Vision Research 51 (2011) 179–186

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Integration of contour and surface information in shape detection

Bart Machilsen, Johan Wagemans *

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

ARTICLE INFO

Article history:
Received 6 May 2010
Received in revised form 9 November 2010

Keywords:
Perceptual grouping
Figure-ground
Shape detection
Texture segregation
Contour integration
Cue combination
Surface perception

ABSTRACT

In studies of shape perception, the detection of contours and the segregation of regions enclosed by these contours have mostly been treated in isolation. However, contours and surfaces somehow need to be combined to create a stable perception of shape. In this study, we used a 2AFC task with arrays of oriented Gabor elements to determine whether and to what extent human observers integrate information from the contour and from the interior surface of a shape embedded in this array. The saliency of the shapes depended on the alignment of Gabors along the shape outline and on the isolinearity of Gabors inside the shape. In two experiments we measured detectability of shapes defined by the contour cue, by the surface cue, and by the combination of both cues. As a first step, we matched performance in the two single-cue conditions. We then compared shape detectability in the double-cue condition with the two equally detectable single-cue conditions. Our results show a clear double-cue benefit: Participants used both cues to detect the shapes. Next, we compared performance in the double-cue condition with the performance predicted by two models of sensory cue combination: a minimum rule (probability summation) and an integration rule (information summation). Results from Experiment 2 indicate that participants applied a combination rule that was better than mere probability summation. We found no evidence against the integration rule.

1. Introduction

Perceptual grouping of local elements into a more global configuration is an important task for our visual system. It guides the process of segregating an ambiguous retinal image into separate regions belonging to an object or to the background. To detect a visual shape, vision relies on the boundaries between different regions and on the surface properties of these regions. Although boundary detection and surface segregation involve different grouping mechanisms (i.e., contour integration versus region grouping), they both contribute to the detection of a visual shape. In the present study, we apply two-dimensional signal detection theory to estimate how well human observers integrate congruent contour and surface information.

1.1. Contour grouping

To detect the outline contour of a shape our visual system makes use of regularities in the input image. Adjacent elements of a shape outline are usually locally aligned. Detection of this collinearity serves as a cue to the presence of a contour (Geisler, Perry, Super, & Gallogly, 2001). Perceptual grouping of collinear elements closely relates to the Gestalt principle of good continuation (Wertheimer, 1923). The importance of good continuation in contour grouping is illustrated in the *pathfinder* paradigm, in which participants have to detect a straight or curved path in a cluttered background (reviewed by Hess, Hayes, and Field (2003)). The spacing and orientations of elements relative to the path orientation influence the grouping strength (Claessens & Wagemans, 2005; Field, Hayes, & Hess, 1993; Li & Gilbert, 2002). Contour integration also depends on more global stimulus properties, such as path length and curvature (Field et al., 1993; Pettet, 1999; Watt, Ledgeway, & Dakin, 2008), and contour closure (Hess, Beaudot, & Mullen, 2001; Kovács & Julesz, 1993; Pettet, McKee, & Grzywacz, 1998).

1.2. Region grouping

The relevance of contour grouping in shape perception is clear: It helps to define the boundaries of a visual shape. However, contour grouping is not the only means to detect a shape. Shapes can also be segregated from a background by region grouping based on regularities within a textured area. Psychophysical studies on texture segregation have revealed the importance of local discontinuities in orientation, size, or density at the boundary between two textures (Julesz, 1981; Nothdurft, 1991, 1992). However, texture segregation is not only influenced by local orientation contrast at the border between textured areas; the

* Corresponding author. Address: Laboratory of Experimental Psychology, University of Leuven (K.U. Leuven), Tiensestraat 102, Bus 3711, B-3000 Leuven, Belgium.

E-mail addresses: bart.machilsen@psy.kuleuven.be (B. Machilsen), johan.wagemans@psy.kuleuven.be (J. Wagemans).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.11.005

homogeneity of orientations within a textured region (i.e., away from the texture border) also affects texture segregation (Giora & Casco, 2007; Harrison & Keeble, 2008).

1.3. Contour information more important than surface information

Elder and Zucker (1998) argued that a grouping algorithm based on the interior surface of a visual shape requires a degree of surface homogeneity that is usually not present in natural images. Grouping based on the outline of a shape is much more robust against fluctuations in brightness and color inherent to natural images. The relative importance of contour grouping over surface segregation is confirmed by neurophysiological studies on figure-ground organization. Lee, Mumford, Romero, and Lamme (1998) found that the interior enhancement effect – i.e., V1 cells responding more vigorously to the interior of a texture-defined figure than to identical stimulation coming from the background (Lamme, 1995) – declines with increasing distance from the boundary. They argued that this is caused by a smoothing-in of the boundary signals.

1.4. Contour detection precedes surface segmentation

Scholte, Jolij, Fahrenfort, and Lamme (2008) used both ERP and fMRI to reveal the spatiotemporal dynamics of contour detection and surface segregation. They were able to dissociate between contour detection and surface segregation by comparing homogeneously oriented texture displays to similar displays in which a central region within a frame did or did not group with the background. Their results show that early visual areas code for texture boundaries around 92 ms after stimulus onset. The first signals related to surface segregation appear 112 ms after stimulus onset in temporal areas. The surface segregation signal appears in occipital areas only after 172 ms. This finding is in agreement with the neural network model of Roelfsema, Lamme, Spekreijse, and Bosch (2002) in which boundary detection is accomplished in a feedforward sweep, while surface filling-in results from reentrant processes.

