The role of eye movements in a contour detection task

Nathalie Van Humbeeck Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

Nadine Schmitt Department of Neurophysics, Institute for Theoretical Physics, University of Bremen, Bremen, Germany

Frouke Hermens School of Psychology, University of Aberdeen, Aberdeen, Scotland, UK

Johan Wagemans Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

Udo A. Ernst Department of Neurophysics, Institute for Theoretical Physics, University of Bremen, Bremen, Germany

Vision combines local feature integration with active viewing processes, such as eye movements, to perceive complex visual scenes. However, it is still unclear how these processes interact and support each other. Here, we investigated how the dynamics of saccadic eye movements interact with contour integration, focusing on situations in which contours are difficult to find or even absent. We recorded observers' eye movements while they searched for a contour embedded in a background of randomly oriented elements. Task difficulty was manipulated by varying the contour's path angle. An association field model of contour integration was employed to predict potential saccade targets by identifying stimulus locations with high contour salience. We found that the number and duration of fixations increased with the increasing path angle of the contour. In addition, fixation duration increased over the course of a trial, and the time course of saccade amplitude depended on the percept of observers. Model fitting revealed that saccades fully compensate for the reduced saliency of peripheral contour targets. Importantly, our model predicted fixation locations to a considerable degree, indicating that observers fixated collinear elements. These results show that contour integration actively guides eye movements and determines their spatial and temporal parameters.

Introduction

Images that enter our eyes are first processed by neurons sensitive to only a small region of visual space and certain features (e.g., orientation, color). Crucial to our understanding of visual perception is the process by which the visual system groups these local neural responses to signal the presence of spatially coherent structures (for a broad review, see Wagemans et al., 2012). A powerful tool to achieve a better understanding of visual grouping is the contour detection paradigm of Field, Hayes, and Hess (1993). In this paradigm, a new type of stimulus was proposed to investigate the grouping of individual elements into global contours. In a typical contour detection experiment, observers are asked to detect the presence of a contour composed of spatially separate Gabor elements embedded in a background of similar but randomly oriented elements. The spatial separation of the distractor elements matches the separation of the contour elements, and therefore, the contour can only emerge from the background due to the collinearly arranged contour elements forming a smooth path (for a recent review, see Hess, May, & Dumoulin, 2013).

The ability of human observers to detect a contour appears best when the contour forms a straight line and decreases as the degree of curvature of the contour, defined by the angle between the successive elements, increases (Field et al., 1993; Watt, Ledgeway, & Dakin,

Citation: Van Humbeeck, N., Schmitt, N., Hermens, F., Wagemans, J., & Ernst, U. A. (2013). The role of eye movements in a contour detection task. *Journal of Vision, 13*(14):5, 1–19, http://www.journalofvision.org/content/13/14/5, doi:10.1167/13.14.5.

doi: 10.1167/13.14.5 Received May 22, 2013; published December 4, 2013 ISSN 1534-7362

2008). To account for this finding, Field et al. proposed that oriented elements interact with neighboring elements through a local "association field." Typically, an association field indicates how strongly an element in the image can be collinearly grouped with its neighboring elements by assigning an association strength to each element in the field (May & Hess, 2007; Watt et al., 2008). The association strength of an element is high when the element is collinearly aligned with nearby elements and low when the distance, curvature, and misalignment from cocircularity between the element and its neighboring elements increases. This concept has been implemented in neural interaction models, involving facilitatory connections between nearly collinear elements and inhibitory interactions between parallel and orthogonal elements (Li, 1998; Mundhenk & Itti, 2002; Yen & Finkel, 1998). Recently, Ernst et al. (2012) developed an association field model that captures human contour detection behavior to a previously unprecedented degree. This model has a well-defined probabilistic interpretation (Williams & Thornber, 2001) and uses an approach that reduces contour integration to an optimal inference problem. More specifically, it computes for each edge element in a stimulus the probability that this edge is part of a contour with certain statistical properties, thereby generating a probability map for contours of increasing length as processing time proceeds.

Studies making use of the contour detection paradigm typically limit the presentation time of the contour stimulus to reduce the influence of eye movements during the task. However, these short presentation times come at a cost. For example, Field et al. (1993) found that observers' detection performance became worse when the presentation time of the contour stimulus was reduced from 1 s to 0.25 s. The absence of eye movements may play an important role in this reduction. Given that the contour path did not always fall in the foveal region of the stimuli, preventing observers from making eye movements could have severely constrained contour detection. The role of eye movements in contour detection is further supported by findings showing that contour detection performance is highly dependent on the peripheral position of the contour (Hess & Dakin, 1997; May & Hess, 2007; Nugent, Keswani, Woods, & Peli, 2003). To account for such effects, May and Hess (2007) proposed a type of association field model in which the association fields are small in the foveal region of the stimulus and large in the periphery. Such a model predicts that elements of highly curved contours located far from the fixation point are more likely to be linked with distractor elements due to larger association fields in the periphery linking elements over larger distances. Consequently, contours located in the periphery are less detectable, especially when they are curved. Eye movements can considerably support this contour-integration process, allowing observers to foveate image elements that may be part of a contour. More specifically, we hypothesized that observers fixate regions of the image that are characterized by high association strengths. Under this assumption, measurements of fixation locations should reveal the internal association field used by observers during contour integration, providing a powerful tool to test models of contour integration that predict a specific association field. We here describe two experiments aimed to test this hypothesis by applying the association field model of Ernst et al. (2012) to examine whether the model's predictions of salient collinear locations in an image match the fixation locations in a contour detection task.

Information processing during contour integration may also be reflected in other spatial and temporal aspects of eye movements. Fixation duration and saccade amplitude (i.e., the distance covered by a saccadic eye movement) have been proposed to indicate distinct modes of global and local processing during image viewing (Pannasch, Helmert, Roth, Herbold, & Walter, 2008; Unema, Pannasch, Joos, & Velichkovsky, 2005; Velichkovsky, Joos, Helmert, & Pannasch, 2005). For instance, in a visual-search study, Over, Hooge, Vlaskamp, and Erkelens (2007) observed an increase in fixation duration and a decrease in saccade amplitude with viewing time. They hypothesized that this time course is the result of a search strategy of the visual system to gradually move from a global search to a more local search at finer spatial scales. The strategy did not depend on whether the exact appearance of the target and background was known in advance or not, suggesting that it reflected a general oculomotor strategy. However, a scene-perception study by Mills, Hollingworth, Van der Stigchel, Hoffman, and Dodd (2011) showed that certain task requirements influence these oculomotor parameters differently, indicating that they are controlled by independent mechanisms instead of one intrinsic coarse-to-fine process (Pannasch et al., 2008; Unema et al., 2005). In this study, it was found that fixation duration gradually increased toward an optimal level of local processing, and saccade amplitude remained relatively stable over time. According to the authors, the size of saccades during a trial depended mainly on whether it was needed to acquire visual information across the entire scene. For instance, saccade amplitude remained high over the course of a trial in a search task in which the target was extremely difficult to detect. One possibility is that the time courses of fixation duration and saccade amplitude, instead of reflecting a built-in coarse-to-fine mechanism, can be strategically adjusted to the difficulty of the task. For instance, a study by Vlaskamp, Over, and Hooge (2005) suggests that

fixation duration and saccade amplitude are closely related to the difficulty of finding a target during visual search. These authors found that fixation duration increased and saccade amplitude decreased with increasing target-distractor similarity (i.e., decreased target saliency), which reflects a more local processing strategy.

In sum, research on contour integration suggests that eye movements can substantially support contour detection, presumably by foveating likely contour candidates in the image. Moreover, certain saccadic characteristics, such as fixation duration and saccade amplitude, seem to be influenced, to a certain degree, by the saliency of a target during search and by the time course of image processing. The purpose of this study was to examine the involvement of eye movements in the process of contour integration, especially in situations in which contours are difficult to find. We investigated how durations of individual fixations and saccade amplitudes vary as a function of time and difficulty of the contour-integration task. In addition, we tested whether an association field model, proposed recently to explain contour integration (see Ernst et al., 2012), can predict fixation locations. Our hypothesis was that subsequent fixations are preferentially directed toward "hot spots" of association strength predicted by an association field. Participants were asked to find a contour hidden in a dense arrangement of oriented Gabor patches and to perform two tasks while their eye movements were tracked. These included indicating whether the contour, which was always presented, was on the left or right of the visual field or to determine whether a contour was present in the Gabor field. Task difficulty was manipulated by varying the path angle of the contour, which influenced the saliency of the contour elements.

Methods

Participants

Twenty-one observers, with normal or corrected-to-normal vision (age range 17–35), participated in each task of the experiment. Two were authors, and the others were psychology students at the University of Leuven. All gave their written informed consent before participating in the experiment, which was approved by the local ethics committee.

Apparatus

The stimuli were displayed on a 22-in. CRT monitor (Iiyama HM204DT A) at a refresh rate of 75 Hz. A Pentium PC (NVIDIA GeForce 7600 GT graphics card) controlled the presentation of the stimuli while a second PC recorded the eye-movement data. Eye movements were measured using the Eyelink II video-based eye tracker (SR Research, Osgoode, ON, Canada), which uses two small cameras mounted to a headband worn by the participants. The Eyelink II has a reported average accuracy of less than 0.5°, which is below the average distance of the elements in our display (i.e., 0.7°), allowing the accurate assignment of fixations to the individual elements in the display. To ensure that participants remained at a fixed distance from the screen and to avoid head movements, which could possibly result in drift in the recorded eye positions due to headband slippage, a chin rest was used, positioned at 60 cm from the CRT screen. The eye tracker recorded the horizontal and vertical eye positions for both eyes at a rate of 250 Hz. When possible, the combined pupil-plus-corneal-reflection mode was used. For three participants, this mode resulted in clear distortions in the recorded eye position because of reflections of the IR illumination to glasses or contact lenses. For these three participants, the pupil-only mode was used instead. Participants' responses were registered by means of a computer keyboard.

