Vision Research 49 (2009) 417–428

Contents lists available at ScienceDirect

Vision Research

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

Identification of everyday objects on the basis of kinetic contours

Katrien Segaert, Geir Eliassen Nygård, Johan Wagemans *

Department of Psychology, Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 4 June 2008
Received in revised form 9 November 2008

Keywords:
Figure-ground segregation
Shape-from-motion
Shape perception
Motion perception
Object perception
Random dot patterns
Individual differences

ABSTRACT

Using kinetic contours derived from everyday objects, we investigated how motion affects object identification. In order not to be distinguishable when static, kinetic contours were made from random dot displays consisting of two regions, inside and outside the object contour. In Experiment 1, the dots were moving in only one of two regions. The objects were identified nearly equally well as soon as the dots either in the figure or in the background started to move. RTs decreased with increasing motion coherence levels and were shorter for complex, less compact objects than for simple, more compact objects. In Experiment 2, objects could be identified when the dots were moving both in the figure and in the background with speed and direction differences between the two. A linear increase in either the speed difference or the direction difference caused a linear decrease in RT for correct identification. In addition, the combination of speed and motion differences appeared to be super-additive.

1. Introduction

We perceive the world around us as an organised whole of surfaces and objects. This requires our visual system to structure the bits and pieces that reach our receptors into larger chunks that belong together as parts of meaningful objects and events. When elements are similar to one another with respect to features such as luminance, colour, orientation, primitive shape, etc., they are grouped. Likewise, when certain regions within the visual field are different from one another regarding these properties, the regions can become segregated from one another. In this respect, perceptual grouping and figure–ground segregation are two sides of the same coin (e.g., Palmer, 2003; Peterson, 2003). In the case of clearly segregated surfaces (defined by luminance or colour differences) or outlines, figural cues such as area, convexity and symmetry, usually determine which of the surfaces are seen as figures and which as ground. The edge is then taken to belong to the figure and the background is seen to continue behind it. In this respect, edge assignment and figure–ground segregation go hand in hand too (e.g., Von der Heydt, Zhou, & Friedman, 2003). When the different grouping and segregation cues tend to balance each other out, there is perceptual multistability in the sense that two organisations can be seen. One organisation usually dominates at one point in time but it can then switch to the other organisation. Because the edge belongs to the figure only, the two figure–ground solutions cannot be seen simultaneously.

Static objects that are similar to their background with respect to colour, luminance and texture cannot be segregated from their background. However, it is well-known that motion is a powerful cue to break this kind of camouflage (e.g., Regan, 2000): as soon as the object itself is set in motion or when the texture within the object is moving, the object is segregated from its background and a clear shape is seen (e.g., Uttal, Spillmann, Stürzel, & Sekuler, 2000). Gestalt psychologists coined a special term for this type of grouping by similarity, namely grouping by common fate (e.g., Wertheimer, 1924/1938). Kinetic shapes or kinetic boundaries—shapes or boundaries that are defined by motion—are a useful tool to study motion as a figure–ground cue. They can be generated in several ways, depending on the type of motion difference between two neighbouring areas: absence versus presence of motion, coherent versus incoherent motion, differences in coherence levels, motion direction, speed, etc. Numerous experiments have investigated a whole range of properties of kinetic boundaries and looked at the effect of manipulations like motion contrast, dot lifetime, motion direction, presentation duration, etc. (e.g., Nakayama, 1985; Sekuler, Watamaniuk, & Blake, 2002). These experiments all investigated the detection or discrimination of stimuli that are relatively simplistic in nature, such as lines, bars and simple geometric shapes. Lorenceau and Boucart (1995) used natural objects defined by moving contours in the presence or absence of a static textured background. They showed that performance on an orientation discrimination task was affected (impaired or facilitated) by spatial parameters (like texture density and orientation or contrast of background texture elements). Grill-Spector, Kushnir, Edelman, Itzchak, & Malach (1998) used natural objects defined by differences in contrast, texture, or motion, but they did not register

* Corresponding author. Fax: +32 16 32 60 99.
E-mail address: johan.wagemans@psy.kuleuven.be (J. Wagemans).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.11.012

behavioural responses and they in fact reported that all objects were equally well recognized. Hence, as far as we know, there have been no psychophysical studies that have looked into the identification of kinetic contours of familiar real-world objects and the variables that affect it.

