Vision Research 49 (2009) 2475–2484

Contents lists available at ScienceDirect

Vision Research

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

The influence of orientation jitter and motion on contour saliency and object identification

Geir Eliassen Nygård, Tina Van Looy, Johan Wagemans *

Laboratory of Experimental Psychology, University of Leuven, Belgium

ARTICLE INFO

Article history:
Received 6 December 2008
Received in revised form 22 July 2009

Keywords:
Perceptual grouping
Figure-ground segregation
Shape-from-motion
Gabor
Good continuation
Motion perception
Shape perception
Object perception

ABSTRACT

One of the ultimate goals of vision research is to understand how some elements are grouped together and differentiated from others to form object representations in a complex visual scene. There exists an extensive literature on this grouping/segmentation problem, but most of the studies have used un-recognizable stimuli that have little to do with object recognition per se. We used Gabor-rendered outlines of real-world objects to study some relationships between bottom-up and top-down processes in both spatial- and motion form perception. We manipulated low-level properties, such as element orientation and local motion, while incorporating higher-level properties, such as object complexity and identity, and found that adding local motion improved overall performance in both object detection and object identification tasks. Adding orientation jitter effectively decreased object detection performance in both static and motion conditions, and increased reaction time for identification in the static condition. Orientation jitter had much less effect on reaction times for identification in the local motion condition than in the static condition. Both contour properties ("good continuation") and object properties (identifiability) had a positive effect on detection and reaction time for identification.

1. Introduction

A fundamental goal of vision is to locate, characterize, and recognize objects. To determine "what" is "where", the visual system must first determine which parts of the image belong together in groups that can be segmented from the background. This is known as the image grouping/segmentation problem, and one of the most important tasks is the extraction of object contours. Since contours are often not well defined along all of their extent (due to partial occlusion), the visual system needs to be able to infer their nature from an incomplete representation. It can make use of several cues to construct a coherent percept, for instance, texture gradients, color, depth information, occlusion, and motion. How the brain combines local information into a global structure, how it computes form from these cues, remains an important issue in visual neuroscience. Most studies trying to address this subject have used simple, non-object stimuli in detection or discrimination tasks, where identification of objects was not necessarily needed. In an attempt to reduce the gap between the extensive literature on grouping/segmentation and object perception (where object identification has to take place), we introduce a new set of stimuli where we use Gabor-rendered outlines of real-world objects. We will first review some of the background on contour integration and form perception, and then present two experiments that explore the influence of jitter and motion on both the detectability and identifiability of real-world object contours.

Several studies have looked at contour integration in both static and motion conditions. In a paper on the role of temporal modulation in visual contour integration, Bex, Simmers, and Dakin (2001) compared the detectability of "snakes" and "ladders" in relation to orientation jitter on the contour elements. "Snakes" are constituted by Gabor elements that are locally aligned with the contour axis, while "ladders" are constituted by elements perpendicular to its axis. For a contour containing six elements placed along a path with an angle of 20° between the segments, they found that the amount of jitter that could be tolerated was approximately doubled when they added motion to the Gabor elements (translation of the carrier sine wave). In a later study, Bex, Simmers, and Dakin (2003) used a different paradigm (where the contour elements consisted of moving dots) and found that, as with static contour images, the visibility of moving contours decreases at high curvature, albeit by less than in the static case. It has also been found that counter-phase temporal flicker can enhance contour detection (Bex et al., 2001). In exploring the influence of spatial frequency and orientation on motion-defined contours, Ledgeway and Hess (2006) found a very broad tuning for spatial frequency, and a relatively narrow tuning for orientation. Motion direction tuning is in comparison relatively broadband (Allman, Miezin, & McGuinness,

* Corresponding author. Address: University of Leuven, Department of Psychology, Tiensestraat 102, B-3000 Leuven, Belgium. Fax: +32 16 32 60 99.
E-mail address: johan.wagemans@psy.kuleuven.be (J. Wagemans).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.08.002

1985; Bex et al., 2001; Bex et al., 2003; Ledgeway & Hess, 2002, 2006).

Traditionally, mechanisms involved in object recognition and mechanisms for encoding object position and motion have been assumed to project ventrally and dorsally from primary visual cortex to infero-temporal and posterior-parietal cortex, respectively (DeYoe & Van Essen, 1988; Livingstone & Hubel, 1987; Ungerleider & Desimone, 1982), constituting the so-called "ventral" and "dorsal" streams, respectively. Such a strong functional distinction between these two pathways is disputed however: several cortical areas in both the ventral and dorsal streams have been implicated in shape-from-motion (Braddick, O'Brien, Wattam-Bell, Atkinson, & Turner, 2000; Murray, Olshausen, & Woods, 2003) and MT/MST, a typical dorsal area dedicated to motion processing, also appears involved in the analysis of object shape (Kourtzi, Bülthoff, Erb, & Grodd, 2002). Moreover, several studies point to a third stream projecting from V1 to lateral occipito-temporal cortex (LOC) that also underlies complex motion perception (Ferber, Humphrey, & Vilis, 2003, 2005; Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Murray et al., 2003). Evidence suggests that the streams consist of a hierarchy of processing stages that transform lower-order stimulus properties into higher-order primitives (Grill-Spector et al., 1998), and anatomical work has revealed reciprocal inter-stream connections at all levels of the visual hierarchy (Felleman & Van Essen, 1991; Van Essen & Maunsell, 1983).

Several studies have set out to find how spatial form and motion form, presumed to be implemented by two independent systems (Ledgeway & Hess, 2006; Lorenceau & Alais, 2001; Rainville & Wilson, 2004), interact when they are presented simultaneously. Lorenceau and Alais (2001), for instance, showed that binding local motions into global object motion depends on spatial form (open vs. closed contour configurations). They suggested that this influence arises in early cortical levels where a spatial-form-based veto of motion integration occurs in the absence of closure. Rainville and Wilson (2004, 2005), on the other hand, argued that the interference is a result of processes further up in the hierarchy involving curvature extraction or overall shape.

The stimuli that have been used in most of the above studies have been relatively simplistic in nature, consisting of geometric figures (squares, circles, polygons), parametric contours (radial frequency patterns), snake-like contour segments, or dot patterns, with little, if any, biological significance. These studies are invaluable because they allow parametric control on all low-level aspects, while restricting the variability and complexity of the shape properties. The downside is, however, that they do not allow to extrapolate the findings to more complex, natural shapes and to link the grouping/segmentation processes to higher-order processes such as object recognition. In an attempt to extend the topics of the ongoing discussion to a domain that is more directly related to object perception, we introduce a new set of stimuli where we use Gabor-rendered outlines of real-world objects.[1] This will enable us to manipulate low-level properties that can be used in models of contour perception, while incorporating higher-level object properties such as complexity and identity. By adding motion we can then study both spatial form and motion form using more familiar stimuli.