Stimuli

The target stimulus (590 × 590 pixels; 23°) consisted of Gabor patches that were in cosine phase with wavelength λ (0.20°), Gaussian envelope of SD $\lambda/2$ (0.10°), and Michelson contrast of 50% (see also Watt et al., 2008). An example of a stimulus can be found in Figure 1, which shows the stimulus sequence during a trial. Each stimulus with a contour target contained a single embedded contour placed in a background of randomly oriented Gabor elements. In the left-right decision task (LR task), Gabor paths were required to be at a distance of at least 50 pixels (2°) from the center of the screen so that they were clearly located either in the left or right half of the Gabor field. The contours in the contour-present conditions of the present-absent task (PA task) could appear in any position of the display. Each contour was defined by a sequence of seven Gabor patches whose orientations were set to an underlying contour "spine" (sequences often referred to as "snakes"; see inset of Figure 1). The positions of each of the Gabors in the path were chosen such that the mean distance between adjacent elements in the contour was the same as the mean distance between adjacent elements in the background. In order to manipulate the saliency of the contour and therefore task difficulty, the path angle of the contour was systematically varied (May & Hess, 2008; Watt et al., 2008). The path angle is defined as the angle between successive elements in the

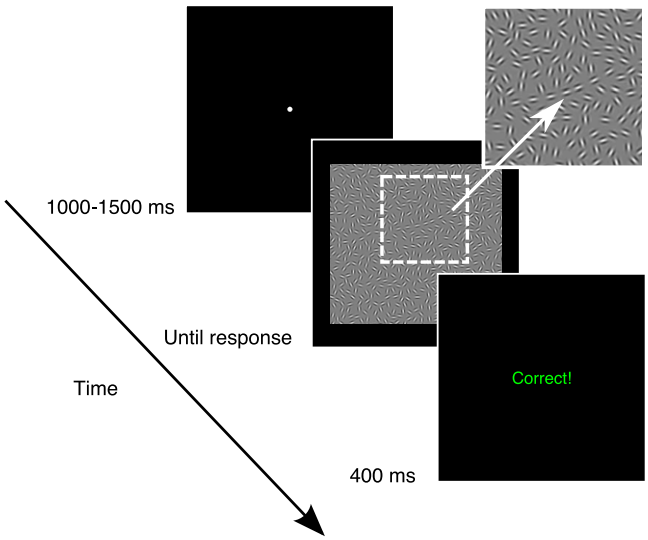

Figure 1. Illustration of the stimulus sequence. A fixation symbol was presented for a randomly selected duration between 1000 and 1500 ms. The target array was then presented until participants pressed a key to indicate whether a salient contour was located in the left or right half of the stimulus field (LR task) or whether it was present or absent (PA task). Feedback was then provided for 400 ms. The inset shows a close-up of the contour path consisting of seven collinear elements.

contour. For example, in the case of three elements, the path angle is defined by the difference in orientation between the lines formed by linking element 1 to element 2 and element 2 to element 3. Four different levels of path angle (0°, 5°, 10°, or 20°) were used. The sign of the path angle could vary randomly within the contour. However, the orientation of the Gabor in a path was always aligned with the path angle (i.e., no additional jitter was used). As indicated earlier, the remainder of the Gabor field was filled with randomly oriented elements placed at random positions in such a way that the mean distance between the center of each element and the center of any other neighboring element was 0.7°. Stimuli in the contour-absent condition of the PA task did not contain a salient contour and consisted only of randomly placed Gabor elements.

Procedure

The experiment started with instructing the participants about the two tasks and determining their dominant eye by asking them to look through a cylindrical object. In the first task, observers were instructed to indicate whether the contour was embedded left or right in the stimulus field (LR task). In the second task, they were asked to indicate whether the contour was present or not (PA task). In both tasks, either response was required in equal proportions. Participants responded by pressing either the letter "q" ("left" or "absent") or the letter "p" ("right" or "present") on a standard QWERTY keyboard. Before the experiment and at regular times during the experiment (after each block of 40 trials, if required), a calibration procedure for the eye tracker was performed. Calibration was repeated until all recorded fixations were aligned on a three by three grid for both eyes and the recordings of the first and the last fixation were at close proximity. Each trial started with a fixation dot, which was presented until the experimenter pressed the space bar to correct for drifts in the eye movement recordings due to small movements of the head. A second, smaller fixation dot followed the drift correction, which was presented for a random duration between 1000 and 1500 ms. Participants were then presented with the stimulus image until they pressed one of two keys on the computer keyboard to indicate their response. During stimulus presentation, they were free to shift their gaze across the display. After each trial, feedback was provided for 400 ms, indicating whether the participant's response was correct or not (see Figure 1 for an example of the stimulus sequence during a trial). After each 40 trials, participants received a message indicating how many trials were remaining. For the LR task, 60 different Gabor fields for each of the four path angles were used, resulting in 240 stimuli per participant. For the PA task, 80 random Gabor fields were used in which no salient contour was present and 20 stimulus fields were used for each of the four path angles, resulting in 160 stimuli per participant. All the participants received the same Gabor fields, so that eye traces for each of the stimuli could be directly compared across the participants. The order of the trials was randomized for each participant.

Eye-movement analysis

The default settings of the Eyelink II eye tracker were used to detect saccades, based on a velocity threshold of 30°/s and an acceleration threshold of 8000°/s^2. Although eye movement recordings were obtained for both eyes, we chose to focus on the recordings for the dominant eye (as determined with the cylindrical object–viewing task) in the analysis. Fixations shorter than 50 ms and longer than 2000 ms, as well as saccades with amplitudes smaller than 0.1°, were excluded from data analysis. The saccade before and the saccade after an excluded fixation were merged to form a new saccade, and the same method was used for fixations before and after removed saccades. The first fixation in each trial was discarded because it results from the preceding fixation dot at the center of the screen used for drift correction. Across all trials of this study, average fixation duration was 275 ms ($SD =$

133) and average saccade amplitude was 5.41° (SD = 4.56).

Statistical analyses were performed on the individual eye movements, which were in total 37,352 in the LR task and 32,193 in the PA task. Eye movements were nested within images and within observers for both tasks. The images were crossed with individuals because each observer saw the same set of Gabor field stimuli. Given that our data set has a hierarchical structure, multilevel models were used in which images and observers were treated as crossed random factors (Hoffman & Rovine, 2007; Locker, Hoffman, & Bovaird, 2007; see Mills et al., 2011 for a similar analysis on eye-movement data). An important advantage of multilevel analysis is that it allows for an unbalanced design as is the case for our data set, in which there are differences in the timing and number of eye movements across observers and trials. Models were estimated using maximum likelihood via the MIXED procedure of SPSS (syntax used to estimate the models is available from the authors upon request). Although the methods are relatively robust against violations of the normality assumption of the underlying distribution, we applied base 10 logarithmic transformations to our data to obtain normal data distributions.

To examine how fixation duration and saccade amplitude changed over the course of a trial, we tested a quadratic model of change over time for these two parameters as exploratory analyses of our data showed a quadratic change over time in fixation duration and saccade amplitude. The time for each trial was first normalized to the interval starting at the time when the first saccade was initiated in the center of the screen and ending when the last saccade ended. This interval was divided into 10 subintervals of equal width, and fixations were sorted into these subintervals according to their normalized starting time. Statistical tests were performed separately for the LR task and the PA task. We included two between-trial predictors. The first predictor was the saliency of a contour. Conditions included the different path angle conditions in both tasks and the contour-absent condition in the PA task. The second predictor was the correctness of the given manual response. In addition, we tested whether these predictors interacted with the linear and quadratic effect of time.

Architecture of contour-integration model

We compared our experiments with predictions from the association field model by Ernst et al. (2012). This contour-integration model was realized by a two-dimensional layer of $i = 1, \ldots, N$ neuronal populations, which were recurrently connected by synaptic weights w_{ij}. For reducing computational complexity, only populations i were used whose receptive fields were centered on the edge elements in a particular stimulus and which had the same preferred orientations.

In order to integrate neighboring edges, the synaptic weights were chosen such that a stronger interaction between nearby element pairs i and j results when these are arranged in a collinear and cocircular fashion, thus realizing an association field. The weights were sampled from a function $A(r, \alpha, \beta)$ defined as

$$A(r,\alpha,\beta) \propto \exp(-r/\lambda)\cosh\Big(1/\sigma_\alpha^2\cos(\beta/2-\alpha) + 4/\sigma_\beta^2\cos(\beta/2)\Big). \quad (1)$$

The radial part of this function decays exponentially with the length constant λ depending on the distance r between two edge elements. The angular part depends on the angles α and β, with β referring to the angle between the orientations of two edge elements and α denoting the angle of the destination edge i from the collinear continuation of the origin edge j. σ_α and σ_β are the scaling constants that determine the form of the association field (panel A of Figure 2). By setting both parameters to small values, contours with quite straight curvatures are favored during integration. By setting the parameters to higher values, contours with larger curvature can also be integrated but with the penalty of making the model less robust against noise. Panel B of Figure 2 shows a typical example of an association field with the parameters chosen for our simulations. Note that the association field is not symmetric and extends in a particular direction. Unidirectionality of the association field turned out to be beneficial for contour integration (Ernst et al., 2012), but it requires that each edge element in a stimulus be represented by two populations i and i' with association fields extending in opposite directions.