1.5. Combining contour and surface information

The literature on border-ownership stresses the importance of combining contour information and surface information: A border between two image regions normally belongs to only one image. Once a boundary is detected the visual system still has to decide on which side of the boundary the figure is situated. This issue is nicely illustrated in the famous Rubin vase-faces figure: The bistable perception of either a vase or two face silhouettes depends on whether the borders are perceived as owned by the inner vase or outer face regions of the image (Rubin, 1921). According to Palmer and Brooks (2008), the notion of border-ownership suggests that borders are perceptually grouped with the region to which they belong. In a series of experiments they showed that boundaries and textures with common features (color, motion, blur, ...) are grouped together.

The importance of both contour and surface information is also evident from the literature on illusory contours (e.g., Fulvio & Singh, 2006; Stanley & Rubin, 2003). A recent neuro-computational model of illusory contours (Kogo, Strecha, Van Gool, & Wagemans, 2010) suggests that the illusory perception of a contour in Kanizsa-type figures emerges as the result of a border-ownership computation that indicates the presence of an illusory surface. In this model, contour completion is not established independently from surface filling-in, but results from constructing surfaces in depth through which the ownership of the contours is defined.

The picture that emerges from the literature is one in which contour information is extracted first, followed by a filling-in of the interior surface. The underlying temporal dynamics indicate that the processes of contour detection and surface filling-in are not independent: Shape perception arises from the grouping of boundaries and regions.

In two psychophysical experiments we investigate how well human observers integrate contour and surface information. With the Gaborized stimuli we can define a shape by its contour, its interior surface, or by a combination of both cues. We equate the shape information provided by the contour and surface cues, and apply the framework of signal detection theory to predict detectability of an embedded shape when both cues are available. By comparing the performance predicted by two models of sensory cue combination with the actual performance in the double-cue condition, we gain insight in how well human observers combine congruent contour and surface information.

2. Methods

We will first provide details on stimulus construction and stimulus presentation, and outline the common framework for the two experiments. Specific details for each experiment will be given in the respective method section.

2.1. Stimulus construction

We used arrays of nonoverlapping Gabor elements on a gray background (Fig. 1). Artificial nonsense shapes were embedded in the arrays. The arrays comprised 496×496 pixels. Each Gabor element was defined as the product of an oriented sine-wave luminance grating (spatial frequency of 3 cycles per degree) and a circular Gaussian envelope (standard deviation of 5.4 arc min).

The embedded shapes were generated by summing 5 radial frequency components (each sine wave having a random phase angle), and plotting the result in polar coordinates (for details see Machilsen, Pauwels, & Wagemans, 2009). This method yielded smoothly curved closed shapes. After rescaling the surface area to one eighth of the array size we colocalized the center of mass of each shape with the center of the array.

Next, we positioned 30 Gabors on equidistant locations along the contour of the shape. These elements had orientations parallel to the local tangent of the shape outline (curvilinear contour elements). Another 30 Gabors were placed in the interior of the shape. Their orientation was parallel to the main axis of the shape (iso-oriented or isolinear surface elements). Finally, we positioned 300 randomly oriented elements outside the shape (random background elements).

To ensure a homogeneous density throughout the Gabor arrays we discarded arrays for which the local density, here defined as the average Euclidean distance from each element to its five nearest neighbors, differed significantly between surface, contour and background elements (three unpaired t-tests, $\alpha = 0.2$). No stimuli were included for which differences in local density between surface, contour, and background elements exceeded 1 arc min. Two thousand arrays, each with a different shape embedded, were selected for the experiment. The average center-to-center distance between Gabor elements was 44 arc min.

The saliency of the embedded shapes could be manipulated by rotating the elements on the contour or in the interior surface of the shape away from their curvilinear resp. isolinear orientations. The shapes were most salient when the orientation of each contour element was parallel to its local tangent (curvilinearity, Fig. 2A), or when the orientation of each interior surface element was parallel to the main axis of the embedded shape (isolinearity, Fig. 2B).

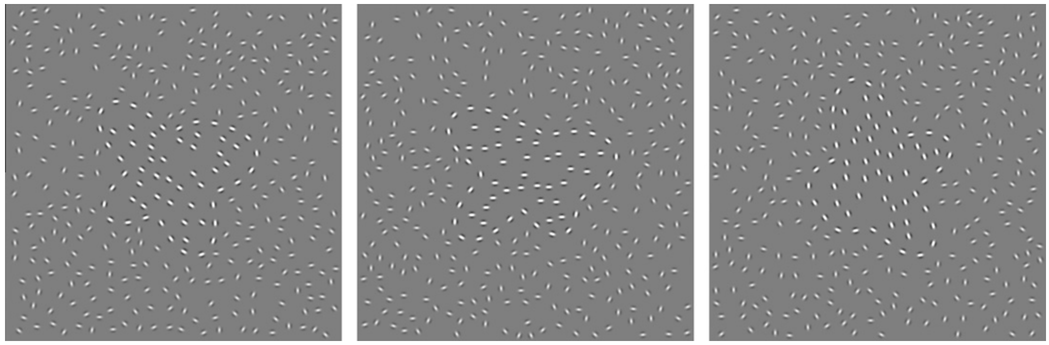

Fig. 1. Example stimuli used in the experiment. In these examples the embedded shapes are defined by collinear contour elements and isolinear surface elements. For illustrative purposes we adjusted the contrast of contour and surface elements to highlight the embedded shapes.