Of course, we do not mean to suggest that one cannot study high-level shape and contour processing with synthetically generated stimuli that are not derived from real-world objects. On the contrary, quite interesting shape perception work has made use of synthetic shapes consisting of combinations of Fourier dimensions (e.g., Cortese & Dyre, 1996), well-controlled contours generated by radial frequency patterns (e.g., Bell, Badcock, Wilson, & Wilkinson, 2007; Wilkinson, Wilson, & Habak, 1998), or even synthetic faces (e.g., Wilson, Loffler, & Wilkinson, 2002). Indeed, we have made use of similar shapes with variable levels of complexity and symmetry in our own work (e.g., Kayaert & Wagemans, 2009; Machilsen, Pauwels, & Wagemans, in press; Op de Beeck, Wagemans, & Vogels, 2003). What we do mean to suggest is that stimuli derived from real-world objects probably also induce some extra processing in the highest levels in the visual hierarchy where contact is made with representations of existing objects and associations with other items in semantic memory become available too. Investigation of how good continuity, orientation jitter, local motion, etc. affect the visual processing at these highest levels of the visual hierarchy, and vice versa, requires the kind of stimuli we introduce here.

In sum, such stimuli will enable us to answer one of the original questions of Gestalt psychology: does grouping help identification when the contour represents a familiar object (Wertheimer, 1938)? And vice versa? In Experiment 1, we examine the role of static and dynamic grouping in the detection of these complex shapes and we include good continuation and identifiability as additional variables of interest. In Experiment 2, we explicitly ask to what extent static and dynamic grouping influence the identification of these Gaborized outlines. In addition to addressing the two-way linkage between grouping and identification, we use these stimuli to be able to generalize the earlier findings obtained with simpler, parametrically controlled stimuli to more complex, natural shapes. In a similar way as for snake detection paradigms, we expect objects with more curved segments to be more difficult to both detect and identify, and that adding orientation jitter will degrade performance in both tasks. Furthermore, we expect motion to enhance both detection and identification of our object contours. We also expect to see an influence of global object properties, for instance, a positive effect for objects that are easier to identify.

2. Experiment 1

In Experiment 1 we asked for detection of objects in a noisy background, defined by contours with variable jitter levels on the orientation of the constituent elements. The Gabor elements were either static or had "local motion", i.e., the phase of the Gabors was translated as to give a motion direction perpendicular to the element's orientation while the Gaussian envelope remained stationary. The objective was to see if and how jitter and motion interact in determining detection.

2.1. Subjects

Subjects ($N = 6$) were three male and three female, aged 20–24, with normal or corrected-to-normal vision. One subject is the second author, the remaining subjects were recruited in the general student population for a paid (per hour) participation, and were naïve regarding the purpose and the details of the experiment.

2.2. Stimuli

The stimuli consisted of Gabor elements that were placed and oriented such that they gave rise to the percept of an object embedded in a background (Fig. 1). The objects were contour versions of 60 items from the Snodgrass and Vanderwart (1980) set of line drawings, which we had first converted into silhouettes and then into outlines (De Winter & Wagemans, 2004; Wagemans et al., 2008). Of course, contours of real-world objects differ on a

[1] We have also performed a parallel series of experiments where we used kinetic dot versions of the same objects/stimuli to look at related questions (see Segaert, Nygård, & Wagemans, 2009).

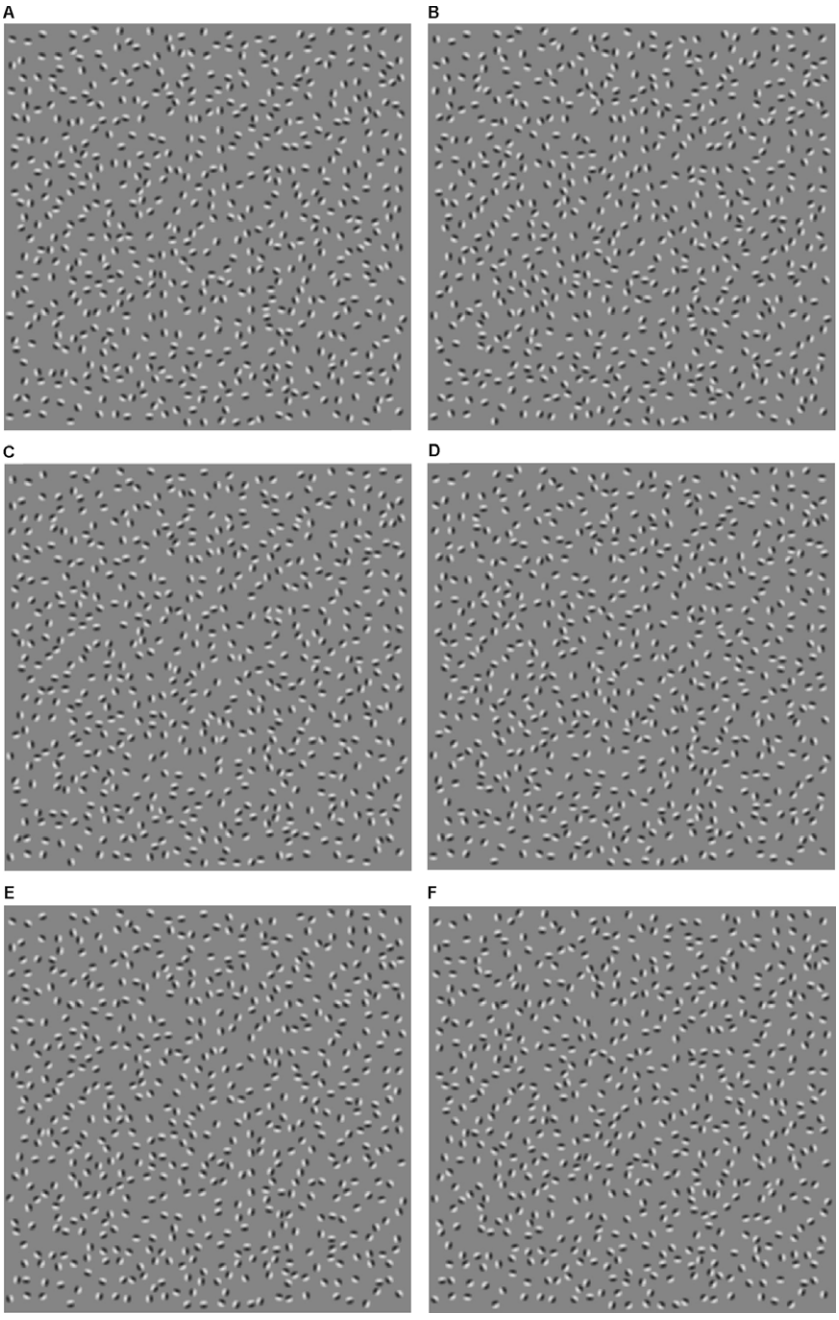

Fig. 1. Examples of the stimuli used in the experiment (the contrast in these examples is set to 100% for clarity). Panels A–F present the same object, a bottle, at different jitter intervals: 0, 60°, 90°, 120°, 180°, 360°, respectively. The 360° jitter interval condition corresponds to what we used as our control stimuli.

number of factors that may have an influence on their perception. We controlled for the overall complexity of the shape (represented by its compactness; see Appendix A) and two variables of interest: the "good continuation factor" of the contour (GCF; see Appendix B), and the identifiability of the object (ID; measured in a separate experiment where subjects were asked to name Gabor-rendered object contours (collinear contour elements embedded in a random background, same as in Fig. 1A) presented for 5 s; see see Nygård & Wagemans, submitted for publication). Our measure for compactness theoretically spans the interval]0, 1] (from an infinitely complex object, which asymptotically has zero compactness, to a circle, which has a compactness of 1), while the compactness for our stimulus set was 0.28 ± 0.13 (mean ± standard deviation; see Fig. A1 for examples). The GCF spanned a range of 0.082–1.85, with a mean and standard deviation of 0.792 ± 0.393 (see Fig. B1 D for examples), while ID ranged from 0 to 1 with a mean and standard deviation of 0.22 ± 0.23.