To account for reduced saliencies of edge elements i in peripheral vision (Foley, Varadharajan, Koh, & Farias, 2007), an additional scaling factor b_i for the afferent input to each population i was introduced. In the contour detection study of Ernst et al. (2012), it was found that b has to decrease with distance E from the fixation spot for the model to be able to successfully explain human contour integration. The required functional dependency of b on eccentricity E was found to be compatible with earlier psychophysical studies (e.g., Foley et al., 2007) and is quantified by

$$b(E,\mu,v) = 1 + 2\mu\left(\left(\frac{E}{E_{max}}\right)^v - \frac{1}{2}\right). \quad (2)$$

Here, μ describes how strongly b varies with eccentricity, and v regulates how steeply b varies with eccentricity. For $v < 1$, b is concave down, and for $v > 1$, b is concave up. Panel C of Figure 2 illustrates an example for such eccentricity scaling by modulating the

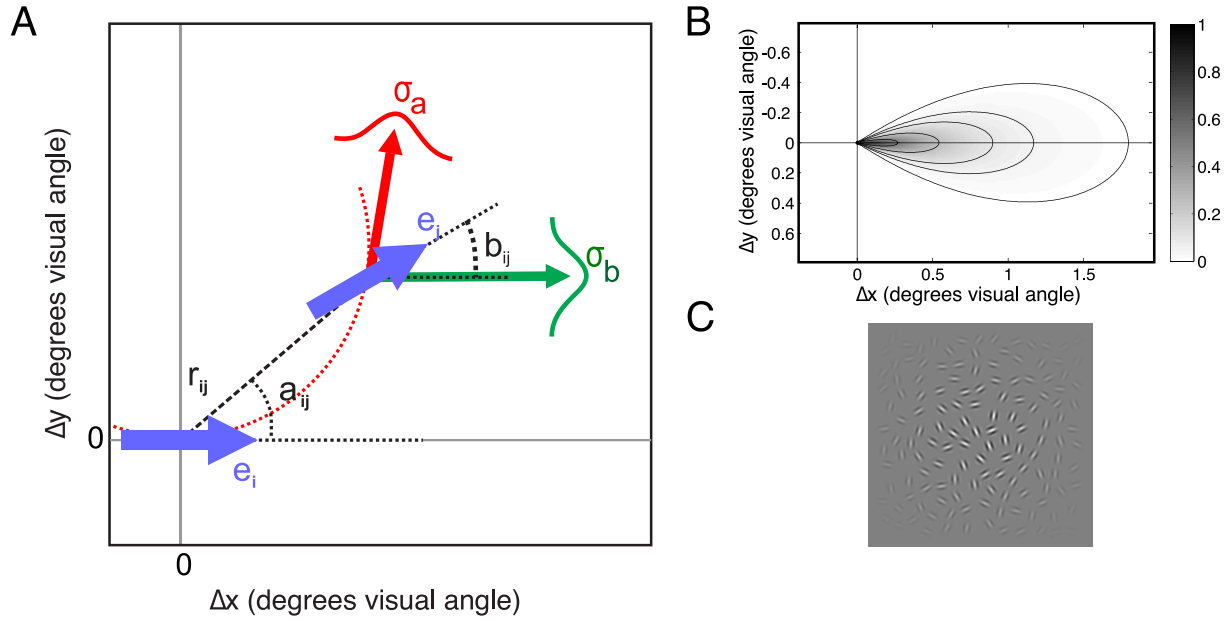

Figure 2. Parameters and shape of the association field and eccentricity scaling. (A) The geometrical relationship between edge elements e_i and e_j is defined by the distance r_{ij} between the elements and the angles α_{ij} and β_{ij}. α_{ij} refers to the angle of edge e_j from the collinear continuation of edge e_i, and β_{ij} is the angle between the orientation of the two edges. The red arrow specifies the direction e_j should have for perfect cocircularity between e_i and e_j. The association field decreases on a length scale σ_α with increasing deviation of the orientation of e_j from this direction. The green arrow depicts the case in which the two edge elements would have the same orientation. The association field decreases on a length scale σ_β with increasing deviation of the orientation of e_j from this direction. Figure adapted from Ernst et al. (2012). (B) The association field, defined as the product of an angular part and a radial part (see Methods section), for an element positioned at the origin and with an orientation equal to zero. It is defined by parameter values fitted to reproduce human contour detection behavior in our LR task. Greyscale values specify the association strength with darker shades indicating locations that give rise to higher association strengths. (C) The association field model also contains a scaling factor b_i for each edge e_i, which determines its visibility. In the image, the degraded visibility of the contour and other elements located in the periphery would be reflected in the model by relatively low b_i values for these elements.

contrast of background and contour elements located in the periphery. Although it may be argued that no scaling for peripheral locations may be required in the presence of eye movements, we included the b parameter in the original fit of our data specifically to test whether observers' eye movements were indeed sufficient to make collinear configurations in the periphery more salient.

Model dynamics

The dynamics of the contour-integration model are described by a time-continuous differential equation for the activation $p_i(t)$ of population i, which is a generalization of the time-discrete model used in Ernst et al. (2012):

$$\tau \frac{dp_i(t)}{dt} = -p_i(t) + \eta(t) g\left[b_i \sum_j w_{ij} p_j(t) \right]. \quad (3)$$

In this equation, $g[\ldots]$ denotes a rectifying neural gain function, and $\eta(t)$ is a normalization factor that is defined by the sum over the synaptic input term in rectangular brackets,

$$\eta(t) = \sum_i b_i \sum_j w_{ij} p_j(t). \quad (4)$$

If contour integration is regarded as a probabilistic inference problem (Ernst et al., 2012), p_i can be interpreted as the likelihood that a contour passes through an edge element in a stimulus. p_i is computed iteratively for contours of increasing lengths. For retaining the analogy to neural networks and neural processes in visual cortex, however, we will refer to this variable as the model's "activation" throughout the following text.

Fitting procedure

To avoid overfitting, we used one part of the available experimental data to perform the fit and tested the model's predictive power on a different part of the data. In particular, we used the contour-detection behavior of all subjects in the LR task to

optimize the parameters of the contour-integration model and then tested the generalizability of the data fit by applying a model with the same parameters for subsequent prediction of saccades in the PA task.

For our fit, we required the model both to reach or surpass human contour detection performances and to reproduce excess correlations among human observers. The term "excess correlations" refers to human decisions, which are either less frequent or more frequent than expected from human mean performance. One example shall illustrate this idea: Assuming that performance of human observers in detecting a contour is 70% correct, we expect, on average, 14 correct decisions for each stimulus when we would have 20 subjects in total. If we find a stimulus with only one correct and 19 incorrect decisions, meaning that observers are systematically wrong, we also expect our model to "fail" on this particular contour. These two criteria were quantified in two measures Z and C, respectively. Z takes a value between zero and one and denotes the percentage of stimulus conditions in which a model reaches or surpasses mean human performance. We required our model to achieve $Z = 1$. C can also take a value between zero and one and is related to the integral of a receiver-operator characteristic (ROC) curve. It quantifies excess correlations between the correct decisions of two (sets of) observers: If $C = 0.5$, identical decisions of the two observers are fully explained by their mean performances. If the number of identical decisions of the two observers increases over this chance level, C becomes larger than 0.5 with a maximum attainable value of one (for details with regard to the precise computation of C, we refer the reader to Ernst et al., 2012). In order to quantify how well a model reproduces human behavior, we compare average excess correlations $C_H = 0.79$ among subjects with average excess correlations C_M between model decisions and subjects, requiring the model to come as close as possible to the "benchmark" value C_H.

Fitting was performed on five parameters, namely σ_α and σ_β (shape of association field), μ and ν (scaling of edge saliency with visual field eccentricity), and λ (scaling of coupling strength with edge distance). For each parameter, a plausible range of values was selected and divided into equidistant intervals. After initialization with a uniform-activation distribution, the model was simulated for a time interval of $T = 10\tau$ for all possible parameter combinations. The model with $Z = 1$ and the highest value of C_M identified by this procedure was subsequently used to predict saccade targets.

Prediction of fixation locations

The model generates a time series of activities $p_i(t)$ for all populations i. Predictions for saccade targets were derived by comparing p with a time-varying threshold $\Theta(t) = \langle p_i(t) \rangle_i + 0.5\sigma(p_i(t))$, where $\langle \ldots \rangle$ denotes the mean and σ the standard deviation of the model activation *at a particular point in time*. If $p_i(t) > \Theta(t)$, the corresponding edge element i was considered a potential saccade target. For all stimuli containing a contour, we also removed all fixations within a radius $R_T = 3.5°$ around the center of mass of the target contour. Removing fixations near the target had two reasons. First, we wanted to find out if the model could predict fixations *prior* to detection of the target contour. Second, we indirectly used this information already during parameter fitting because saccade trajectories end near a contour if it is detected by an observer.

For assessing how well the model predicts the remaining fixations, we counted the relative number of fixation spots $t_{pos}(R)$ ("true positives"), which were within a range R of at least one potential saccade target identified by the model. This procedure was also performed for the same model prediction compared with fixations from a *different* stimulus, thus giving a relative number of "false positives" $f_{pos}(R)$. By relating t_{pos} with f_{pos} via the free parameter R, we obtained ROCs whose integral indicates how well the model can predict the fixation spots (0.5 is chance level).

The ROCs were first computed for each observer individually and then averaged by taking the maximum value achieved over simulation time t for each observer. This procedure accounts for the possibility that different subjects might choose saccade targets from different stages of the contour-integration process: Although one observer might make saccades only to contours of at least five elements, a different observer might judge contours of three already aligned edge elements as sufficiently "interesting" for performing a saccade and for conducting a further inspection of its neighborhood.

Results

Task performance

Mean performance on the contour-integration task decreased as the path angle of the target contour increased. For increasing path angles 0°, 5°, 10°, and 20°, the average proportion correct across observers equaled, respectively, 97%, 96%, 90%, and 74% in the LR task and 92%, 88%, 77%, and 56% in the PA task. For the condition in which no contour was present in the display, mean performance was 81% correct. Box plots of the number of fixations and response times as a function of path angle for both the LR task and the PA task are shown in the left and right panels of Figure 3,

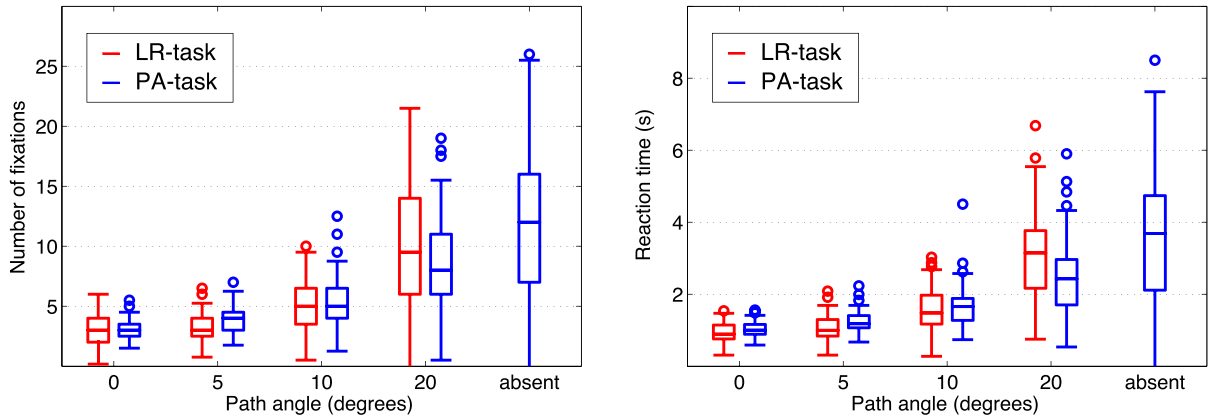

Figure 3. Box plots of number of fixations (left panel) and reaction times (right panel) as a function of the path angle of the contour (including the condition in which no salient contour is present in the PA task). As the path angle increases, the number of fixations and reaction times increase. Data from the LR task are shown in red, and data from the PA task are shown in blue.