The goal of the present study is to develop stimuli with kinetic contours derived from real-world objects and to test some straightforward aspects of their perception. The first experiment will examine the effect of motion coherence of figural and background dots on the identification of everyday objects derived from kinetic contours. The second experiment will examine the effect of linear increases in the direction and speed differences between moving dots in figure and background. Moreover, this experiment will investigate whether the speed and direction differences are combined additively in the identification of everyday objects derived from kinetic contours. In both experiments, we will measure the time it takes to be able to identify the stimuli as the dependent variable of interest. This measure taps into identification as the output level of preceding processing levels like grouping, figure–ground segregation, edge assignment, etc. We assume that these processes are required to establish some shape percept that can then be interpreted as a known object but we do not want to distinguish between all the different processes leading up to an identification response in this study. Hence, compared to previous work on the role of motion as a figure–ground cue, which has investigated the effects of low-level motion cues on detection or discrimination, this study will examine identifiability of kinetic contours defined by the same motion cues. It will be interesting to find out whether the motion perception principles discovered in the earlier psychophysical work will generalize to this more high-level task. However, it is not our ambition in the present study to disentangle the influences of the different variables on the different processes that are involved.

2. General methods and materials

2.1. Subjects

Subjects were undergraduate students from the University of Leuven between 20 and 23 years old. They had normal or corrected-to-normal vision and signed an informed consent form. They were all naïve regarding the purpose and the details of the experiments.

2.2. Apparatus

The stimuli were generated on a Macintosh Z1-9.1 computer with Matlab Version 5.2 and were presented on a Sony CRT screen with 1024×768 pixel resolution and a refresh rate of 85 Hz. Subjects sat in a darkened room with their head in a chin rest at 57 cm from the screen. They were looking at the stimuli through an aperture with a diameter of 30 cm.

2.3. Stimuli

The stimuli are derived from a subset of the line drawings provided by Snodgrass and Vanderwart (1980) which have been converted into silhouettes and outlines by Wagemans et al. (2008) (see also De Winter & Wagemans, 2004). The object outlines used to make the stimuli in this study are depicted in Fig. 1. These objects differ from each other on a number of factors that may influence perception. One of these factors is the complexity of the shape, which we measured as the inverse of compactness. Compactness is calculated by dividing the area of the object by the area of a circle with the same contour length as the object. Thus, the most compact object is a circle, with a compactness value of 1, while an infinitely complex object has a compactness value that asymptotically reaches 0. The compactness values of the 20 objects (ranging between .66 and .16) are included in Appendix A.

Using moving dot patterns, kinetic contours were derived from these object outlines. Random dot displays consisted of two regions, inside and outside the object contour. When the dots started to move, either inside or outside the contour or in both regions but with a difference between the regions in the speed and/or the direction of motion, the contour could be seen and the object could be identified. The outline of the object served as an edge for the appearance and disappearance of the moving dots, previously called "accretion" and "deletion" of surface elements (Gibson, Kaplan, Reynolds, & Wheeler, 1969; Kaplan, 1969). The two regions of the random dot display could not be distinguished when the dots were static, which is the reason why this is called a dynamic occlusion cue. (Readers can verify this by looking at some example movies on our website:

http://ppw.kuleuven.be/labexppsy/newSite/groepen/index.php?group=1&sublink=topics#5).