We made the stimuli by placing Gabor elements on the contour, and then inside and around the contour (for more details, see Nygård & Wagemans, submitted for publication). In short, the process consisted of centering the object on a constant size square grid, and filling-in empty cells with more elements. Contour elements were equally spaced along the length of the contour, with an element separation of 1.86 times the wavelength of the Gabor elements (ES = 1.86). We chose a cell size (CS) in function of element separation so that, on average, there would be approximately one element per cell:

$$CS = \frac{2 * ES}{1 + \sqrt{2}} \quad (1)$$

Local density cues were avoided by using the same distribution for object- and background elements. We achieved this by copying and pasting the position of elements in a random cell containing contour elements to an empty cell. The number of elements in each cell would vary with the length of the contour segment in that cell. Furthermore, segment length is correlated with complexity: the more complex the segment, the longer it is, and therefore the more room to place elements. Since the objects had a big range in complexity, this gave rise to some variability in the number of elements in the display. We had a total of 602 ± 19.8 elements in the images, of which 42.3 ± 12.9 were Gabors pertaining to the contour (mean ± standard deviation).

We made sure that the elements did not overlap as a result of the copy-and-paste procedure (this could potentially happen at the borders of the cells). This was done through iterative steps: in case of overlap, we would randomly choose a different cell containing contour elements. This was repeated until there was no overlap, or all the contour elements had been sampled. The background cell was left empty if overlap could not be avoided. To supervise the quality of our stimuli we counted the number of empty cells, and calculated the distances between contour, surface, and noise patches (within and between groups). The allowed number of empty cells was limited to 1% of the total number of cells and the difference in average distance within and between the groups of patches was limited to 5%. Stimuli exceeding these criteria were rejected and only valid stimuli were used in the experiment.

Gabors can conceptually be seen as a sine wave in a Gaussian envelope. They were odd symmetric and defined by:

$$g(x, y, \theta) = \sin(2\pi f(x \sin\theta + y \cos\theta)) * e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

where (x, y) is the distance from the element center, θ is the element orientation, f is the spatial frequency, and σ is the space constant. We chose a spatial frequency of 2 cycles per degree (cpd), a space constant equal to a fourth of the wavelength, and a Michelson contrast of 50%. The elements were placed on a uniform gray background, and the average luminance of the display was 25 cd/m^2.

We manipulated the perceptual salience of the contour by adding jitter to the orientation of the elements positioned on the contour. In the most salient condition there was no jitter at all. The orientation was then equal to that of the tangent on the local contour where the element was positioned. We added jitter by taking a random sample from a uniform distribution spanning 60°, 90°, 120°, 180°, or 360°[2] for experiment 1, centered on the orientation of the local tangent (see Fig. 2). A condition where all the elements had a random orientation was used as a control (in essence this is the same as the condition above where 360° of jitter was added; see Fig. 1 for examples of our stimuli). All Gabor elements, i.e., those pertaining to contour, background, and control stimuli, were either static or had "local motion" (i.e., the phase of the Gabors was translated as to give a motion direction perpendicular to the element's orientation (speed = 3.5 deg/s) while the Gaussian envelope remained stationary), giving a total of 12 different conditions (six jitter levels for both static and motion conditions).

2.3. Procedure

Stimuli were generated offline and displayed on a Sony computer monitor with a refresh rate of 85 Hz and at a resolution of 1024 × 768 pixels. The monitor was calibrated, linearized, and checked using a Minolta 110 L spot meter. The stimuli subtended 14 × 14° of visual angle, and were presented in a two sequential

[2] Note that, in the case of jitter values greater than or equal to 180°, this transformation can also be seen as a change in both spatial phase and orientation of the Gabor elements.

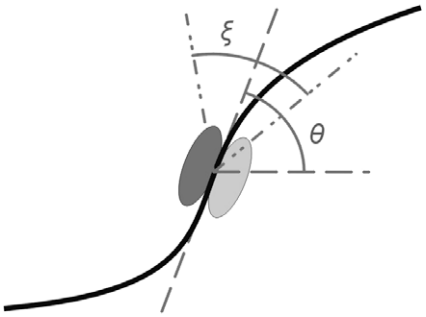

Fig. 2. Illustration of how the orientation of the Gabor elements on the contour was set. θ is the orientation of the local tangent of the contour, and ξ represents the interval in which jitter angles could be chosen (ξ could be 0°, 60°, 90°, 120°, 180°, or 360°, centered on θ).

intervals of 941 ms, preceded by a fixation cross for 500 ms and with an inter-stimulus-interval of 500 ms. Only one of the intervals contained the contour of an object and the other contained a control stimulus. The subject was positioned with a chin rest 57 cm from the monitor in a darkened room, and their task was to indicate in which interval there was an object present (2-alternative-forced-choice or 2-AFC task). The 12 conditions were repeated two times each for all 60 contour stimuli for a total of 1440 trials per subject. Static and motion conditions were blocked, while all jitter conditions were randomized within blocks. Total duration was approximately 3 h, and breaks were allowed at the subject's leisure. Auditory feedback was given (high and low tone for correct and for wrong answers, respectively).

2.4. Results

Fig. 3 shows the data for all subjects, plotted as performance against jitter interval for both the static (open symbols) and the local motion (filled symbols) condition. We fitted a logistic function to this data to model object detection performance, Ψ, using the following formulae:

$$\Psi(X_1) = \gamma + (1 - \gamma) * F(X_1) \quad (3)$$

$$F(X_1) = \frac{e^{\beta_0 + \beta_1 X_1}}{1 + e^{\beta_0 + \beta_1 X_1}} \quad (4)$$

where X_1 is the jitter interval, Y is a design parameter reflecting the minimum performance (or chance level, equal to 0.5), F is a logistic

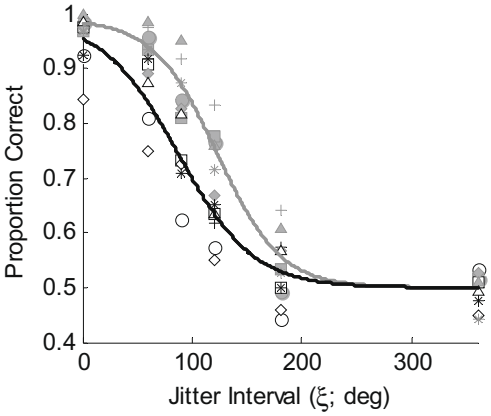

Fig. 3. Plot of the object detection performance as a function of the jitter interval. Different symbols represent different subjects; open symbols and the black line reflect the data and fitted model (Eqs. (3) and (4)), respectively, for the static Gabor condition, while filled symbols and the grey line reflect the data and fitted model for the local motion Gabor condition.

function where β_0 and β_1 reflect the intercept and the slope of the function. The fits are shown for both static (black line) and local motion (grey line) conditions. Visual evaluation suggests that there is a main effect of jitter interval on detection performance: the more jitter, the poorer the performance, eventually dropping to chance level as the level of jitter approaches 180–200°. In addition, it seems that local motion enhances object detection performance by approximately 5–15% compared to the static condition.

