doi:10.1038/srep00920)
27. Le Meur O, Coutrot A. 2016 Introducing context-dependent and spatially-variant viewing biases in saccadic models. *Vis. Res.* **121**, 72–84. (doi:10.1016/j.visres.2016.01.005)
28. Motter BC, Belky EJ. 1998 The guidance of eye movements during active visual search. *Vis. Res.* **38**, 1805–1815. (doi:10.1016/S0042-6989(97)00349-0)
29. Brockmann D, Geisel T. 2000 The ecology of gaze shifts. *Neurocomputing* **32**, 643–650. (doi:10.1016/S0925-2312(00)00227-7)
30. Boccignone G, Ferraro M. 2004 Modelling gaze shift as a constrained random walk. *Phys. A Stat. Mech. Appl.* **331**, 207–218. (doi:10.1016/j.physa.2003.09.011)
31. Engbert R, Mergenthaler K. 2006 Microsaccades are triggered by low retinal image slip. *Proc. Natl Acad. Sci. USA* **103**, 7192–7197. (doi:10.1073/pnas.0509557103)
32. Hubel DH, Wensveen J, Wick B. 1995 *Eye, brain, and vision*. New York, NY: Scientific American Library.
33. Handford M. 1997 *Where's Wally? The wonder book*. London, UK: Walker Books.
34. Motter BC, Simoni DA. 2008 Changes in the functional visual field during search with and without eye movements. *Vis. Res.* **48**, 2382–2393. (doi:10.1016/j.visres.2008.07.020)
35. Geisler WS, Chou KL. 1995 Separation of low-level and high-level factors in complex tasks: visual search. *Psychol. Rev.* **102**, 356. (doi:10.1037/0033-295X.102.2.356)
36. Young AH, Hulleman J. 2013 Eye movements reveal how task difficulty moulds visual search. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 168. (doi:10.1037/a0028679)
37. Hulleman J, Olivers CN. 2014 Search through complex motion displays does not break down under spatial memory load. *Psychon. Bull. Rev.* **21**, 652–658. (doi:10.3758/s13423-013-0537-6)
38. Henderson JM, Weeks Jr PA, Hollingworth A. 1999 The effects of semantic consistency on eye movements during complex scene viewing. *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 210. (doi:10.1037/0096-1523.25.1.210)
39. Wolfe JM, Horowitz TS. 2004 What attributes guide the deployment of visual attention and how do they do it? *Nat. Rev. Neurosci.* **5**, 495–501. (doi:10.1038/nrn1411)
40. Wolfe JM, Vo MLH, Evans KK, Greene MR. 2011 Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* **15**, 77–84. (doi:10.1016/j.tics.2010.12.001)
41. Noton D, Stark L. 1971 Scanpaths in eye movements during pattern perception. *Science* **171**, 308–311. (doi:10.1126/science.171.3968.308)
42. Noton D, Stark L. 1971 Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vis. Res.* **11**, 929–942. (doi:10.1016/0042-6989(71)90213-6)
43. Zangemeister WH, Sherman K, Stark L. 1995 Evidence for a global scanpath strategy in viewing abstract compared with realistic images. *Neuropsychologia* **33**, 1009–1025. (doi:10.1016/0028-3932(95)00014-T)
44. Klein RM. 1988 Inhibitory tagging system facilitates visual search. *Nature* **334**, 430–431. (doi:10.1038/334430a0)
45. Klein RM, MacInnes WJ. 1999 Inhibition of return is a foraging facilitator in visual search. *Psychol. Sci.* **10**, 346–352. (doi:10.1111/1467-9280.00166)
46. Klein RM. 2000 Inhibition of return. *Trends Cogn. Sci.* **4**, 138–147. (doi:10.1016/S1364-6613(00)01452-2)
47. Hooge ITC, Over EA, van Wezel RJ, Frens MA. 2005 Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vis. Res.* **45**, 1901–1908. (doi:10.1016/j.visres.2005.01.030)
48. Horowitz TS, Wolfe JM. 1998 Visual search has no memory. *Nature* **394**, 575–577. (doi:10.1038/29068)
49. Egaña JJ, Devia C, Mayol R, Parrini J, Orellana G, Ruiz A, Maldonado PE. 2013 Small saccades and image complexity during free viewing of natural images in schizophrenia. *Front. Psychiatry* **4**, 37. (doi:10.3389/fpsy.2013.00037)
50. Pingel TJ, Schinazi VR. 2014 The relationship between scale and strategy in search-based wayfinding. *Cartograph. Perspect.* **33**–45. (doi:10.14714/cp.77.1232)
51. Palyulin VV, Chechkin AV, Metzler R. 2014 Lévy flights do not always optimize random blind search for sparse targets. *Proc. Natl Acad. Sci. USA* **111**, 2931–2936. (doi:10.1073/pnas.1320424111)
52. SR Research. 2010 *EyeLink 1000 user manual*. Mississauga, Canada: SR Research Ltd.
53. Ollson P. 2007 Real-time and offline filters for eye tracking. Master's degree project, KTH Electrical Engineering, Stockholm, Sweden.