The proportion of dots moving in the same direction, the so-called motion coherence level, could be made to vary between 0 (all random) and 1 (all coherent). The dots had a size of 3 pixels and a limited lifetime expressed in stimulus frames. We have tested different lifetimes in our pilot experiments and we have varied it in this study during the training phase of Experiment 1 and between Experiments 1 and 2 (see below), to adjust it at comfortable levels for our identification task. However, we have not manipulated it systematically, which could be done if one were interested in the temporal integration window needed for identification. The stimuli (object and background) spanned 14° by 14° of visual angle, while the objects spanned 6–12° in horizontal and vertical direction. To necessitate a wide distribution of attention, objects were presented away from fixation. More specifically, the center of the object did not coincide with the screen center, but was placed on a randomly chosen point on the circumference of a circle with 5° of visual angle. A fixation cross in the center of the screen preceded the presentation of each stimulus.

2.4. Procedure

The specific procedures that differ between the two experiments will be explained in the respective method sections, while the general aspects are described here. Both experiments used an identification task. In the instructions both speed and accuracy of identification were emphasized. Subjects were asked to press a key as soon as they identified the object, which made the stimulus disappear from the screen. RT from the onset of the stimulus until button press was recorded. Then a response window appeared on the screen and subjects were required to type the name of the object they believed to have identified so that the correctness of their response could be assessed. Stimuli were presented for a maximum of 5200 ms, so if subjects' RTs were slower, no RT was recorded. These trials were not taken into account for the analysis nor were trials in which the identification was incorrect.

3. Experiment 1

The first experiment examined the effect of motion coherence of figural and background dots on the identification of everyday objects derived from kinetic contours. The kinetic shapes were defined by placing a region with motion next to a static region (i.e., if the dots were moving in the figure, then the background dots were static and vice versa).

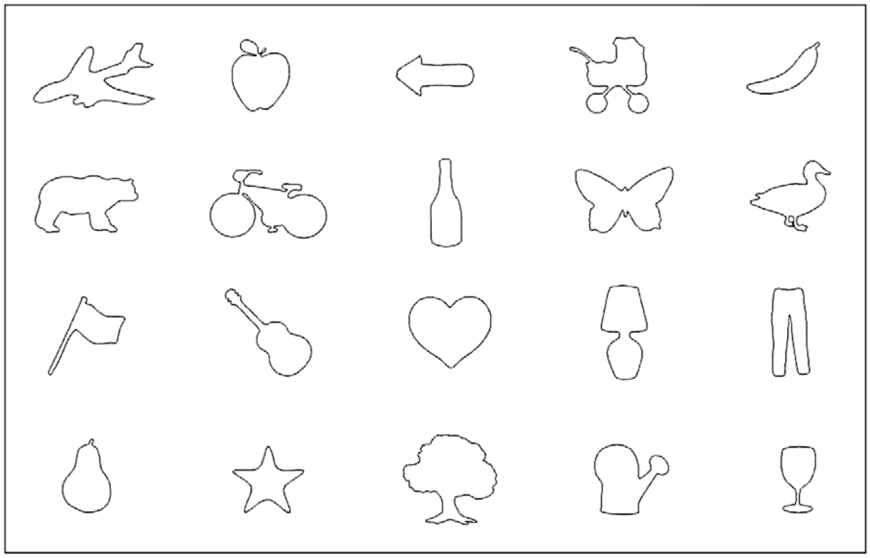

Fig. 1. The 20 objects used to create kinetic contours to be used as experimental stimuli, from left to right, and from top to bottom, in alphabetical order of their basic-level name (see Appendix A).