In a statistical analysis we performed a logistic regression on a model containing the parameters manipulated in the paradigm (jitter interval and static vs. motion Gabors), plus the two additional stimulus parameters GCF and ID. The logistic function, F, was now of the form:

$$F(X_1,X_2,X_3,X_4) = \frac{e^{\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4}}{1+e^{\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4}} \quad (5)$$

where X_1, X_2, X_3, and X_4 are the variables for static vs. local motion condition (categorical variable), jitter interval, "good continuation factor" (GCF), and object identifiability (ID). β_0 reflects the intercept of the function, and β_1 to β_4 are the regression parameters to be estimated for the respective variables. Fig. 4 presents detection performance in relation to our stimulus parameters GCF and ID. All the parameters were significant, as can be seen in Table 1.

This analysis assumes that there are no outliers in terms of subject performance, i.e., that there are no significant individual differences. When we look at Fig. 3, we see that there are slight variations in different subjects' maximum performance (at jitter = 0) for the static condition, something that would influence the shape of a fitted psychometric function. To test for the importance of these variations, we studied if the variance of the intercept and slope of the individual subjects significantly differs from that of the mean by performing a multilevel analysis. We used the same model as above, but where the power of the exponential was modified to separately include a random effect a_0 for the intercept and a_1 to a_4 for the slopes. Below is an example of how this was implemented for a_0 (for a_1, a_2, a_3, and a_4: proceed as for a_0 by substituting the respective β_n by $(\beta_n + a_n)$):

$$F(X_1,X_2,X_3,X_4) = \frac{e^{(\beta_0+a_0)+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4}}{1+e^{(\beta_0+a_0)+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4}} \quad (6)$$

a_0 to a_4 are drawn from a normal distribution N with a mean equal to zero and an estimated variance. The results are presented in Table 2: none of the random effects reached a significant level (smallest p-value was 0.1735), an indication that there were no statistically significant individual differences. Pooling across subjects, therefore, seems justified.

Table 1
Overview of the parameter estimates for the main effects and some of their statistical properties, based on the model described in Eqs. (3)–(5). All the main parameters are highly significant.

	DF	Estimate (β)	Std. error	t-Value	$p > \|t\|$
Intercept	8640	1.4154	0.1535	9.22	<.0001
Stat. vs. Mot.	8640	1.2267	0.116	10.57	<.0001
Jitter	8640	−0.03135	0.001448	−21.65	<.0001
GCF	8640	1.2612	0.1599	7.89	<.0001
ID	8640	0.9066	0.2579	3.51	0.0004

Table 2
Overview of the estimates of the random effects for the intercept (a_0) and slopes (a_1 to a_4) of the logistic function, and some of their statistical properties, based on the model described in Eq. (6). Note that none of the parameters reached significant values (smallest p-value was 0.1735).

	DF	Estimate	Std. error	t-Value	$p > \|t\|$
a_0	5	0.2103	0.1326	1.59	0.1735
a_1	5	0.000013	8.66E-06	1.45	0.2065
a_2	5	0.1628	0.1140	1.43	0.2126
a_3	5	0.6140	0.4491	1.37	0.2298
a_4	5	0.000036	0.000024	1.52	0.1883

As mentioned in the visual evaluation of Fig. 3, and as confirmed in Table 1, the addition of motion enhances object detection performance. In Fig. 3, it would seem that the enhancement is slightly more pronounced at the intermediate jitter levels (i.e., 90 and 120). To test this we added second-order interaction effects to our model and found that the interaction between jitter and motion was not statistically significant ($p = 0.4179$ in the full model). When we perform a stepwise backward regression on the full model, all interactions except the one between jitter and GCF fall away (Table 3; $p < 0.0001$). We therefore have no evidence to support that the performance increase by adding motion is different for different jitter values.

Table 4 presents the odds ratios, equal to the exponential of the β-values, from the former analysis (presented in Table 3). Odds ratios (OR) are multiplicative factors that show by how much the odds of correct detection increase (if OR > 1) or decrease (if OR < 1) for each unit of increase in the corresponding variable. For instance, the estimated OR for the motion categorical variable is 3.41, meaning that, after control for jitter interval, GCF, and ID, going from static to local motion Gabors increases odds of object

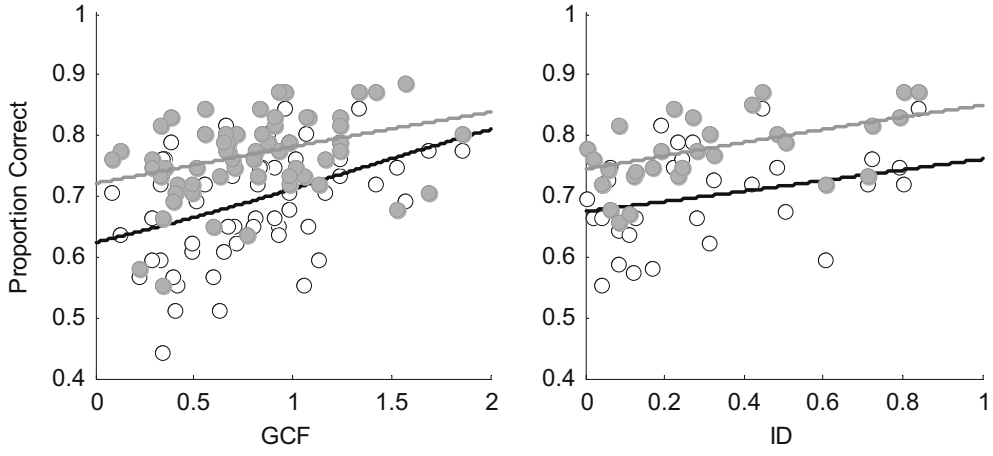

Fig. 4. Plot of the object detection performance as a function of the stimulus parameters "good continuation factor" (GCF; left panel) and object identifiability (ID; right panel). Open circles and the black line reflect the data and fitted model, respectively, for the static Gabor condition, while filled symbols and the grey line reflect the data and fitted model for the local motion Gabor condition.

Table 3
Overview of the parameter estimates for the reduced model (main effects and two-way interactions) and the statistical properties, based on the model described in Eqs. (3)–(5).