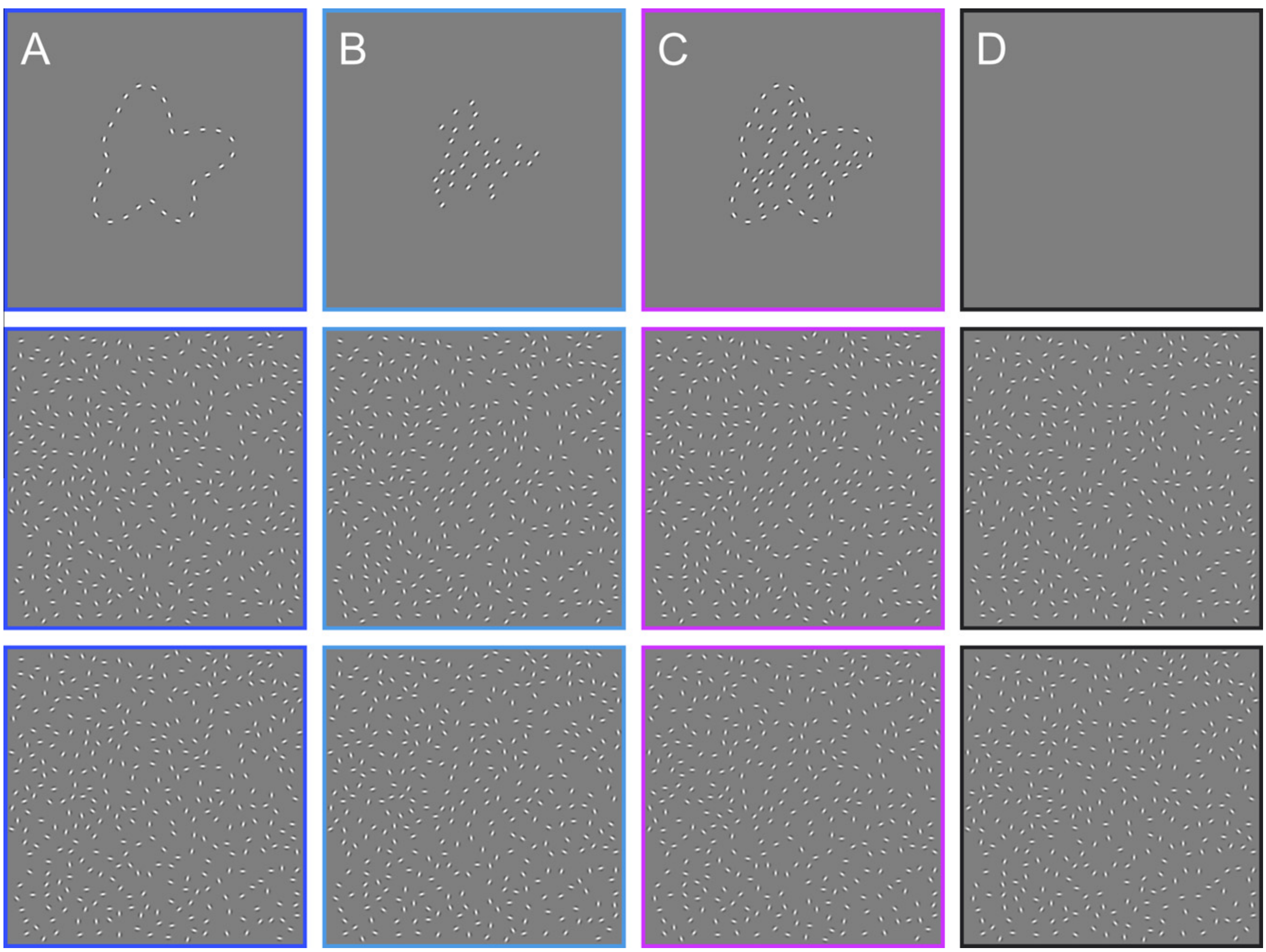

Fig. 2. Example stimuli used in the experiment. The top row depicts the relevant Gabor elements for each stimulus type. For target stimuli (A–B–C) the embedded shape is defined by (A) curvilinearity of the contour elements, (B) isolinearity of the surface elements, or (C) a combination of both. For a non-target stimulus (D) all elements have random orientations. Examples for each stimulus type without orientation jitter are shown in the middle row. The saliency of the embedded shape is reduced by adding orientation jitter to the elements (bottom row, 20 degrees of orientation jitter).

Harrison and Feldman (2009) have shown that texture segregation improved when texture elements were aligned with the skeletal axis of the shape. In the double-cue condition the shapes were defined by both the curvilinearity of contour elements and the isolinearity of surface elements (Fig. 2C). For the non-target displays in our 2AFC experiment the orientation of each element was sampled from a Uniform distribution between 0° and 180° (Fig. 2D).