respectively. Repeated-measures ANOVA revealed a significant difference between path-angle conditions in log response times, Greenhouse-Geisser correction, $F(1.74, 34.86) = 224.18$, $p < 0.001$ for the LR task and $F(1.71, 34.26) = 89.27$, $p < 0.001$ for the PA task, as well as log number of fixations, Greenhouse-Geisser correction, $F(2.33, 46.65) = 170.58$, $p < 0.001$ for the LR task and $F(1.79, 35.82) = 74.47$, $p < 0.001$ for the PA task. Pairwise comparisons between path-angle conditions showed that all differences were significant after Bonferroni adjustment (requiring $p < 0.008$ for the LR task and $p < 0.005$ for the PA task) except the difference in log number of fixations between the smallest path angles in the LR task ($p = 0.074$).

Eye-movement behavior

To account for random variations between trials (i.e., images) and between observers, the regression coefficients in our multilevel models predicting fixation duration and saccade amplitude were treated as random variables. For the prediction of fixation duration, significant sources of random variation after controlling for the predictors' contour saliency and correctness were a random intercept and slope for observers and a random intercept for trials. For the prediction of saccade amplitude, we included a random intercept, slope, and quadratic term for observers and a random intercept for trials.

The left and right panels of Figure 4 show the change in log fixation duration over the course of a trial as a function of contour saliency for the LR task and PA task, respectively. Fixation durations increased gradually with the rate of change decelerating over time and slightly decreased toward the end of the trial. The last fixation in each trial appeared considerably longer than the preceding fixations due to the overlap with the observers' responses, and these fixations were excluded from further analysis of fixation duration. For both tasks, multilevel analysis showed a significant positive linear time effect, $F(1, 191.55) = 200.27$, $p < 0.001$ for the LR task and $F(1, 832.71) = 230.82$, $p < 0.001$ for the PA task, and negative quadratic time effect, $F(1, 32{,}322.09) = 183.82$, $p < 0.001$ for the LR task and $F(1, 28{,}586.34) = 150.81$, $p < 0.001$ for the PA task, indicating that the linear effect of time decelerated significantly over the course of the trial.

In addition, there was a main effect of contour saliency for the LR task, $F(3, 2{,}324.74) = 20.20$, $p < 0.001$, and the PA task, $F(4, 6{,}614.24) = 8.27$, $p < 0.001$. The effect size of contour saliency in predicting fixation duration was assessed using pseudo-R^2 statistics (Singer & Willett, 2003), indicating the change in the variance components after including the effect of contour saliency in the model. As to be expected, including the predictor contour saliency caused only a small change in the random variance of the intercept and linear time effect for *observers* as each observer received all path-angle conditions. However, contour saliency accounted for 51% and 62% of the random intercept variance for *stimuli* in the LR task and the PA task, respectively. Moreover, contour saliency did not interact with the linear and quadratic effects of time, meaning that the effect of contour saliency remained the same over the course of a trial. There were no significant differences in log fixation duration between correct and incorrect trials for the LR task, $F(1, 32{,}589.16) = 1.07$, $p = 0.301$, or the PA task, $F(1, 28{,}428.23) = 0.35$, $p = 0.553$. The fitted curves in the left and right panels of Figure 4 show the pattern of change in log fixation duration across time intervals as predicted by a quadratic model of time that includes a main effect of contour saliency. In a separate analysis, we combined the data sets of the LR task and the PA task and found no difference in log

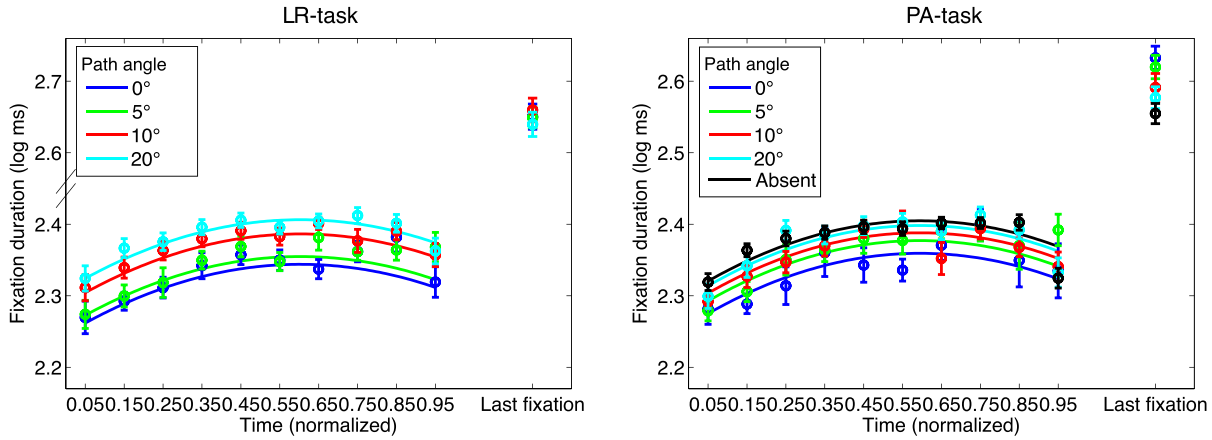

Figure 4. Log fixation duration in the LR task (left panel) and the PA task (right panel) as a function of normalized viewing time can be predicted by a quadratic model of time, including a main effect of contour saliency as indicated by the solid lines. The data fits suggest that fixation durations increased gradually and slightly decreased toward the end of the trial. Different colors indicate the contour saliency levels (i.e., the different path-angle conditions and the contour-absent condition of the PA task, see legend). Error bars represent the standard error of the mean across observers.

fixation duration between the two tasks, $F(1, 52.84) = 0.62$, $p = 0.434$.

Panel A of Figure 5 shows the effect of viewing time on log saccade amplitude as a function of contour saliency for the LR task. The pattern of change in log saccade amplitude suggests a quadratic pattern of change with saccade amplitude peaking at around 40% of the time in the trial and decreasing toward the end of the trial. Indeed, the dependence of saccade amplitude on time can be predicted by a quadratic model of viewing time as shown by the fitted curve in panel A of Figure 5. Our analyses showed that the linear, $F(1, 24.12) = 17.13$, $p < 0.001$, as well as the quadratic time effect were significant, $F(1, 24.25) = 42.25$, $p < 0.001$. The main effect of contour saliency failed to reach significance, $F(3, 12{,}268.96) = 2.41$, $p = 0.065$.

Panel B of Figure 5 shows the dependence of log saccade amplitude on viewing time as a function of contour saliency and correctness of response for the PA task. Here, log saccade amplitude seems to follow two trajectories: The blue curves depict the saccade amplitudes for trials in which no contour was perceived (i.e., trials in which observers either correctly or incorrectly indicated the absence of a contour). In these trials, saccade amplitude gradually increased over the course of the trial. The red curves show saccade amplitudes when observers reported finding a contour. Here, saccade amplitude followed roughly the same time course as in the LR task. Thus, the time course of saccade amplitude seemed to depend on the contour saliency condition (i.e., present or absent conditions) in which the observers gave a correct or false response. This is reflected by a nearly significant interaction between correctness, contour saliency, and the linear time effect, $F(4, 32{,}058.75) = 2.35$, $p = 0.052$, and a highly significant interaction between correctness, contour saliency, and the quadratic time effect, $F(4, 31{,}963.95) = 11.90$, $p < 0.001$. As a measure of effect size, including this interaction between contour saliency and correctness reduced the random intercept variance for stimuli by 57%. The bold red and blue curves in Panel B of Figure 5 show the fits of a quadratic model of time in which the effect of reporting a contour interacts with the linear and quadratic effects of time. When only considering trials in which a contour was present in the PA task, the distance of fixation to the contour follows a similar time course to that of saccade amplitude. Panel C of Figure 5 shows the mean distance to the contour as a function of normalized time. For trials in which they correctly perceived the contour, observers progressively made fixations closer to the contour whereas they gradually fixated locations further from the contour for trials in which they did not perceive the contour. The data sets of the LR task and the PA task were combined into a further analysis, and a significant difference in log saccade amplitude was found between the two tasks, $F(1, 52.06) = 8.01$, $p = 0.007$: Saccade amplitude was, on average, larger for the LR task than for the PA task. The difference in saccade amplitude distributions between the LR task and the PA task is also obvious from panel D of Figure 5, where it can be seen that long saccades prevail in the LR task.