	DF	Estimate (β)	Std. error	t-Value	$p > \|t\|$
Intercept	8640	0.1887	0.2664	0.71	0.4787
Stat. vs. Mot.	8640	1.2285	0.1161	10.58	<.0001
Jitter	8640	−0.01665	0.002912	−5.72	<.0001
GCF	8640	3.1246	0.4080	7.66	<.0001
ID	8640	0.7598	0.2777	2.74	0.0062
Jitter ∗ GCF	8640	−0.02039	0.004094	−4.98	<.0001

Table 4
Odds ratios (OR) of the main effect parameter estimates (β). OR = exp(β).

	Estimate	Lower	Upper
Stat. vs. Mot.	3.4161	2.7205	4.289
Jitter	0.9835	0.9779	0.989
GCF	22.7508	10.2257	50.617
ID	2.1378	1.2404	3.684

detection by 3.41 times, or 341%. Similarly, after control for static vs. local motion, GCF, and ID, the odds decrease by 0.984 times, or 1.60%, for each degree of jitter added to the Gabor orientation. After control for static vs. local motion Gabors, jitter interval, and ID, the odds of correct object detection increase by 22.75 times for each unit increase in GCF, and finally, after control for static vs. local motion Gabors, jitter interval, and GCF, the odds of correct object detection increase by 2.14 times for each unit increase in ID.

2.5. Discussion

The present study used Gabor elements arranged so that they give rise to both spatial form and motion form, where the shapes were based on outlines of real-world objects. In line with previous research, we have seen that jitter degrades object detection, while adding motion has the opposite effect. Ledgeway, Hess, and Geisler (2005) found that motion increased the contour detectability by approximately 10% in relation to the static condition, well in line with our results of a 10–15% improvement. We have also seen that object identity, as well as contour-specific properties expressed through the "good continuation factor" (GCF), play a role on detection. We tested for simple interaction effects of all the model parameters and found that jitter and GCF were the only ones that had a significant interaction. We find it worth noting that motion and jitter did not have a significant interaction; some possible implications of this will be discussed later.

3. Experiment 2

In Experiment 2 we go one step further in exploiting the potential of Gaborized contours. This experiment differs from Experiment 1 in that we asked for identification of the objects instead of detection. Furthermore, we changed the paradigm from a 2-AFC to a reaction time experiment. The objective was to see if jitter and motion had a similar influence on identification as it did on detection. This provides an additional kind of generalization of the earlier findings with simpler stimuli to more complex, natural shapes and it allows for a more direct link between grouping and recognition.

3.1. Subjects

Subjects ($N = 6$) were two male and four female, aged 20–24, with normal or corrected-to-normal vision. One subject is the second author; the remaining subjects were recruited in the general student population for a paid (per hour) participation, and were naïve regarding the purpose and the details of the experiment (different from those that participated in Experiment 1).

3.2. Stimuli and Procedure

Stimuli were produced in the same manner as for Experiment 1, except that the jitter intervals now were 0°, 20°, 40° and 60°. Lower jitter levels were now employed because identification of degraded stimuli is a much more demanding task than detection (this may not be the case for categorization of high-quality pictures, see Thorpe, Fize, & Marlot, 1996). We reduced the object set to 20 items, selected from the 60 objects used in Experiment 1, chosen to have a fairly even distribution of compactness (0.36 ± 0.11), identifiability (0.32 ± 0.20), and "good continuation" (0.928 ± 0.396).

Observers were first introduced to the objects by displaying them for 2 s with the correct object name, and then trained prior to the main experiment. The training trials started with a fixation cross for 300 ms, followed by displaying the stimulus until the observer pressed a button or until 5 s had elapsed. The task was to press a button as soon as the object could be identified and then to type the name of the object. Objects and jitter conditions (four jitter levels) were presented in random order within blocks of static and motion stimuli, and training was terminated when the observer had correctly identified each object twice.

The main experiment followed the same paradigm as the training session. Every condition was repeated four times for each of 20 objects for a total of 640 trials per subject and a total duration of approximately 2 h. Subjects were free to take breaks at their leisure.

3.3. Results

Based on preliminary analysis, one of the six subjects was rejected because of outlying reaction times and many errors (only 84% correct identification; criterion for rejection was set to average performance below 95% correct). Only the correct trials were included in the following analysis.

Fig. 5 presents the main effects of jitter on reaction time for identification for the static (open symbols) and the local motion condition (closed symbols). The lines represent simple linear regressions (black for static and grey for local motion). We see an effect of jitter, at least for the static condition, where reaction time increases as jitter is added. We also see a clear effect of adding motion (lower reaction times than the static case), but here jitter seems to have less effect. The different slopes for the static and

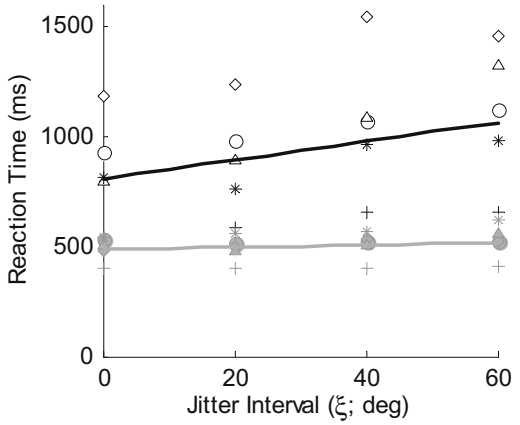

Fig. 5. Plot of the reaction time for object identification task as a function of the jitter interval. Different symbols represent different subjects; open symbols and the black line reflect the data and fitted model for the static Gabor condition, while filled symbols and the grey line reflect the data and fitted model for the local motion Gabor condition.

Table 5
Type III effects for the repeated-measures ANOVA (main effects) on reaction time for object identification.

Effect	Num DF	Den DF	*F*-Value	Pr > *F*
Stat. vs. Mot.	1	3060	2044.64	<.0001
Jitter	1	3060	82.55	<.0001
GCF	1	3060	38.09	<.0001
ID	1	3060	21.89	<.0001

Table 6
Type III effects for the repeated-measures ANOVA (reduced model; stepwise backward elimination) on reaction time for object identification.

Effect	Num DF	Den DF	*F*-Value	Pr > *F*
Stat. vs. Mot.	1	3057	362.08	<.0001
Jitter	1	3057	90.32	<.0001
GCF	1	3057	42.46	<.0001
ID	1	3057	23.92	<.0001
Jitter * Mot.	1	3057	37.26	<.0001
GCF * Mot.	1	3057	31.20	<.0001
ID * Mot.	1	3057	10.97	0.0009

motion conditions seem to indicate an interaction between motion and orientation jitter.

For the statistical analysis we performed a repeated-measures ANOVA over subjects (same factors as in Eq. (5)). The results for the main effects are presented in Table 5, and, again, all were highly significant ($p < 0.0001$).