2.2. Stimulus presentation

Experiments were run in the MATLAB environment using the Psychophysics Toolbox extensions (Brainard, 1997). Stimuli were displayed on a 21″ Sony GDM-F520 CRT monitor with a spatial resolution of 1152 × 864 pixels and a temporal resolution of 85 Hz. At a viewing distance of 57 cm the stimuli subtended approximately

18 degrees of visual angle. The room was darkened during the experiment.

2.3. Cue combination and signal detection theory

In signal detection theory one assumes that the internal representation of sensory information always shows some amount of uncertainty. Hence, the internal response to a signal is represented by a distribution rather than by a single point. The central panel of Fig. 3 illustrates this for the case of a compound signal defined by two cues. In our double-cue condition the shape signal consists of two congruent sources of shape information: contour and surface. To discriminate between a compound target (signal) and non-target (noise) distribution (Fig. 3, central panel), an observer can apply different cue-combination rules. One strategy might be to rely on one single cue, and ignore the other cue. This decision rule (Fig. 3A and B) is called *vetoing* (Bülthoff & Mallot, 1988) or *decisional separability* (Macmillan & Creelman, 1991).

Another strategy is to decide on the presence of a compound signal if either of the two components exceeds its own decision threshold (Fig. 3C). This rule can be regarded as a minimalistic approach to cue integration: The benefit from having a double-cue signal is merely due to having two chances to discriminate the target from the non-target. This decision rule is therefore called *probability summation* (*PS*).

Finally, a more efficient way to discriminate the signal and noise distributions would be to use the linear classifier shown in Fig. 3D. This strategy is called the *integration rule* (Macmillan & Creelman, 1991), or *information summation* (*IS*; Madison, Thompson, Kersten, Shirley, & Smits, 2001). This classifier is the optimal linear classifier under the assumption of equal variance between the two distributions. The IS rule maximizes the hit rate (answering 'signal' when the trial was actually a signal trial), while minimizing the false alarm rate (answering 'signal' when the trial was actually a noise trial).

2.4. Task

We measured shape detectability in a 2AFC task, with one interval showing a target stimulus defined by curvilinearity of contour elements (single-cue contour condition), isolinearity of surface elements (single-cue surface condition), or a combination of both (double-cue condition), and the other interval showing a non-target stimulus with no visual shape.

Each trial consisted of a fixation cross, followed by a target and a non-target stimulus, each presented for 150 ms, and separated by 300 ms visual masks (Fig. 4). The order of target and non-target stimuli was random. After each trial the participant pushed a button to indicate the target interval containing the visual shape. Auditory feedback was provided after each response. Stimuli were presented in blocks of 50 trials. All trials within a block belonged to the same condition. At the onset of each block participants were informed about the condition of trials within that block. The order of blocks was fully randomized.

2.5. Calibration of detection thresholds

Prior to the main experiment, detectability of the two single-cue conditions was individually calibrated. For each participant we ensured that the single-cue contour and single-cue surface stimuli were equally detectable. In Experiment 1 we equated the performance (percentage correct) for the two single-cue conditions by adjusting the number of contour elements. In Experiment 2 we used a staircase procedure to estimate the amount of orientation jitter required to yield a pre-specified detection threshold. This was done separately for the two single-cue conditions. The outcome of this calibration (number of contour elements for Experiment 1 and orientation jitter for Experiment 2) was used to generate the stimuli for the main experiment. For the double-cue stimuli the orientation jitter was a pairwise combination of the orientation jitter added to the single-cue stimuli. Having equally detectable cues avoids a vetoing strategy, and encourages observers to exploit both shape cues in the double-cue condition. A similar approach was used by Meinhardt, Persike, Mesenholl, and Hagemann (2006) to study the combination of spatial frequency and orientation cues.

2.6. Estimating the probability summation and information summation rules

The performance in the two single-cue conditions was used to predict the performance on the double-cue trials when observers apply the PS rule or the IS rule.

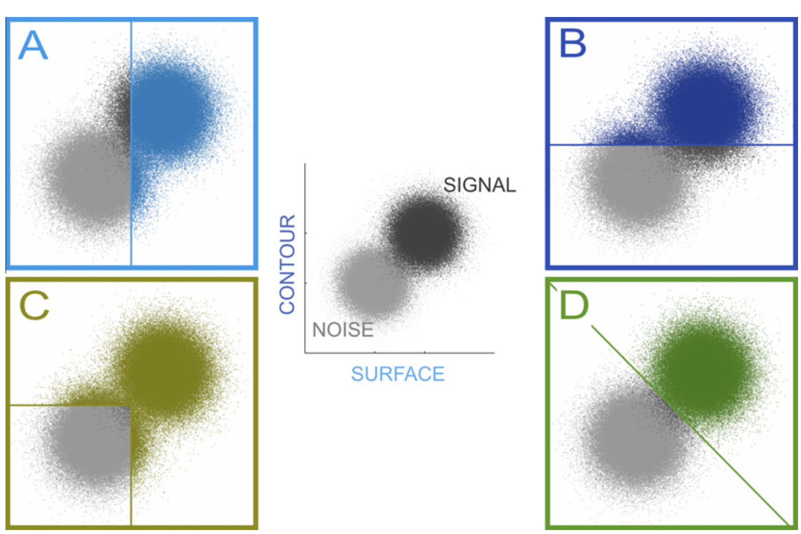

Fig. 3. To discriminate the signal distribution from the noise distribution (central panel) different decision rules can be employed. *Vetoing*: one single cue is used (A: surface; B: contour) to decide on the presence of a signal (colored area). *Probability summation* (*PS*): a signal is perceived if either cue exceeds its own decision threshold (C). *Information summation* (*IS*): the diagonal line is the optimal linear classifier between the signal and noise distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