3.1. Introduction

An important dynamic factor that determines perceptual organisation is the well-known Gestalt principle of "common fate", which states that elements moving coherently in the same direction and at the same speed have a strong tendency to be grouped together (Uttal et al., 2000; Wertheimer, 1924/1938). These grouping processes are typically studied by means of the motion coherence paradigm. In this paradigm random dot kinematograms are used, in which a number of dots (called "signal dots") move coherently together, while the other dots (called "noise dots") move randomly. Since each dot in the kinematogram can be controlled independently, it is possible to manipulate the coherence levels of the display by adding randomly moving dots and then to examine how well the participants perform a discrimination task under different noise levels. Performance in a motion perception task with such a stimulus (e.g., discriminating between two opposite motion directions) is then related to the so-called "coherence threshold", expressed as the percentage of dots moving in the same direction at the same speed (e.g., Grzywacz, Watamaniuk, & McKee, 1995; Watamaniuk, McKee, & Grzywacz, 1995; Watamaniuk & Sekuler, 1992). The effect of motion coherence on motion perception has been studied frequently, also in the context of perception–action coupling (Bleumers et al., 2006; Ceux et al., 2005, 2006).

Depending on the speed and on the duration of the coherent motion, and on the area and visual field of the visual stimulus, the coherence thresholds for human and non-human primate observers have been found to be within the range of 2% and 30% (e.g., Britten, Shadlen, Newsome, & Movshon, 1992; Croner & Albright, 1997). Because we used an identification paradigm, the coherence level was made to vary between .3 (well above threshold) and .9 (close to maximum coherence). There are several ways in which noise can be added to moving dot patterns: random-position noise, random-walk noise or random-direction noise (Scase, Braddick, & Raymond, 1996). Scase and colleagues compared the coherence threshold for directional judgements of the three methods of adding noise and found that the results were not significantly affected by the choice of noise. In our experiment we added random-direction noise. Although it is now well-established that figure–ground segregation can be based entirely on temporal information (e.g., Blake & Lee, 2005; Kandil & Fahle, 2004), we were not interested in providing pure motion cues. We thus measured the incremental effect of motion coherence over the effect of areal dynamics and dynamic borders that are present at all coherence levels. In line with the extensive psychophysical literature on motion detection and discrimination in noise, we predict that the identification of objects based on kinetic contours is also still possible when noise is added but that the identification will be substantially more difficult with low coherence levels than with high coherence levels.

In theory, figure–ground segregation can be region-based or edge-based, i.e., based on the similarities and continuities within a region, or based on the dissimilarities and discontinuities at the edge between figure and background. While Gestalt psychologists have emphasized the role of similarities, psychophysical studies have shown that both mechanisms are important (Møller & Hurlbert, 1996), although segregation can still occur when one of both mechanisms is rendered impossible to be used (Smith & Curran, 2000). In our experiment, one region was always static while the other one contained motion albeit with variable coherence levels. Because the contrast between a static and a moving region is always present at all coherence levels, the edge defined by this contrast will always be visible. Because the grouping within the region with the moving dots will become more difficult with lower coherence levels, the strength of region-based figure–ground segregation will vary with the coherence level. In other words, the decreasing function of identification times with increasing coherence will reflect the role of region-based grouping more than edge-based segregation.

The effect of the second variable that was manipulated in this experiment, whether the region where the dots are moving is the figure or the background, is difficult to predict. Poom and Börjesson (2004) compared bar-motion (motion in the figure) with flank-motion (motion in the background) in a path detection task. They found a similar preference for bar-motion and flank-motion, although detection was slightly easier with bar-motion. Based on this finding, one could predict that the identification for objects defined by figural motion and objects defined by background motion are equally easy or difficult. However, findings based on detection tasks with this type of simple stimuli are not necessarily applicable to identification tasks with everyday objects as stimuli. In general, if one assumes that more attention is directed to the figure than to the background when looking at the world around us, one could argue that identification will be faster with motion in the figure and a static background. However, since motion within a figure is quite infrequent while a static figure on a moving background is typically present at the retinal level whenever tracking a moving object, one could also argue that the identification will be faster with a static figure and motion in the background.