Like in the previous experiment, we would like to look at the possible interactions between the explanatory variables. Fig. 5 already hinted at different effects of jitter on the static and motion condition, and statistics from a full model including simple interaction effects also confirmed this (Jitter $\times$ Motion; $p < 0.0001$). Reducing the model to include only significant effects (Table 6), we see that there is also a significant interaction between motion and GCF, as well as motion and ID ($p < 0.0001$ and $p = 0.009$, respectively).

3.4. Discussion

In the same way as for Experiment 1, we used Gabor elements arranged so that they give rise to both spatial form and motion form, here to serve the task of identification. We see that adding jitter has a negative impact on performance (longer reaction time for correct identification), while adding motion has the opposite effect. Looking at Fig. 5, one could be concerned that we are experiencing a floor effect in the local motion condition: the difference between the average reaction time for 0° and 60° of jitter is only 33 ms and the spread of the data points is a lot smaller than for the static condition. We performed an additional repeated-measures ANOVA for the local motion trials only, and a stepwise backward elimination on the full model (including two-way interaction effects) revealed jitter and ID as significant factors in the reduced model ($p = 0.0001$ and $p = 0.0268$, respectively). This does not entirely exclude the possible influence of a ceiling effect, but it shows that jitter still plays a role for moving Gabors within the tested interval. As in Experiment 1, both contour-specific properties, expressed through the "good continuation factor" (GCF), and object identity (ID) influenced the reaction times for correct identification (see Fig. 6).

4. General discussion

In these two experiments we have seen that both semi-local contour properties and global object properties can play a role in both object detection and identification. As expected, local motion enhanced performance in both tasks. Several explanations could account for this. For instance, the direction selectivity of some motion-sensitive neurons is broader than the orientation selectivity of neurons that encode orientation from static structure (Allman et al., 1985), making them less sensitive to the directional jitter that inherently ensues from our manipulation of element orientation (since the local motion direction is perpendicular to the element orientation). This could enable the motion system to perceptually link elements where the static, orientation selective cells were unable to do so. It is also possible that motion processing strengthens, and perhaps extends, the lateral interactions in early visual cortex through feedback and/or inter-stream connections (Braddick, O'Brien, Wattam-Bell, Atkinson, & Turner, 2000; Felleman & Van Essen, 1991; Ferber et al., 2003, 2005; Grill-Spector et al., 1998; Murray et al., 2003; Van Essen & Maunsell, 1983). It can also be argued that the moving stimulus produced a more sustained response (as opposed to a static stimulus that may have a more transient on/off response), thereby sustaining lateral interactions.

Given the significant exchange of information between major visual pathways, it is relevant to ask how spatial form and motion form interact in these two tasks. Interestingly, we found slightly different results for the effect of motion according to whether the task was to detect or to identify objects: the interaction between jitter and motion was non-significant in the first case, but

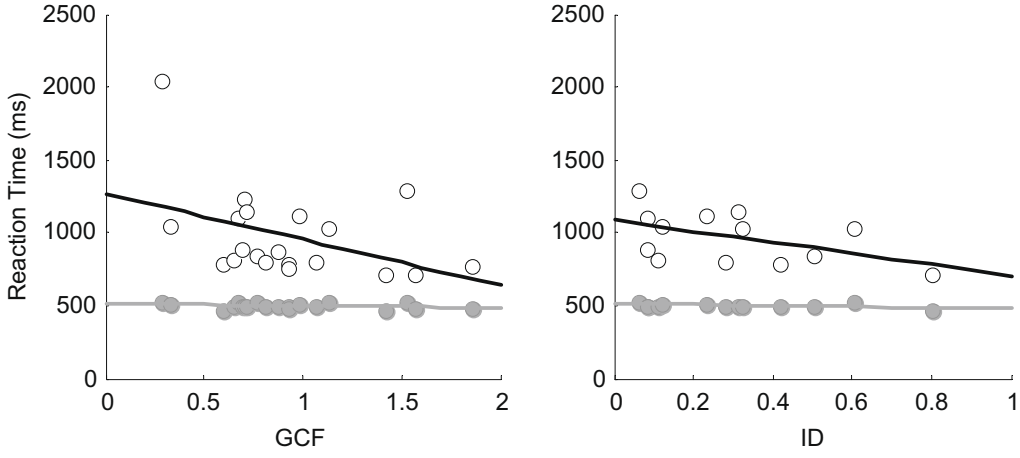

Fig. 6. Plot of the reaction time for object identification task as a function of the stimulus parameters "good continuation factor" (GCF; left panel) and object identifiability (ID; right panel). Open circles and the black line reflect the data and fitted model, respectively, for the static Gabor condition, while filled symbols and the grey line reflect the data and fitted model for the local motion Gabor condition.

significant in the latter. These analyses were based on a largely different range of jitter intervals in the two experiments, so in order to increase the comparability, we reanalyzed the data from Experiment 1 in the range of 0–60° of added jitter. We did not find a significant effect for an interaction between jitter and motion ($p = 0.84$), and ID fell away as a significant main effect ($p = 0.15$). However, since the tasks were different, and our data for Experiment 1 is sparse in the range of 0–60° of added jitter, these results should be treated with care. On the other hand, others have found similar results using static stimuli composed of Gabors of different contrasts and spatial frequencies: Meinhardt, Schmidt, Persike, and Roers (2004) found that feature synergy (from cue combination) depended on both visibility and objecthood, and that cue combination had a more beneficial effect in an identification task than in a detection task (Meinhardt, Persike, Mesenholl, & Hagemann, 2006). In our results, the estimate for the interaction between jitter and motion in the identification experiment is positive (0.00374, $p < 0.0001$), meaning that the effect of adding jitter is stronger in the static condition than in the motion condition, i.e., that the increase in reaction times is higher (as we also saw in Fig. 5). The same is true for GCF and ID, but with opposite sign: the reaction time decreases more in the static condition when GCF or ID increase (estimates are −0.4688 and −0.2323; $p < 0.0001$ and $p = 0.0009$, respectively). Therefore, it seems that the addition of local motion not only has a facilitatory effect, but also that this effect is stronger when the stimuli are more difficult to see (i.e., when more jitter is added in the identification experiment). Our results are thus in line with those of Meinhardt et al. (2004, 2006).

Previous research suggests that the mechanisms responsible for motion form have different characteristics in relation to those responsible for spatial form (Ledgeway & Hess, 2006; Lorenceau & Alais, 2001; Rainville & Wilson, 2004). When Ledgeway and Hess (2006) studied the spatial frequency and orientation selectivity for the mechanisms that extract motion-defined contours, they found some differences in relation to static contours: the spatial frequency tuning for motion contours was very broad, and they remained very detectable when their orientation was oblique to the contour (i.e., all elements had the same orientation, but not in the direction of the contour backbone). However, this result cannot be extrapolated to explain our results, since all our elements had the same spatial frequency and a non-uniform orientation. Ledgeway and Hess (2006) further found a difference in performance when alternating elements were aligned along the contour (every other element was misaligned): subjects were at chance level for static contours and around 75% correct for motion contours. These results are below probability summation for static alternating elements, and the authors suggested that this could be due to some inhibitory process that interferes with the static contour grouping. In our case, assuming that there is more cue summation in figure identification than in figure detection supports the view that spatial form and motion form have different characteristics. It seems that the motion cue adds more certainty to the judgment of form (up to the identity level) than to the judgment of the mere difference of target and distracter. Adding motion increases local salience, and therefore improves target detectability. However, and more importantly, it seems to even more enhance grouping and linking of similar fragments across space to allow integration into a global shape. Hence, it would seem that it is the benefit of form completion from feature integration that explains the stronger cue summation effect in object identification compared to object detection. Apparently, promoting form completion by adding a motion cue serves figure-ground segregation and form identification, resulting in a more stable object vision in noisy environments. This effect can also be demonstrated with other features than orientation and motion (Meinhardt et al., 2006).