Prediction of fixation locations

Before applying the association field model to our eye movement data, we performed a parameter fit of the model on the observed contour detection perfor-

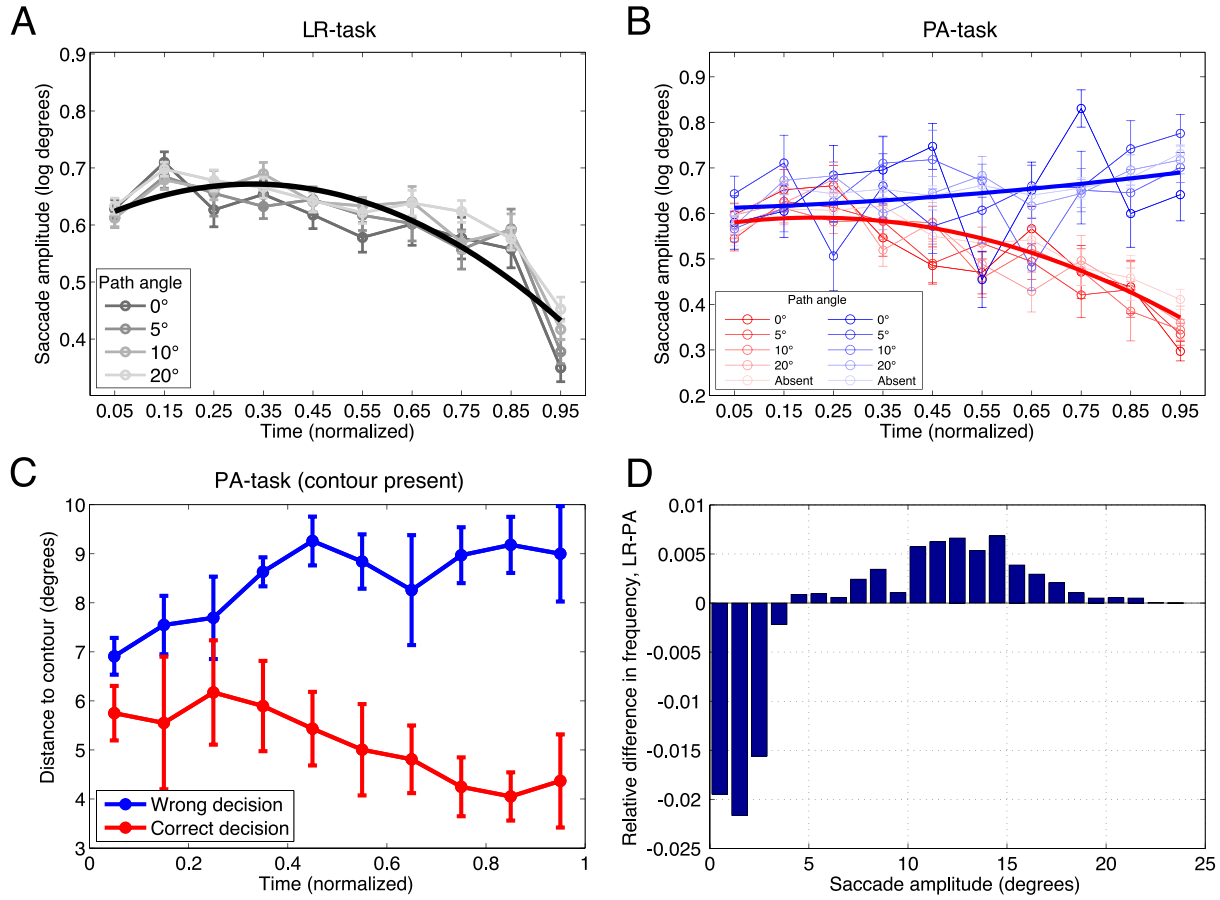

Figure 5. Log saccade amplitude in dependence on normalized time during a trial. Error bars indicate the standard error of the mean across observers. (A) For the LR task, log saccade amplitude peaks at around 0.40 normalized time and decreases toward the end of the trial. This can be predicted by a quadratic model of time as indicated by the bold black curve. Different shades of grey indicate different path angles of the contours. (B) For the PA task, the time course of saccade amplitude depended on the percepts of the observers (i.e., whether they reported the contour to be present or absent). The blue curves show log saccade amplitudes (all path angles and contour-absent condition) when no contour was perceived, and the red curves show log saccade amplitudes when observers reported having perceived a contour. The time course of saccade amplitude can be predicted by a quadratic model of time in which the effect of perceiving the contour interacts with the linear and quadratic effect of time as indicated by the bold curves. Different shades of blue and red indicate either different path-angle conditions or the contour-absent condition. (C) Mean distance of fixation to the contour in dependence on normalized time during a trial for the contour-present conditions in the PA task. Observers progressively made fixations closer to the contour in case of a correct decision whereas they fixated further from the contour over the course of a trial in case of a wrong decision. (D) Difference in saccade amplitude distributions between the LR task and the PA task for trials in which a contour was present.

mance as described in the Methods section. An optimal fit was achieved when elements were assumed to have the same visibility over the whole display (flat eccentricity scaling function, $\mu = 0$). This result is consistent with the idea that elements and contours at the borders of the screen, which would have lower visibility when observers were not allowed to make eye movements, become more salient when saccades are allowed to inspect locations in the periphery. For the association field parameters, which determine the coupling matrix w_{ij}, we found optimal values of $\sigma_\alpha = 0.15$, $\sigma_\beta = 0.24$, and $\lambda = 0.3°$ (left panel of Figure 6).

To determine whether saccades are made to collinear structures that are salient according to our contour-integration model, we compared actual fixations with model predictions. Figure 7 shows such a comparison for four examples from each of the experimental conditions. For the purpose of visualization, we constructed an activity distribution from the discrete activations p_i by using a convolution with a Gaussian profile (half-width 2° of visual angle). The neural activity distribution from the association field is shown as a normalized heat map with red indicating high activity. Fixations of all subjects are superimposed on

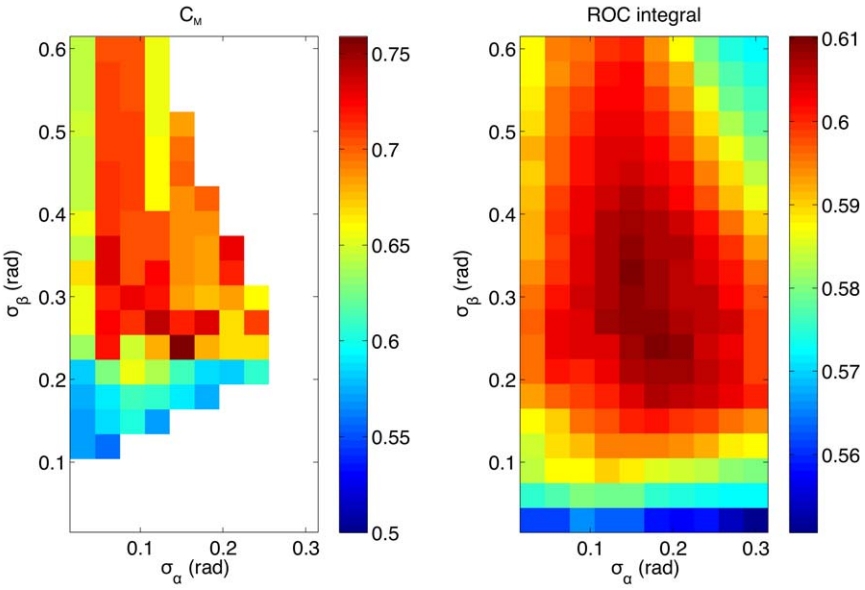

Figure 6. Fitting the model to psychophysical performance and eye-movement data. Left panel: Excess correlations quantified by C_M for different models with association-field parameters σ_α and σ_β varied independently. The other parameters were held constant at their optimal values. The color scale indicates how well the model reproduces human excess correlations. Parameter combinations for which model performance was inferior to human performance are left white. Right panel: ROC integral for different models with association-field parameters σ_α and σ_β varied independently. The color scale indicates how well the model predicts human-fixation locations.

these activity maps, shown as black stars. It can be seen that hot spots of model activity, in general, match well with dense clusters of fixation spots, but there are also examples in which participants fixate regions of low model activity. This impression of the data from these examples is confirmed by a correlation analysis for all stimuli in the PA task in which contours are absent (Figure 8). For this analysis, fixation locations and model predictions for each stimulus were first convolved with a Gaussian profile of half-width 1° of visual angle. Then, a correlation coefficient over the resulting two-dimensional distributions was computed. The distribution of correlation coefficients from this data set (blue bars in Figure 8) is centered around 0.3 and well separated from the distribution of coefficients for surrogate data, for which we compared model predictions to fixations made for a randomly selected, *different* stimulus (red bars in Figure 8).

As a more direct quantitative measure of the predictive power of the model, we computed a ROC that relates the amount of "true positives" to "false positives" in dependence on a decision criterion R. In this case, R is given by the radius around a candidate edge in a stimulus. A candidate edge is an edge for which the model predicts an activation that is higher than a certain threshold (see Methods). If a fixation spot falls within a radius R around a candidate edge, it is taken as a "true positive." If a fixation spot of a saccade trajectory for a randomly chosen, different stimulus (with the same contour path angle) falls within a radius R around a candidate edge, it is taken as a "false positive." Loosely speaking, the integral over the ROC then quantifies how well fixation spots are predicted by the model compared to "chance" prediction performance.

In this analysis, we obtained a value of 0.608 for the LR task, a value of 0.64 for the PA task when the contour was present, and a value of 0.583 for the PA task when the contour was absent. These numbers are significantly different from the value that would be obtained by chance (i.e., 0.5 with a threshold of 0.525 for $p < 0.05$). The ROC analysis therefore shows that the model's performance is clearly above chance level, meaning that activation in an association-field model predicts fixations in a contour detection task well. However, not all human saccades can be explained on the basis of the model. Predictions are worse for stimuli containing only background elements. For these stimuli, the high percentage of false positives emerging in a confined stimulus arrangement with a large number of fixations is problematic, reducing the ROC integral considerably. Hence, the reported ROC values may underestimate the true predictive power of the model. As a consistency check for our initial model fit, we performed an independent search to identify parameter sets maximizing the ROC integral values. This procedure revealed a similar optimal parameter range as found by the fit using human performance and