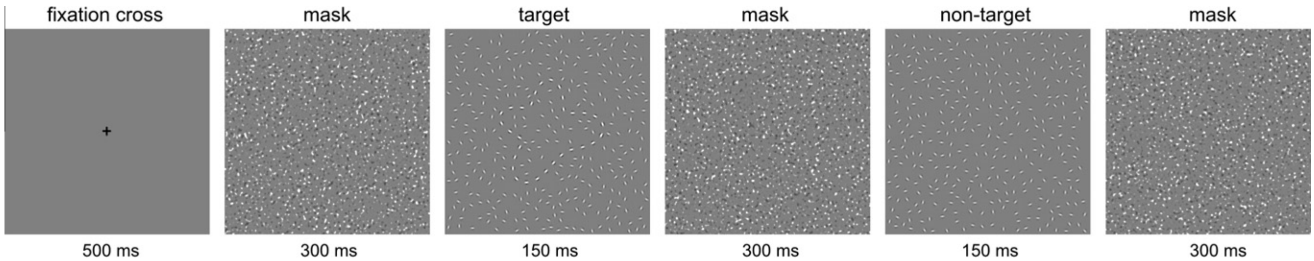

Fig. 4. Time course of a single trial. We increased the contrast of the contour elements in the target stimulus for illustrative purposes.

Under the *PS rule*, the expected hit rate H in the double-cue condition was calculated as $1-(1-H_c)(1-H_s)$, and the expected false alarm rate F was calculated as $1-(1-F_c)(1-F_s)$, with H_c and H_s the observed hit rates for the contour and surface condition, respectively (Macmillan & Creelman, 1991). With a known number of signal trials (N_s) and noise trials (N_n), the predicted performance under the PS rule was calculated as $\frac{H \cdot N_s + (1-F) \cdot N_n}{N_s + N_n}$.

Under the *IS rule*, the expected signal detectability (d') in the double-cue condition was calculated as $\sqrt{(d'_c)^2 + (d'_s)^2}$, with $d'_c = \frac{1}{\sqrt{2}}[z(H_c) - z(F_c)]$ detectability in the contour condition, and $d'_s = \frac{1}{\sqrt{2}}[z(H_s) - z(F_s)]$ detectability in the surface condition. Assuming equal variance in the signal and noise distributions, $\Phi(d'/\sqrt{2})$ transformed the predicted d' value to a proportion correct (Macmillan & Creelman, 1991).

2.7. Participants

Fourteen participants volunteered for this study. All had normal or corrected-to-normal vision. Eight collaborators from our lab participated in Experiment 1, five undergraduate students and the first author participated in Experiment 2. All participants, except the first author, were unaware of the purpose of the experiment. Each participant gave informed consent for the study. The study was approved by the KULeuven Ethics Committee.

3. Experiment 1

3.1. Procedure

In the first experiment we added the same amount of orientation jitter to the contour elements and to the surface elements. Nine different levels of orientation jitter were used, sampled from a Gaussian distribution around 0 with a standard deviation of 15, 25, 30, 35, 40, 45, 50, 55, or 65 degrees of arc. To ensure equal detectability in the contour and surface conditions, we ran a calibration procedure in which the number of contour elements was individually adjusted. We kept the number of surface elements constant (30), while varying the number of contour elements between 26 and 34. Within this limited range we could avoid differences in local density between surface, contour and background elements. Increasing the number of contour elements improves the grouping of the contour elements (Field et al., 1993; McKendrick, Weymouth, & Battista, 2010). After a short demo, training phase, and calibration, the main experiment was run with 2700 trials per participant (9×100 per condition).

3.2. Analysis

A logistic regression analysis predicting detection performance with condition and log-transformed orientation jitter as fixed effects was performed on the data. The analysis allowed for a random intercept and random slope for each participant. To account for occasional lapses the upper bound $(1-\lambda)$ of detection performance was estimated from the data, with λ restricted between 0 and 0.05. In a first analysis we compared the two single-cue conditions with the double-cue condition. The double-cue condition was compared with the PS prediction and with the IS prediction in two separate analyses. For illustrative purposes, the results from all three analyses are plotted in a single graph. Whenever the interaction between condition and orientation jitter did not reach significance, we repeated the analysis without the interaction term. All analyses were performed with SAS procedure NLMIXED (SAS version 9.2).

3.3. Results

Fig. 5 displays the observed proportion correct for the two single-cue conditions and the double-cue condition. Regression lines were plotted through the data points. To emphasize the predictive nature of the IS rule and PS rule, we did not plot the estimated data points, but only the regression through these points.