In addition to the effect of coherence level and the location of the moving elements (in the figure or in the background), it will be interesting to compare different objects. We used contours derived from 20 different everyday objects, which varied regarding properties like complexity (e.g., degree of curvature variation along the contour). Previous experiments with a detection task have looked at the effect of curvature. In a path detection task with path elements consisting of moving dots, Bex, Simmers, and Dakin (2003) showed that the visibility of the moving paths decreased at high curvature. Also in path detection tasks with oriented Gabors, it was found that performance deteriorated as the path curvature increased (Field, Hayes, & Hess, 1993; Ledgeway, Hess, & Geisler, 2005). These experiments are not ideal to infer predictions about the effect of object complexity. The first reason is that these experiments use a detection task. Another, more important reason is that the effect of curvature of a line fragment likely is very different from the effect of curvature along the contour of an object. In our previous studies on identification (and segmentation) of silhouette or outline versions of line drawings of everyday objects (e.g., De Winter & Wagemans, 2004; De Winter & Wagemans, 2006; Wagemans et al., 2008), and even in fragmented versions of them (Panis, De Winter, Vandekerckhove, & Wagemans, 2008; Panis & Wagemans, in press), complexity was measured as the inverse of compactness and it always turned out to play an important role. These studies have found evidence for a negative relationship between compactness and performance in identification and segmentation tasks. More simple and compact shapes have a higher structural similarity, which is an advantage during grouping but a disadvantage during object recognition (Gerlach, Law, & Paulson, 2006; Panis & Wagemans, in press). Based on these results, we predict that identification will be easier for objects with low compactness values than for objects with high compactness values (i.e., more circle-like objects).

Summing up, the first experiment investigates the effect of motion coherence, the effect of the region where the dots are moving (i.e., figure versus background), and the effect of object compactness on the RT for correct identification of everyday objects that are derived from kinetic contours.

3.2. Methods

3.2.1. Subjects

Twenty subjects (10 male, 10 female) volunteered to participate. All were undergraduate students from a variety of different faculties (10 from humanities and behavioural sciences, 10 from natural and biomedical sciences).

3.2.2. Stimuli

The dots were moving either inside or outside the contour with an average speed of 3.9° of visual angle per second. They moved horizontally with a limited lifetime from left to right or from right to left and changed direction every 1300 ms (max. 4 cycles for the max. stimulus duration). The objects used to make the kinetic contours are the 20 objects depicted in Fig. 1 (i.e., airplane, apple, arrow, baby carriage, banana, bear, bicycle, bottle, butterfly, duck, flag, guitar, heart, lamp, pants, pear, star, tree, watering can, and wine glass). These objects were selected because they were easy to identify based on their silhouette and outline versions (see Wagemans et al., 2008) and they contained enough variation in object category (e.g., animals versus artefacts) and in compactness (see Appendix A).

3.2.3. Procedure

Prior to the start of the experiment subjects received some preparation and training in the task. First, the 20 objects were presented to the subjects one after another accompanied by the name of the object. The coherence level was fixed at .9 and the lifetime of the dots was 10 stimulus frames, an optimal condition for the subjects to be confronted with these objects for the first time. All 20 objects were presented in this way twice, once with moving dots in the figure and once with moving dots in the background. The presentation order was random. Second, a training phase followed in which the same 20 objects were presented again—this time without the name of the object accompanying them. The task during this training phase was as described above: to identify the object as quickly as possible. The stimuli were presented both in the condition with dots moving in the figure and in the condition with dots moving in the background, with a motion coherence level of .9 and dots having a lifetime of five stimulus frames. The presentation order was random. The training phase ended when each object was identified correctly twice in each condition, so the training lasted for a minimum of 80 trials. The purpose of this training was to familiarize subjects with the task and the stimuli. After this preparation and training, the actual experiment started, again using the same task. All of the 20 objects were presented in eight conditions, consisting of a factorial combination of two variables: motion coherence was varied at four levels (i.e., .3, .5, .7 and .9) and the location where the dots are moving was varied at two levels (i.e., figure or background). These 160 stimuli were presented to each subject twice, so there were 320 experimental trials which were presented in a random order. The lifetime of the dots was five stimulus frames. The entire experiment lasted approximately 1 h.