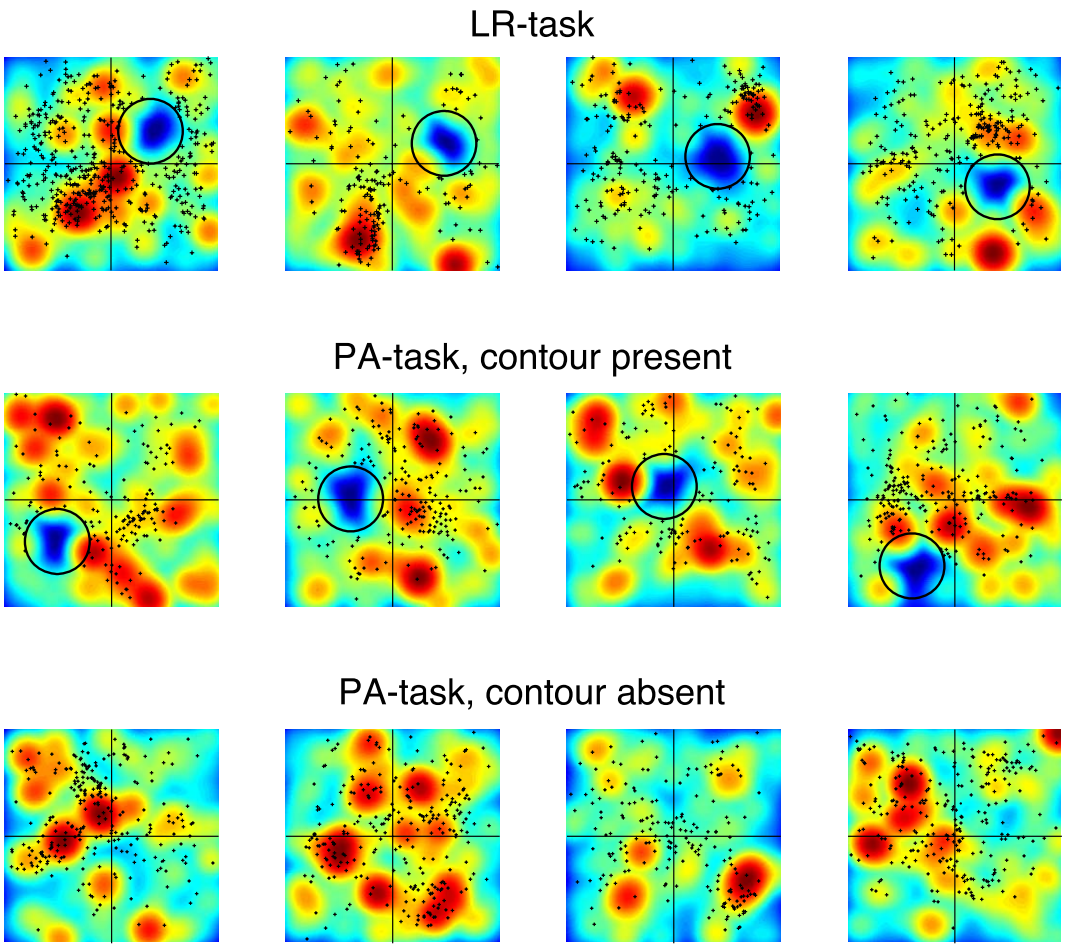

Figure 7. Heat maps of model activity for a subset of the stimuli in the LR task (top row), the contour-present condition in the PA task (middle row), and the contour-absent condition in the PA task (bottom row). Fixations of all observers are indicated as black stars. High activity is colored in red, and low activity is displayed in blue. The black lines indicate the vertical and horizontal meridian of the display. The black circle encloses the region around the target contour from which fixations were removed before comparing the fixation locations to model predictions. The plots suggest that fixation spots, in general, tend to cluster at locations of high model activity although, in some cases, participants fixate regions where model activity is low.

excess correlations as the relevant measures (right panel of Figure 6).

These analyses can be extended to individual participants as illustrated in the top left panel of Figure 9 showing ROC integrals for each of the individual observers over time. Whereas the model predicts fixations for certain subjects well, for others its predictions barely exceed chance level. This analysis also suggests that the maximum predictive performance for each observer is reached at different points in (simulation) time. This is reflected by the absolute values on the time axis, measured in units of the model's time constant τ, which can be regarded as an indicator of how many elements the model has integrated into a potential contour at these points in time. For example, a time of 2τ means that any combination of three edge elements in the stimulus has been integrated by evaluating any set of two links between it. As observers' maxima range from about 2τ to 5τ, it seems that human saccades are best predicted by identifying contour configurations of three to six elements. This is also evident from the bottom left panel of Figure 9, which shows the average ROC integral across observers as a function of time. The right panel of Figure 9 demonstrates that the individual differences in predictive power were preserved across experiments. If fixations of an individual observer were well predicted in the LR task, they were also well predicted in the PA task (both for contour-present and contour-absent conditions). For the stimuli in which a contour was present, we also found a negative correlation between ROC integral and mean number of fixations per observer (slopes -0.0053 and -0.0056 for a linear regression with $R^2 = 0.36$ and $R^2 = 0.30$ for the LR task and PA task, respectively, data not shown).

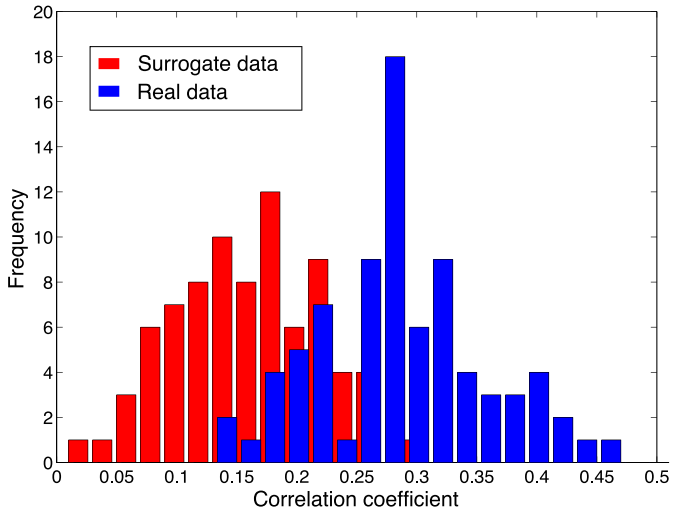

Figure 8. Distributions of correlation coefficients indicating how well our model predicts fixation locations for real (blue) and surrogate (red) data of the contour-absent condition in the PA task. It can be seen that the distribution of correlation coefficients for the real data is well separated from the distribution of coefficients for the surrogate data.

Discussion

We conducted two contour detection tasks to investigate how saccadic eye movements and contour integration combine to support a perceptual decision. In both tasks, observers were free to move their eyes across the image to search for the contour. The difficulty of grouping contour elements into a global percept of a contour was manipulated by the degree of curvature (path angle) of the contour while the number of elements in the contour was kept constant. Main results common to both tasks were

- Gabor fields containing a relatively salient contour (i.e., with a small path angle) led to small numbers of eye movements, which each had short fixation durations, resulting in fast response times. The opposite was found for Gabor fields containing a contour with a large curvature, with which large numbers of fixations and longer fixation durations as well as long response times were found.
- For all path angles, fixation durations increased during the course of a trial whereas the time course of saccade amplitudes depended on whether or not a potential contour could be perceived. When subjects indicated that they could not identify a contour (independently of whether a contour was present or absent), saccade amplitudes increased over a trial. When participants reported perceiving a contour, saccade amplitudes decreased.
- Many fixations could be predicted on the basis of saliency maps obtained from an association-field model of contour integration (Ernst et al., 2012). This suggests that measurements of fixation locations can be used to probe the observers' internal representation of the association field. Fitted model parameters on the basis of behavioral responses indicated that saccades compensate for the decrease in contour saliency usually observed in peripheral vision.

These results will be discussed in more detail in the following sections.

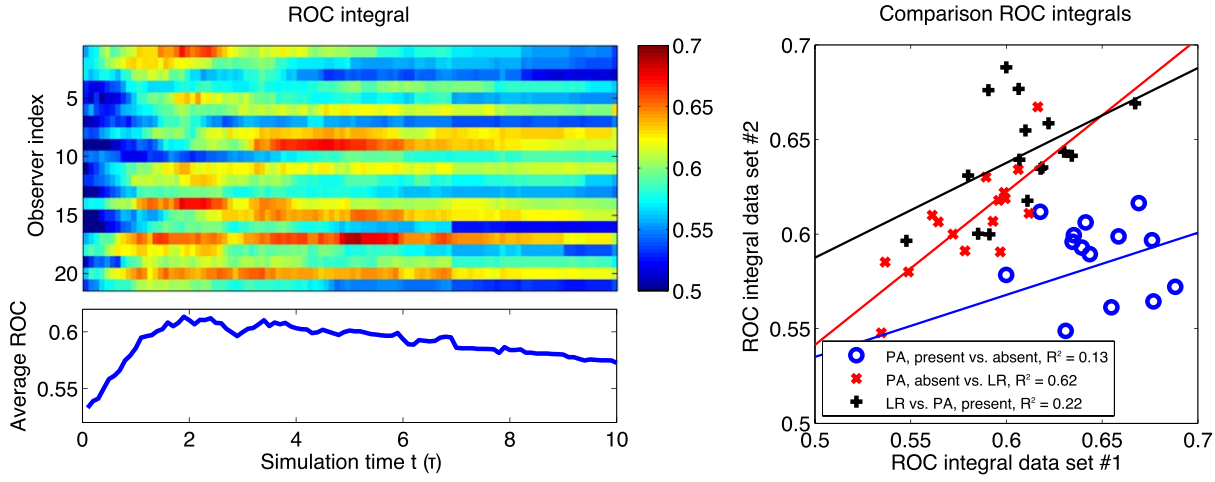

Figure 9. Left panel: ROC integral, indicating the performance of the optimal model in predicting eye fixations in the PA task in the contour-absent condition for different observers in dependence on simulation time *t*. Overall performance of the model and the time interval in which the model has the highest predictive power differ greatly between subjects. ROC integrals averaged across observers as a function of simulation time are shown in the lower graph. Right panel: Correlation between the ROC integrals of observers in the LR task and the contour-present and contour-absent conditions of the PA task.

Effect of contour saliency

The strongest effects of the saliency of the contour were found on the number of fixations and the overall reaction time, which both increased with the path angle of the contour. The highest reaction times and number of fixations were observed for the condition in which no salient contour was present in the PA task. In addition, we found a significant increase in fixation durations with decreasing contour saliency in both the LR task and the PA task. As the observed variations in fixation duration are much smaller, the reaction time is mainly determined by, and thus proportional to, the number of fixations. This finding has also been observed in other studies on visual search (McCarley et al., 2006; Zelinsky & Sheinberg, 1995).