No difference between the two single-cue conditions was found ($F(1,6) = 0.04$; $p = 0.85$), reflecting the validity of the calibration method. Performance in the double-cue condition was significantly better than in the two single-cue conditions ($F(1,6) = 48.21$;

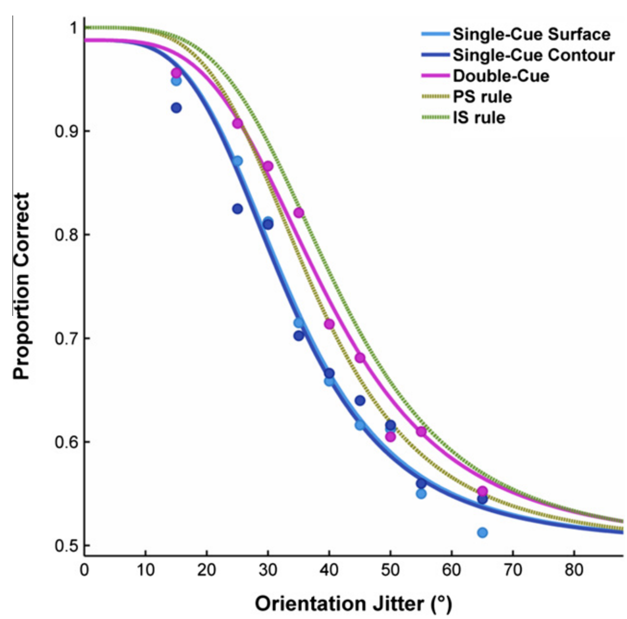

Fig. 5. Results from Experiment 1. Proportion correct as a function of orientation jitter for the two single-cue conditions and the double-cue condition, along with the predictions from the PS and IS rule. The plotted data points are the observed proportion correct for the three conditions, averaged over participants ($n = 8$). Regression lines are fitted through these data points. For the PS and IS rules the predicted data points are not plotted.

$p < 0.001$). The double-cue condition yielded performance levels which were in-between the two combination rules, but did not differ significantly from either the PS rule ($F(1, 6) = 2.76$; $p = 0.15$) or the IS rule ($F(1, 6) = 3.59$; $p = 0.11$).

The results from Experiment 1 show a clear double-cue benefit. Performance in the double-cue condition is significantly better than in the two single-cue conditions. This indicates that congruent information from the contour and the interior surface of a shape is somehow combined to detect the shape in a cluttered background. However, from the observed double-cue benefit we cannot estimate how well the two cues are combined. For that, we need benchmarks against which to compare the double-cue performance. The minimalistic PS rule and the optimal IS rule, calculated from the single-cue data, provide such benchmarks. The double-cue performance falls in-between the predictions from the PS and IS rule, but the results from Experiment 1 do not allow us to differentiate between the two combination rules.

4. Experiment 2

To better discriminate between the two cue-combination rules, we ran a second experiment, with a more flexible calibration procedure, and more trials per data point.

4.1. Procedure

In Experiment 1 the performance on contour and surface trials was equated by varying the number of contour elements. In Experiment 2 the number of contour elements was kept constant at 30 Gabors. To calibrate performance on surface and contour trials we now allowed different jitter levels for surface and contour elements. By manipulating orientation jitter we could employ a more flexible adaptive procedure.

In Experiment 1 we sampled the entire psychometric function. In order to better discriminate between the two cue-combination rules, we now only focus on the central region of the psychometric function. For each participant we search for these jitter levels that yield detection thresholds of 85%, 75%, and 65%. This is done separately for the two single-cue conditions, using an adaptive procedure (Watson & Pelli, 1983) with two interleaved staircases involving 250 trials for each threshold and each condition.

In the main experiment each participant ran 4500 trials (500 trials per data point). For each detection threshold, the amount of orientation jitter obtained in the calibration experiment was used. For the double-cue stimuli, the orientation jitter added to the elements matched that of the single-cue stimuli.

4.2. Analysis

A logistic regression analysis predicting detection performance with condition and detection threshold as discrete predictors was performed on the data. The analysis allowed for a random intercept for participant. Again, we first compared the two single-cue conditions with the double-cue condition, and the double-cue condition with the estimates from the PS and IS rule.

4.3. Results

Fig. 6 displays the observed proportion correct for the two single-cue conditions and the double-cue condition, as well as the predicted proportion correct for the PS rule and the IS rule. Again, the difference between the two single-cue conditions was not significant ($F(1, 5) = 0.03$; $p = 0.87$), reflecting that performance on the single-cue trials was calibrated prior to the main experiment. The double-cue condition was again significantly better than the two single-cue conditions ($F(1, 5) = 42.00$; $p = 0.001$), and did not differ significantly from the IS rule ($F(1, 5) = 0.11$; $p = 0.76$). This time the double-cue performance exceeded the prediction from the PS rule ($F(1, 5) = 10.58$; $p = 0.02$). Note that the first author who participated in Experiment 2 was included in the above analysis. Repeating the analysis with only the five naive participants yielded similar results. Again, there was no significant difference between the two single-cue conditions ($F(1, 4) = 0.14$; $p = 0.73$). The double-cue condition was significantly better than the single-cue conditions ($F(1, 4) = 43.68$; $p = 0.003$), and better than the PS rule ($F(1, 4) = 14.57$; $p = 0.02$). The difference between the double-cue condition and the IS rule was not significant ($F(1, 4) = 0.34$; $p = 0.59$).