5. Conclusion

In the first experiment we replicated results from many detection studies with our real-world object contour stimuli: adding orientation jitter to contour elements deteriorates object detection, while adding motion has the opposite effect. In the second experiment we extended these results to an object identification task: reaction times for identification increased as orientation jitter was added, while adding motion drastically decreased reaction times. In both experiments we were able to verify that both contour properties, such as "good continuation", and object properties, such as identity, also influence performance. Our results thus support the long-standing Gestalt notion that grouping helps identification, and vice versa. We consider these results as a first step in a more extensive research program in which we try to connect perceptual grouping, figure-ground organization, and object recognition. An additional step in this endeavor will be to investigate grouping and segmentation in stimuli that are equally complex as the present ones without being identifiable. Provided that it would be possible to equate all kinds of low-level and mid-level properties, this should allow for a more stringent test of the role of identifiability on grouping and segmentation than we could do in the present study.

Acknowledgments

This research was supported by a research grant from the European Research and Training Network: "Perception for Recognition and Action" (HPRN-CT-2002-00226) and from the Research Fund of the University of Leuven (GOA 2005/03). The manuscript was prepared with support from the Methusalem grant awarded to the last author (METH/08/02). We would like to thank Pascal Mamassian for help with stimulus construction software, Bart Ons and Francis Tuerlinckx for advice on data analysis, and two anonymous reviewers and the editor for their helpful comments on the previous version of this paper.

Appendix A. Compactness as a measure of complexity

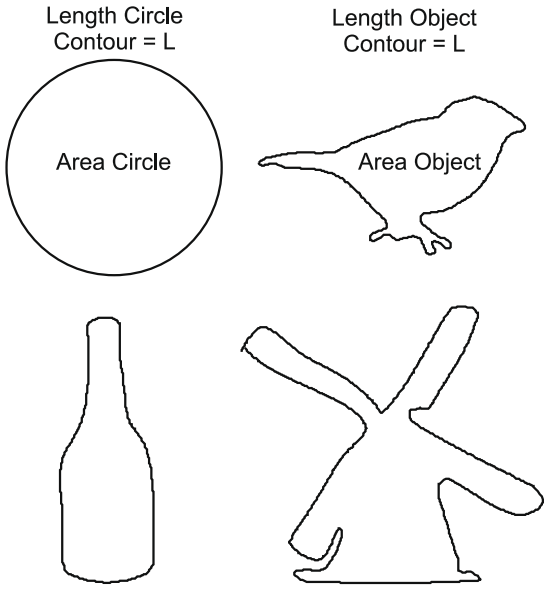

Fig. A1. Compactness was calculated by dividing the area of the object by the area of a circle with the same contour length as the object (Compactness = Area Object/Area Circle). The circle is by definition the most compact object possible (and thus is the least complex; compactness = 1). The bottle, the bird, and the windmill are examples of relatively compact (0.48), medium compact (0.28), and low compact (0.14) objects, respectively. The identifiability of these three objects is 0.13, 0.60, and 0.44, respectively (see Methods on how identifiability was measured).

Appendix B. "Good continuation factor

We defined "good continuation" by a double criterion. First, based on the assumption that human visual interpolation of contours is much like a piecewise spline (Feldman, 1997; Warren, Maloney, & Landy, 2002, 2004), we looked at four contour elements at a time and checked whether the angles between segments joining the center of consecutive elements were less than 30° (Fig. B1, panel A). Then, we checked whether the difference in orientation of two consecutive elements was smaller than 30° (Fig. B1, panel B). The window looking at the four contour elements was then moved along the contour (by one element at a time), and the procedure was repeated (Fig. B1, panel C). "Good continuation" for the stimulus was defined as the number of times both criteria were satisfied within a window. Doing so, the model can (qualitatively) account for findings from earlier studies that have shown that longer contours are more salient than shorter ones (Field, Hayes, & Hess, 1993), since each element has the possibility to contribute four times to the evaluation of GCF (once each time the element falls within the window). In Panel D of Fig. B1 the most salient Gabor elements are highlighted according to this method (brightest elements contributed four times to GCF, the next brightest three times, and so onto the darkest elements, which contributed zero times). The "good continuation factor" (GCF) was then calculated by normalizing the "good continuation" value in relation to contour length (i.e. dividing the "good continuation" value by the number of elements on the contour).