The effect of path angle on fixation durations is supported by previous visual search studies showing increased fixation durations with decreasing target saliency (Hooge & Erkelens, 1996; Jacobs, 1986). For instance, Hooge and Erkelens asked observers to search for a circle among six letter Cs (i.e., rings containing a gap) and found longer fixation durations when the distractor elements had smaller gap sizes, i.e., when the target was less salient. Notably, the increase of fixation duration for higher path angles in our study existed from the beginning of a trial, despite the facts that trials with different contour path angles were randomized and observers could not know the saliency of the contour in a given trial in advance. As the background itself did not reveal any information about the path angle of the embedded contour, it is an indication that ongoing contour-integration processes had already extracted information about the contour (possibly computing an initial but still incomplete saliency map) and that this information was used to directly control the temporal parameters of the eye movements. As a consequence, more time is spent on examining edge configurations in the vicinity of a fixation location when no clear contour candidates are present in the initial assessment of a stimulus. Hence, fixation durations seem to be under the immediate control of the information available in the stimulus, consistent with evidence showing that the duration of some fixations increased when the onset of the stimulus at the saccadic landing position was delayed (Henderson & Pierce, 2008; Henderson & Smith, 2009). Contrary to our finding, Hooge and Erkelens observed that target saliency did not influence fixation durations when it changed from trial to trial. According to the authors, the unpredictability of task difficulty prevented observers from correctly adjusting their fixation durations because estimation of time needed at fixation was based on previous trials. A possible reason for the difference between our findings and those of Hooge and Erkelens may be the amount of processing required to solve the task. Our tasks were relatively complex and involved the grouping of local elements into contours whereas the study by Hooge and Erkelens involved a much simpler search task in which the target only differed from background elements with respect to one feature (i.e., a gap). In our study, it might be more important to adjust fixation durations at the beginning of each trial so that more processing resources can be allocated to fixation locations in difficult trials. In fact, Hooge and Erkelens acknowledged the potential importance of the first fixation as the time when the difficulty of the current trial is assessed. It should be noted that fixation durations were approximately the same for correct and incorrect trials, indicating that fixation duration did not depend on whether a correct perceptual decision was made at the end of the trial. Moreover, we found no difference in fixation durations between the LR task and the PA task. This is not surprising because the contour stimuli used in the two tasks were very similar, requiring the same amount of processing at the fixation location. Our stimuli did differ between tasks in spatial constraints related to the location of the contour, which influenced the amplitude of saccades (see below).

Dynamics of eye movements

Fixation durations increased in the beginning of the trial, followed by an interval of more stable durations and a slight decrease toward the end of the trial. This pattern of change was found across both tasks and all path-angle conditions. In parallel, the saccade amplitude gradually decreased in the LR task and the PA task when a contour was perceived, indicating that observers switched from a global and coarse-scale to a more local and fine-scale search strategy, presumably when observers encountered a likely contour candidate. This interpretation is supported by the observation that the mean distance of fixation to the target contour showed a similar decay over time when this contour was correctly detected. The time course of saccade amplitude is in line with a number of studies showing that information about the gist of the scene can be processed very quickly after the onset of a stimulus and can guide initial saccades to regions where the target is likely to be found so that these regions can be further explored using smaller saccades (Eckstein, Drescher, & Shimozaki, 2006; Neider & Zelinsky, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006). It should be noted that in the LR task, saccade amplitude first slightly increased during the first fixations. A possible reason is that observers initially make saccades to locations close to the fixation symbol present prior to stimulus presentation and only move toward more peripheral locations after a few fixations.

The decrease of saccade amplitude occurring together with an increase of fixation duration over time seems to reflect a general coarse-to-fine search strategy that has been frequently reported in earlier studies (Antes, 1974; Castelhano, Mack, & Henderson, 2009; Over et al., 2007; Pannasch et al., 2008; Scinto, Pillalamarri, & Karsh, 1986; Unema et al., 2005). The gradual decrease in saccade amplitudes, however, was not found when observers reported not perceiving a contour in the PA task. In these trials, observers continued to make large saccades throughout the course of the trial. Thus, a coarse-to-fine search process only seems to occur when a target candidate is found at the end of the trial. The coarse-to-fine processing strategy therefore does not seem to be a default mechanism of the visual system as has been suggested by others (Over et al., 2007; Pannasch et al., 2008; Unema et al., 2005). A possible reason for the discrepancy between our findings and previous studies reporting a coarse-to-fine time course of saccade amplitude during visual search (Castelhano et al., 2009; Mills et al., 2011; Over et al., 2007; Scinto et al., 1986) could be that in these earlier studies the search target was eventually found in the majority of the trials. Presumably, observers made smaller saccades near the end of a trial to process the image region in which the target was found in more detail. In contrast, in a visual search study in which the search target was rarely detected, observers made large saccades, which remained relatively constant over the course of the trial (Mills et al., 2011). It was argued that this result was due to the fact that the images contained no contextual information that could guide participants' search behavior. Observers might have therefore adopted a strategy of making large saccades to visit as many locations as possible because more in-depth processing of the scene did not help in finding the target location. In our study, it is likely that for trials in which observers did not perceive a contour there were no image regions that attracted their attention for further scrutiny, resulting in observers using a similar strategy as in the study by Mills et al. Consistent with this idea, the small increase in saccade amplitude for trials in which observers reported not perceiving a contour can be predicted by a random search model that visits stimulus locations that have not been investigated before until the whole visual field has been searched (simulations not shown). In consequence, the slight increase in saccade amplitude over time probably does not reflect an active adaptation of the search strategy but is rather caused by geometrical constraints. The notion of a rather random search is also supported by the lower predictive power ($R = 0.583$) of the model for contour-absent stimuli than for stimuli in which a contour is present ($R = 0.64$).

It is important to emphasize that the dynamics of saccade amplitudes in the PA task did not depend so much on the actual stimulus content (contour present or absent) but rather on the percept of observers (contour perceived or not perceived). A similar difference in the dynamics of saccade amplitudes was not found in the LR task. This is presumably due to the fact that there was always a contour present in the LR task, causing observers to be "forced" to find a contour. As a consequence, they are more likely to "home in" on a potential target even if it is not the contour that was originally hidden in the stimulus. In the PA task, on the other hand, observers are "allowed" to make the decision that no contour is present, resulting in trials in which no detailed processing of particular image regions takes place and long saccades prevail. Overall, saccade amplitudes were larger in the LR task than in the PA task. A possible reason is the existence of a division zone at the center of the display in the LR task. Observers do not expect contours in this zone and therefore seldom make eye fixations toward it. This mental division of the display in two hemifields may have caused larger saccade amplitudes when observers jump between hemifields.

In sum, fixation durations appear to be determined by the saliency of a present contour from the beginning of the trial and systematically increase over the course of the trial, resulting in more detailed processing. In contrast, the time course of saccade amplitudes depends on the potential locations of the target contour and whether or not the contour can be perceived, irrespective of the saliency of a present contour. These results imply that fixation duration and saccade amplitude are not controlled by a common mechanism, as has been suggested by others (Pannasch et al., 2008; Unema et al., 2005), but instead that task and stimulus factors influence these parameters differently. A similar view has been put forward by Mills et al. (2011), arguing that the increase of fixation duration reflects a gradual build-up toward a more complex representation of the image whereas the pattern of change in saccade amplitude depends on the utility of visual information available in the display.

Predictive power of contour-integration model

The contour-integration model was calibrated and tested on independent data. For determining model parameters, we used the decisions of observers in the LR task (deciding whether the contour was in the left or right hemifield). Based on previous findings showing that the range of parameter values of the association field reaching or exceeding mean human performance is very broad (Ernst et al., 2012; Watt et al., 2008), the model was required to also reproduce correlations in

human contour detection behavior rather than only mean human performance, which resulted in a better parameter-selection criterion. The predictive power of the model was then quantified on the eye-tracking data, and we explicitly excluded eye movements in the vicinity of the target contour (which are indicative of all correct decisions made). The model was able to predict observers' fixation locations well above chance level in both the LR task and the PA task. As a consistency check, we compared model performances in reproducing behavioral decisions and in predicting eye movements. This analysis confirmed that the parameters found during model calibration are also optimal for predicting fixation locations. These results strengthen the validity of our model as it not only reproduces human response behavior, but also predicts the salient image locations at which observers allocate their attention during contour integration. Hence, our modeling approach seems to be able to reconstruct the contour-integration process that guides both perceptual decisions and eye-movement behavior.

The ROC analysis and visual comparison of fixation locations with model predictions demonstrated that many but not all fixations were predicted. The worst performance was actually found for the PA task when contours were absent, which we attribute to a more random search strategy (see above). ROC integral values range between 0.58 and 0.64, which is in the same range as previously reported values from other models predicting fixations during image viewing (i.e., from 0.55 to 0.70; e.g., Betz, Kietzmann, Wilming, & König, 2010; Einhäuser, Spain, & Perona, 2008; Renninger, Verghese, & Coughlan, 2007). Importantly, such eye-movement models include additional assumptions concerning the sequence of fixation locations (e.g., winner-take-all principle, inhibition of return to previously visited fixation locations). Given that our model is based on a simple association-field computation producing one saliency map, performance is surprisingly good compared to that of more complex models.

Moreover, the model might appear worse than it actually is because the comparison to the real data involves several technical difficulties. First, fixations often do not go directly to a target but instead first go into the vicinity of a (putative) contour (about 2°–3° of visual angle). Thus, the comparison radius of model "hot spots" to fixations must be large, causing also the number of false positives in the surrogate data to increase. Second, model activity evolves over time and might predict different fixation locations at different times in the simulation. The reason is the recurrent nature of the contour-integration process: Over time, longer and longer contour candidates are found, and shorter segments decrease in saliency. Nevertheless, even in the PA task for stimuli without a contour, the correlation between model predictions and experimental data was much higher than chance level. This observation was confirmed by computing the distribution of correlations between saliency and fixation maps, which has a distance of $d' \approx 2$ to surrogate data. It is important to note that our goal was not to predict the temporal sequence of fixations during a trial as this would require a more complex model that takes into account a number of assumptions concerning the planning and monitoring of subsequent eye movements (see Itti & Koch, 2001; Torralba et al., 2006; Zelinsky & Sheinberg, 1997 for examples of computational eye-movement models). Rather, we aimed to test whether a model providing a saliency map based on a single association field is able to predict both contour-detection performance as well as the locations likely to be fixated over the entire course of a trial.