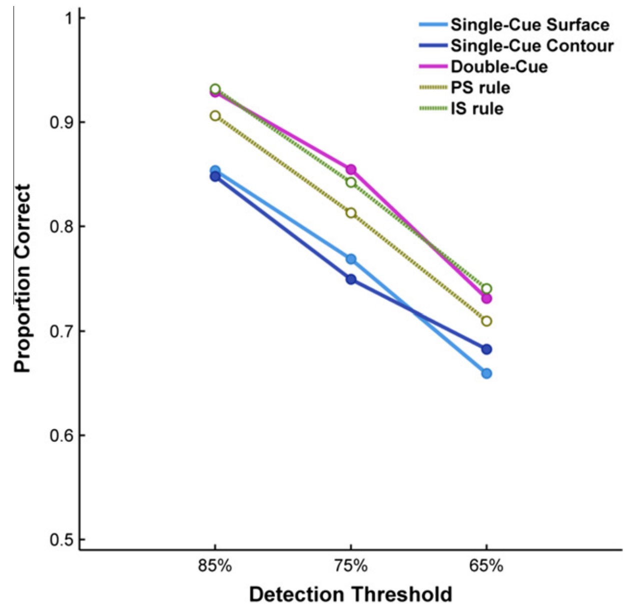

Fig. 6. Results from Experiment 2. Proportion correct as a function of detection threshold for the two single-cue conditions and the double-cue condition (filled circles, solid lines), along with the predictions from the PS and IS rule (open circles, dotted lines). The plotted data points are the observed proportion correct for the three conditions, averaged over participants ($n = 6$).

From the results of Experiment 2, it is possible to differentiate between the two cue-combination rules. Overall, observers do better than mere probability summation when presented with congruent surface and contour information. We find no evidence against the IS rule.

4.4. Discussion

The present study investigated to what extent human observers combine information from the contour and from the interior surface of a shape. Although contour detection and surface segregation are important steps towards shape perception, most studies have treated them in isolation. As already pointed out by Elder and Zucker (1998), separating the roles of contour and surface cues in shape detection is challenging because the two cues are usually present together. In our stimuli we could disentangle contour and surface information. Using the framework of signal detection theory, we were able to test how well human observers combine congruent contour and surface information.

Both experiments presented here showed a clear double-cue benefit: Shape detectability was significantly better when the two cues were presented together. The results from Experiment 1 did not allow us to differentiate between a probability summation rule and an information summation rule. In Experiment 2 we increased the number of trials per data point and used a more controlled approach to calibrate single-cue detectability. In this experiment we did find evidence that information about the surface and the contour of the embedded shape was combined more optimally than predicted by probability summation. We did not find evidence against optimal linear cue combination.

In the calibration phase that preceded the actual experiment, we adopted the approach of Meinhardt et al. (2006) to individually calibrate detectability in the two single-cue conditions. This enabled us to construct double-cue displays in which the two cues were equally detectable. Having equally detectable cues increases the potential gain of employing both cues. The necessity of an individual calibration approach was reflected in the interindividual variability in relative detectability of the two single cues. To equate detection performance for the two cues, the required number of contour elements in Experiment 1 ranged from 30 to 34 (with a fixed number of 30 surface elements). In Experiment 2, all participants could tolerate more orientation jitter to the surface elements than to the contour elements, probably due to the more peripheral position of contour elements (Nugent, Keswani, Woods, & Peli, 2003). However, the relative tolerance to orientation jitter differed a lot between participants. One participant could only tolerate 7 degrees of orientation jitter more on the surface elements than on the contour elements to reach the detection threshold, while another participant could tolerate an extra 22 degrees of orientation jitter on the surface compared to the contour elements.

One might argue that the surface and contour cues were only partially disentangled in our stimuli. Indeed, in the single-cue contour displays the contour elements still delineated the interior surface of the shape, even when the surface elements were randomly oriented. Similarly, in the single-cue surface displays, the isolinear interior elements could also give rise to the perception of an implicit boundary, even though the contour elements had random orientations. However, this limitation does not invalidate our main conclusions. Assuming that single-cue displays were indeed partly contaminated by the (implicit) presence of the other cue, implies that the measured detectability of a single cue is an overestimate of the true detectability. Hence, the predicted performance for the PS and IS rules would also be overestimated. When comparing the observed performance in the double-cue condition with the overestimated PS and IS predictions, we would still conclude that we have no evidence against optimal cue integration, and that performance in the double-cue condition is at least equal to (Experiment 1) or better than (Experiment 2) probability summation.

The double-cue benefit observed in this study shows that observers can combine surface and contour information to decide on the presence of a shape in the display. Our results do not permit us to draw strong conclusions about the mechanisms underlying this cue combination. Some authors have argued that surface and contour information is processed by distinct mechanisms (Elder & Zucker, 1998; Rogers-Ramachandran & Ramachandran, 1998). Our results are consistent with this view of independent processing, assuming that participants employ an optimal combination rule to integrate the two independent cues at a later stage of visual processing. However, recent computational and imaging studies rather suggest interactions between the processing of surface and contour information (Kogo et al., 2010; Scholte et al., 2008). If surfaces and contours are not processed independently, i.e. if their internal noise distributions are correlated, the performance predicted by an optimal linear combination rule under the independency assumption, can be reached by assuming a non-linear decision rule (in case of positively correlated noise distributions) or a suboptimal linear decision rule (in case of negatively correlated noise distributions). The present study does not allow us to differentiate between these possibilities.