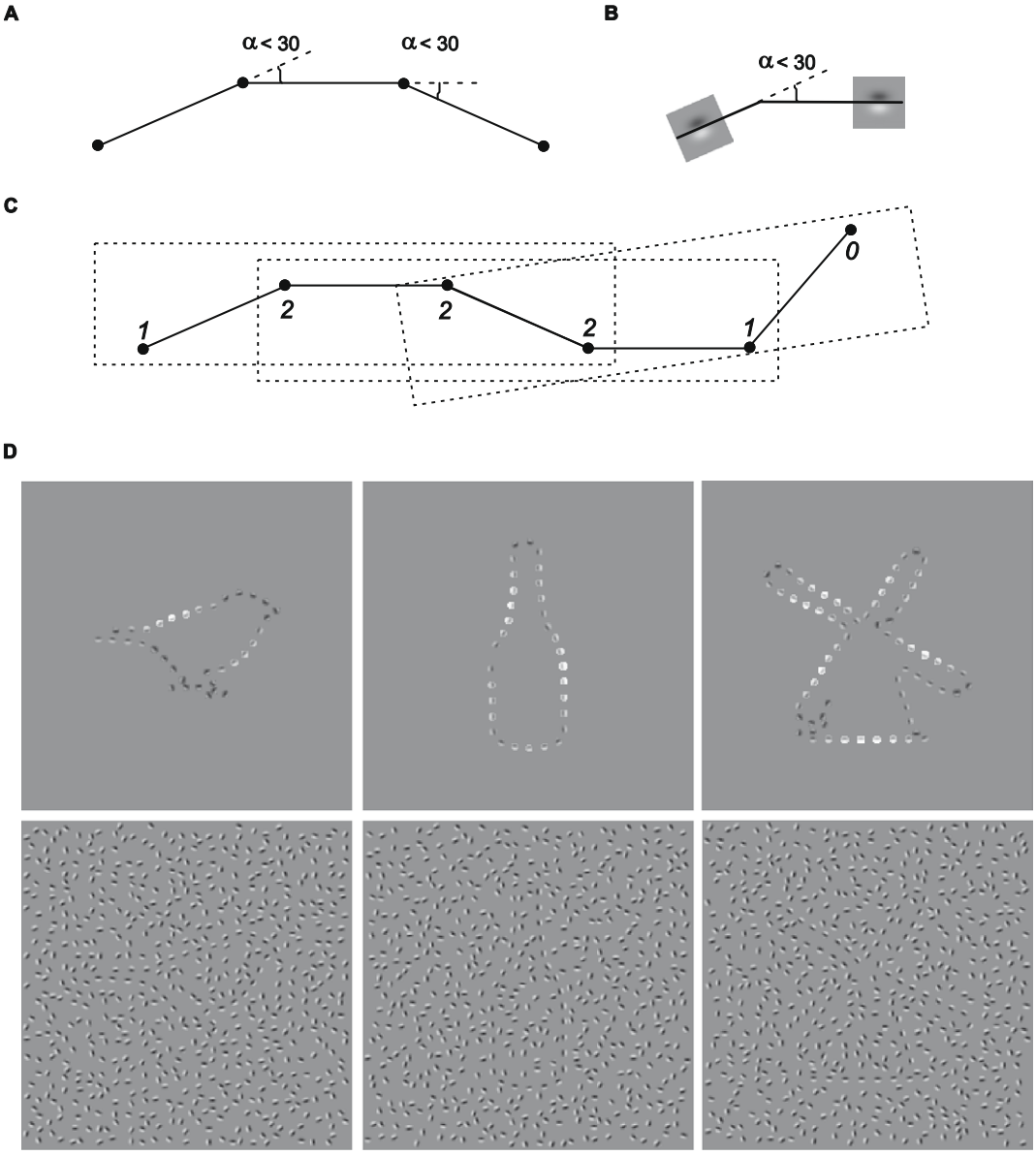

Fig. B1. Panels A–C: Method for constructing the "Good Continuation Factor" (GCF; see also text). Panel D, top row: examples of GCF for three objects. Lighter elements indicate that they contributed one or more (maximum four) times to GCF. GCF for the bird, bottle, and windmill is 0.497, 1.13, and 0.960, respectively. Panel D, bottom row: the objects from the top row embedded in the background.

References

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*, 105–126.

Bell, J., Badcock, D. R., Wilson, H., & Wilkinson, W. (2007). Detection of shape in radial frequency contours: Independence of local and global form information. *Vision Research, 47*, 1518–1522.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2001). Snakes and ladders: The role of temporal modulation in visual contour integration. *Vision Research, 42*, 653–659.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2003). Grouping local directional signals into moving contours. *Vision Research, 43*, 2141–2153.

Braddick, O. J., O'Brien, J. M. D., Wattam-Bell, J., Atkinson, J., & Turner, R. (2000). Form and motion coherence activate independent, but not dorsal/ventral segregated, networks in the human brain. *Current Biology, 10*, 731–734.

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

De Winter, J., & Wagemans, J. (2004). Contour-based object identification and segmentation: Stimuli, norms and data, and software tools. *Behavior Research Methods, Instruments & Computers, 36*(4), 604–624.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences, 11*(5), 219–226.

Feldman, J. (1997). Curvilinearity, covariance, and regularity in perceptual groups. *Vision Research, 37*(20), 2835–2848.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.

Ferber, S., Humphrey, G. K., & Vilis, T. (2003). The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cerebral Cortex, 13*, 716–721.

Ferber, S., Humphrey, G. K., & Vilis, T. (2005). Segregation and persistence of form in the lateral occipital complex. *Neuropsychologia, 43*, 41–51.

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.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 4*, 316–328.

Kayaert, G., & Wagemans, J. (2009). Delayed shape matching benefits from simplicity and symmetry. *Vision Research, 49*, 708–717.

Kourtzi, Z., Bülthoff, H. H., Erb, M., & Grodd, W. (2002). Object-selective responses in the human motion area MT/MST. *Nature Neuroscience, 5*, 17–18.

Ledgeway, T., & Hess, R. F. (2002). Rules for combining the outputs of local motion detectors to define simple contours. *Vision Research, 42*, 653–659.

Ledgeway, T., & Hess, R. F. (2006). The spatial frequency and orientation selectivity of the mechanisms that extract motion-defined contours. *Vision Research, 46*, 568–578.

Ledgeway, T., Hess, R. F., & Geisler, W. S. (2005). Grouping local orientation and direction signals to extract spatial contours; empirical tests of "association field" models of contour integration. *Vision Research, 45*, 2511–2522.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*(11), 3416–3468.

Lorenceau, J., & Alais, D. (2001). Form constraints in motion binding. *Nature Neuroscience, 4*, 745–751.

Machilsen, B., Pauwels, M., & Wagemans, J. (in press). The role of vertical mirror-symmetry in visual shape detection. *Journal of Vision*.

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.

Meinhardt, G., Schmidt, M., Persike, M., & Roers, B. (2004). Feature synergy depends on feature contrast and objecthood. *Vision Research, 44*, 1843–1850.

Murray, S. O., Olshausen, B. A., & Woods, D. L. (2003). Processing shape, motion, and three-dimensional shape-from-motion in the human cortex. *Cerebral Cortex, 13*, 508–516.

Nygård, G.E. & Wagemans, J. (submitted for publication). *Contour grouping and texture segregation in shapes derived from everyday objects: Detection, discrimination and identification*..

Op de Beeck, H., Wagemans, J., & Vogels, R. (2003). The effect of category learning on the representation of shape: Dimensions can be biased, but not differentiated. *Journal of Experimental Psychology: General, 132*, 491–511.

Rainville, S. J. M., & Wilson, H. R. (2004). The influence of motion-defined form on the perception of spatially-defined form. *Vision Research, 44*, 1065–1077.

Rainville, S. J. M., & Wilson, H. R. (2005). Global shape coding for motion-defined radial-frequency contours. *Vision Research, 45*, 3189–3201.

Segaert, K., Nygård, G. E., & Wagemans, J. (2009). Identification everyday objects on the basis of kinetic contours. *Vision Research, 49*, 417–428.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 174–215.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Ungerleider, L. G., & Desimone, R. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. E. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–580). Cambridge, MA: MIT Press.

Van Essen, D. C., & Maunsell, J. H. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences, 6*, 370–375.

Wagemans, J., De Winter, J., Op de Beeck, H., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of everyday objects on the basis of silhouette and outline versions. *Perception, 37*, 207–244.

Warren, P. E., Maloney, L. T., & Landy, M. S. (2002). Interpolating sampled contours in 3D: Analyses of variability and bias. *Vision Research, 42*, 2431–2446.

Warren, P. E., Maloney, L. T., & Landy, M. S. (2004). Interpolating sampled contours in 3D: Perturbation analyses. *Vision Research, 44*, 815–832.

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). London, UK: Routledge & Kegan. Originally published in German (1923).

Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research, 38*, 3555–3568.

Wilson, H. R., Loffler, G., & Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research, 42*, 2909–2923.