Differences between observers

There are strong interindividual differences between observers both in the experimental data (e.g., number of fixations made) and in the model's predictive power. These differences are consistent across data sets (i.e., LR task and PA task). Another prominent finding is a negative correlation between the mean number of fixations an observer makes and the predictive power of the model. A possible explanation is that observers whose data is not well predicted by the model tend to make eye movements not toward putative contour candidates with high saliency but to more "random" locations in the visual field. Interestingly, the mean number of fixations also correlated with the performance of observers (slopes 0.014 and 0.009 for a linear fit with $R^2 = 0.16$ and $R^2 = 0.07$ for the LR task and PA task, respectively, data not shown). These dependencies suggest that making more eye movements, including less salient locations in the visual field as well, is a more successful strategy to find contour targets. A further difference between individual observers is the model's integration time needed until the best match to their behavior is reached. This difference might indicate that observers differ in their criteria for contour candidates, i.e., they may require different lengths of aligned edge configurations to be salient and to be used as targets for a saccade.

Saccades compensate for reduced saliency in peripheral vision

In our procedure to establish the best parameters for the contour-integration model, we also allowed for an optional scaling of edge saliency with eccentricity. The inclusion of this option was based on several visual

search studies that have reported that detection performance of a target declines at higher eccentricities due to loss of spatial resolution (Carrasco, Evert, Chang, & Katz, 1995; Carrasco & Frieder, 1997; Carrasco & Yeshurun, 1998; Geisler & Chou, 1995; Geisler, Perry, & Najemnik, 2006; Scialfa & Joffe, 1998). In addition, the crowding literature shows that it is particularly difficult to identify a target in the periphery when it is surrounded by nearby distractor stimuli (Bouma, 1970; Levi, 2008). Eye movements help to explore these peripheral target locations and bring them to the foveal region, thereby reducing crowding and allowing for detailed spatial processing of these locations (Harrison, Mattingley, & Remington, 2013; Vlaskamp & Hooge, 2006; Wertheim, Hooge, Krikke, & Johnson, 2006). This latter observation is in agreement with our finding that contour detection behavior was best explained when scaling of saliency with eccentricity was absent, indicating that observers' eye movements can fully compensate for the reduced contour saliency in peripheral vision.

Conclusions

A crucial task of everyday vision involves the grouping of spatially separate elements into more global structures, such as contours. Eye movements play an important role in the detection of contours, allowing observers to fixate elements that belong to potential contours. The aim of our study was to examine how spatial and temporal aspects of saccadic eye movements interact with the process of contour integration. A relatively simple model of contour integration, based on an association field with parameters that best matched human contour-integration performance, could successfully predict a considerable amount of fixation locations before the observers found the contour. This indicates that the observers' fixations are drawn to collinear structures, determined by the saliency map of the model. Model fitting further revealed that saccadic eye movements equilibrate the saliency of peripheral contour candidates across the visual field. As such, our study provides a more stringent test of the validity of the model by requiring it to predict the locations where observers fixate during a contour-integration task rather than requiring it to merely reproduce contour detection performance. Moreover, we find that observers' eye movements are adjusted to the difficulty of the contour-integration task: Temporal and spatial oculomotor parameters, such as individual fixation durations and saccade amplitudes, follow a specific time course that is strategically adjusted to the saliency and the percept of a contour in a given trial. The latter findings have important implications for current computational models of eye movement behavior during visual search. These models are typically used to predict the positions of fixations but only incorporate dynamical aspects of eye movements to a limited extent. Our study shows that spatial and temporal aspects of eye movements can be used to advance our understanding of visual-integration processes by providing rich data sets that allow for the refinement of computational models of these processes.

Keywords: contour integration, eye movements, fixation duration, saccade amplitude, association field, saliency

Acknowledgments

This research was supported by a Research Foundation – Flanders (Fonds Wetenschappelijk Onderzoek – Vlaanderen) fellowship granted to Nathalie Van Humbeeck and by a Methusalem grant to Johan Wagemans (METH/08/02). We would like to thank Roger Watt for his help with stimulus construction.

Commercial relationships: none.
Corresponding author: Nathalie Van Humbeeck.
Email: nathalie.vanhumbeeck@ppw.kuleuven.be.
Address: Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium.

References

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70.

Betz, T., Kietzmann, T. C., Wilming, N., & König, P. (2010). Investigating task-dependent top-down effects on overt visual attention. *Journal of Vision, 10*(3):15, 1–14, http://journalofvision.org/content/10/3/15, doi:10.1167/10.3.15. [PubMed] [Article]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–178.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Attention, Perception, & Psychophysics*, *57*(8), 1241–1261.

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research*, *37*(1), 63–82.

Carrasco, M., & Yeshurun, Y. (1998). The contribution of covert attention to the set-size and eccentricity

effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 673–692.

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision, 9*(3):6, 1–15, http://journalofvision.org/content/9/3/6, doi:10.1167/9.3.6. [PubMed] [Article]

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science, 17*(11), 973–980.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14):18, 1–26, http://journalofvision.org/content/8/14/18, doi:10.1167/8.14.18. [PubMed] [Article]

Ernst, U. A., Mandon, S., Schinkel-Bielefeld, N., Neitzel, S. D., Kreiter, A. K., & Pawelzik, K. R. (2012). Optimality of human contour integration. *PLoS Computational Biology*, *8*(5), el002520.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*(2), 173–193.

Foley, J. M., Varadharajan, S., Koh, C. C., & Farias, M. C. Q. (2007). Detection of Gabor patterns of different sizes, shapes, phases and eccentricities. *Vision Research*, *47*(1), 85–107.

Geisler, W. S., & Chou, K.-L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102*(2), 356–378.

Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision, 6*(9):1, 858–873, http://journalofvision.org/content/6/9/l, doi:10.1167/6.9.1. [PubMed] [Article]

Harrison, W. J., Mattingley, J. B., & Remington, R. W. (2013). Eye movement targets are released from visual crowding. *Journal of Neuroscience*, *33*(7), 2927–2933.

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review*, *15*(3), 566–573.

Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, *17*(6–7), 1055–1082.

Hess, R. F., & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature*, *390*(6660), 602–603.

Hess, R. F., May, K. A., & Dumoulin, S. O. (2013). Contour integration: Psychophysical, neurophysiological and computational perspectives. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization* (in press).Oxford, UK: Oxford University Press.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*(1), 101–117.

Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Attention, Perception, & Psychophysics*, *58*(7), 969–976.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203.

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Attention, Perception, & Psychophysics, 39*(1), 47–58.

Levi, D. M. (2008). Crowding an essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, *10*(4), 903–940.

Locker, L., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, *39*(4), 723–730.

May, K. A., & Hess, R. F. (2007). Ladder contours are undetectable in the periphery: A crowding effect? *Journal of Vision, 7*(13):9, 1–15, http://journalofvision.org/content/7/13/9, doi:10.1167/7.13.9. [PubMed] [Article]

May, K. A., & Hess, R. F. (2008). Effects of element separation and carrier wavelength on detection of snakes and ladders: Implications for models of contour integration. *Journal of Vision, 8*(13):4, 1–23, http://journalofvision.org/content/8/13/4, doi:10.1167/8.13.4. [PubMed] [Article]

McCarley, J. S., Kramer, A. F., Boot, W. R., Peterson, M. S., Wang, R. F., & Irwin, D. E. (2006). Oculomotor behaviour in visual search for multiple targets. *Visual Cognition*, *14*(4–8), 685–703.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8):17, 1–15, http://journalofvision.org/content/11/8/17, doi:10.1167/11.8.17. [PubMed] [Article]

Mundhenk, T. N., & Itti, L. (2002). A model of contour integration in early visual cortex. In *Second international workshop on biologically motivated computer vision* (pp. 80–89). Heidelberg: Springer.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, *46*(5), 614–621.

Nugent, A. K., Keswani, R. N., Woods, R. L., & Peli, E. (2003). Contour integration in peripheral vision reduces gradually with eccentricity. *Vision Research*, *43*(23), 2427–2437.

Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research, 47*(17), 2272–2280.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A. K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*(2), 1–19.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision, 7*(3):6, 1–17, http://journalofvision.org/content/7/3/6, doi:10.1167/7.3.6. [PubMed] [Article]

Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Attention, Perception, & Psychophysics*, *60*(6), 1067–1082.

Scinto, L. F. M., Pillalamarri, R., & Karsh, R. (1986). Cognitive strategies for visual search. *Acta Psychologica*, *62*(3), 263–292.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786.

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*(3), 473–494.

Velichkovsky, B. M., Joos, M., Helmert, J. R., & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the xxvii conference of the cognitive science society* (pp. 2283–2288). Mahwah, NJ.

Vlaskamp, B. N. S., & Hooge, I. T. C. (2006). Crowding degrades saccadic search performance. *Vision Research, 46*(3), 417–425.

Vlaskamp, B. N. S., Over, E. A. B., & Hooge, I. T. C. (2005). Saccadic search performance: The effect of element spacing. *Experimental Brain Research, 167*(2), 246–259.

Wagemans, J., Elder, J., Kubovy, M., Palmer, S., Peterson, M., Singh, M., & von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172–1217.

Watt, R., Ledgeway, T., & Dakin, S. C. (2008). Families of models for Gabor paths demonstrate the importance of spatial adjacency. *Journal of Vision, 8*(7):23, 1–19, http://journalofvision.org/content/8/7/23, doi:10.1167/8.7.23. [PubMed] [Article]

Wertheim, A. H., Hooge, I. T. C., Krikke, K., & Johnson, A. (2006). How important is lateral masking in visual search? *Experimental Brain Research, 170*(3), 387–402.

Williams, L. R., & Thornber, K. K. (2001). Orientation, scale, and discontinuity as emergent properties of illusory contour shape. *Neural Computation*, *13*(8), 1683–1711.

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, *38*(5), 719–741.

Zelinsky, G. J., & Sheinberg, D. L. (1995). Why some search tasks take longer than others: Using eye movements to redefine reaction times. In R. K. J. Findlay & R. Walker (Eds.), *Eye movement research: Mechanisms, processes and applications* (Vol. 6) (pp. 325–336). Amsterdam: Elsevier Science.

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(1), 244–262.