In short, our results show that participants do combine congruent surface and contour information to decide on the presence of a shape, at least when the individual cues are matched for detectability, and that this combination is as good as (Experiment 1) or better than (Experiment 2) predicted by a probability summation rule.

5. Conclusion

By using a controlled set of artificial stimuli in which information about the contour and interior surface was disentangled, we were able to show that human observers use both sources of information to detect a visual shape. The framework of signal detection theory provided benchmarks against which to evaluate the performance in the double-cue condition. We have shown that human observers can integrate congruent information from the interior surface of a shape and its surrounding contour more optimally than predicted by a probability summation rule. We did not find evidence against optimal linear combination of surface and contour information.

Acknowledgments

We wish to thank L. Cuturi for assistance in the data collection, and C. Casco, R. Watt, T. Putzeys, and two anonymous reviewers for valuable comments on an earlier version of the manuscript. This research was supported by a Methusalem Grant from the Flemish Government (METH/08/02).

References

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

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, 5*, 1749–1758.

Claessens, P. M., & Wagemans, J. (2005). Perceptual grouping in Gabor lattices: Proximity and alignment. *Perception & Psychophysics, 67*, 1446–1459.

Elder, J., & Zucker, S. (1998). Evidence for boundary-specific grouping. *Vision Research, 38*, 143–152.

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*, 173–193.

Fulvio, J. M., & Singh, M. (2006). Surface geometry influences the shape of illusory contours. *Acta Psychologica, 123*, 20–40.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research, 41*, 711–724.

Giora, E., & Casco, C. (2007). Region- and edge-based configurational effects in texture segmentation. *Vision Research, 47*, 879–886.

Harrison, S. J., & Feldman, J. (2009). The influence of shape and skeletal axis structure on texture perception. *Journal of Vision, 9*(6), 1–21 (article no. 13).

Harrison, S. J., & Keeble, D. R. T. (2008). Within-texture collinearity improves human texture segmentation. *Vision Research, 48*, 1955–1964.

Hess, R. F., Beaudot, W. H. A., & Mullen, K. T. (2001). Dynamics of contour integration. *Vision Research, 41*, 1023–1037.

Hess, R. F., Hayes, A., & Field, D. J. (2003). Contour integration and cortical processing. *The Journal of Physiology, 97*, 105–119.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290*, 91–97.

Kogo, N., Strecha, C., Van Gool, L., & Wagemans, J. (2010). Surface construction by a 2-D differentiation–integration process: A neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. *Psychological Review, 117*, 406–439.

Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure–ground segmentation. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 7495–7497.

Lamme, V. (1995). The neurophysiology of figure–ground segregation in primary visual cortex. *The Journal of Neuroscience, 15*, 1605–1615.

Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research, 38*, 2429–2454.

Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear interactions. *Journal of Neurophysiology, 88*, 2846–2856.

Machilsen, B., Pauwels, M., & Wagemans, J. (2009). The role of vertical mirror symmetry in visual shape detection. *Journal of Vision, 9*(11), 1–11 (article no. 12). Available from http://journalofvision.org/9/12/11.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Madison, C., Thompson, W., Kersten, D., Shirley, P., & Smits, B. (2001). Use of interreflection and shadow for surface contact. *Perception & Psychophysics, 63*, 187–194.

McKendrick, A. M., Weymouth, A. E., & Battista, J. (2010). The effect of normal aging on closed contour shape discrimination. *Journal of Vision, 10*(1), 1–9 (article no. 2).

Meinhardt, G., Persike, M., Mesenholl, B., & Hagemann, C. (2006). Cue combination in a combined feature contrast detection and figure identification task. *Vision Research, 46*, 3977–3993.

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31*, 1073–1078.

Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics, 52*, 355–375.

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

Palmer, S. E., & Brooks, J. L. (2008). Edge-region grouping in figure–ground organization and depth perception. *Journal of Experimental Psychology: Human Perception & Performance, 2008*(34), 1353–1371.

Pettet, M. W. (1999). Shape and contour detection. *Vision Research, 39*, 551–557.

Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research, 38*, 865–879.

Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H., & Bosch, H. (2002). Figure–ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience, 14*, 525–537.

Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research, 38*, 71–77.

Rubin, E. (1921). *Visuell Wahrgenommene Figuren: Studien in psychologischer Analyse*. Copenhagen: Glydenalske.

Scholte, H. S., Jolij, J., Fahrenfort, J. J., & Lamme, V. A. F. (2008). Feedforward and recurrent processing in scene segmentation: Electroencephalography and functional magnetic resonance imaging. *Journal of Cognitive Neuroscience, 20*, 2097–2109.

Stanley, D. A., & Rubin, N. (2003). FMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron, 37*, 323–331.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.

Watt, R., Ledgeway, T., & Dakin, S. C. (2008). Families of models for Gabor paths demonstrate the importance of spatial adjacency. *Journal of Vision, 8*(23), 1–19 (article no. 7).

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung, 4*, 301–350.