3.3. Results

3.3.1. Preliminary analyses

The average number of missed and error trials across the 20 subjects was 33 on 320 experimental trials (SD = 26.25). Three subjects of the 20 were considered outliers with regards to the number of missed and error trials (67, 72 and 111 trials). They admitted to being too tired to concentrate and to press too quickly and to guess to shorten the duration of the experiment. The data of these three subjects were excluded from further analyses.

The subjects whose data were subjected to further analysis had an average of 24 missed and error trials on 320 experimental trials (SD = 13.15). There were slightly more missed and error trials in the conditions with background motion (3.86% of the experimental trials) than in the conditions with figural motion (3.64% of the experimental trials) but an ANOVA with the number of missed and error trials as the dependent measure showed this difference to be not significant ($F_{1,159} < 1$). The number of missed and error trials was mainly determined by the coherence level ($F_{3,159} = 13.73$, $p < .0001$, $R^2_{\text{coherence.partial}} = .23$) and the compactness of the objects ($F_{17,159} = 8.95$, $p < .0001$, $R^2_{\text{comp.partial}} = .53$). The higher the coherence level, the fewer missed and error trials (2.87%, 2.06%, 1.67% and .90% for the experimental trials with a coherence level of .3, .5, .7 and .9, respectively). The number of missed and error trials was also not divided evenly across the objects. Some objects were clearly more difficult to identify than others (on 5440 experimental trials there were 47 missed and error trials with the apple, 45 with the pear and 45 with the duck), while there were almost no missed and error trials with other objects (one with the guitar, three with the arrow and three with the bicycle). The correlation of the number of missed and error trials with the compactness of the objects was .21 ($p = .3650$).

3.3.2. Analyses

A repeated-measures analysis of variance (ANOVA) was performed with location of motion, coherence level of motion and object compactness as within subject-factors and subjects as a random factor, and with the logarithm of RT of the correct responses as the dependent measure. Where necessary due to deviations from sphericity, the Greenhouse–Geisser correction was applied. Significant predictors of the RT for correct identification were the motion coherence level ($F_{3,4878} = 186.86$, $p < .0001$, $\hat{\omega}^2 = .6442$) and the compactness of the objects ($F_{1,4878} = 55.65$, $p < .0001$, $\hat{\omega}^2 = .3822$). The identification was significantly faster for higher coherence levels than for lower coherence levels and identification was also faster for more complex, less compact objects (e.g., plane and bicycle) than for simpler, more compact objects (e.g., apple and pear), with a correlation of .1539 between RT and compactness of the object ($p < .0001$).

The RTs tended to be somewhat shorter when motion was located in figure (1623 ms with SD = .86) compared to when motion was located in the background (1644 ms with SD = .78) but this effect was not significant ($F_{1,4878} = 1.24$, $p = .2828$, $\hat{\omega}^2 = .0045$). Location of motion interacted significantly with motion coherence level ($F_{1,4878} = 14.82$, $p < .0001$, $\hat{\omega}^2 = .0985$). At the highest coherence level of .9 identification was faster when the dots were moving in the figure compared to the background, while at the lower coherence levels the effect was smaller and even turns around at .3 (see Fig. 2a). Contrast tests using the error variance at each coherence level (Scheffé) showed that only at a coherence level of .9 there was a significant effect of the location where the dots were moving ($F_{1,16} = 12.59$, $p = .0027$). The effect was not significant at a coherence level of .7 ($F_{1,16} = 1.44$, $p = .2471$), .5 ($F_{1,16} = 1.85$, $p = .1928$) and .3 ($F_{1,16} = 2.80$, $p = .1139$).

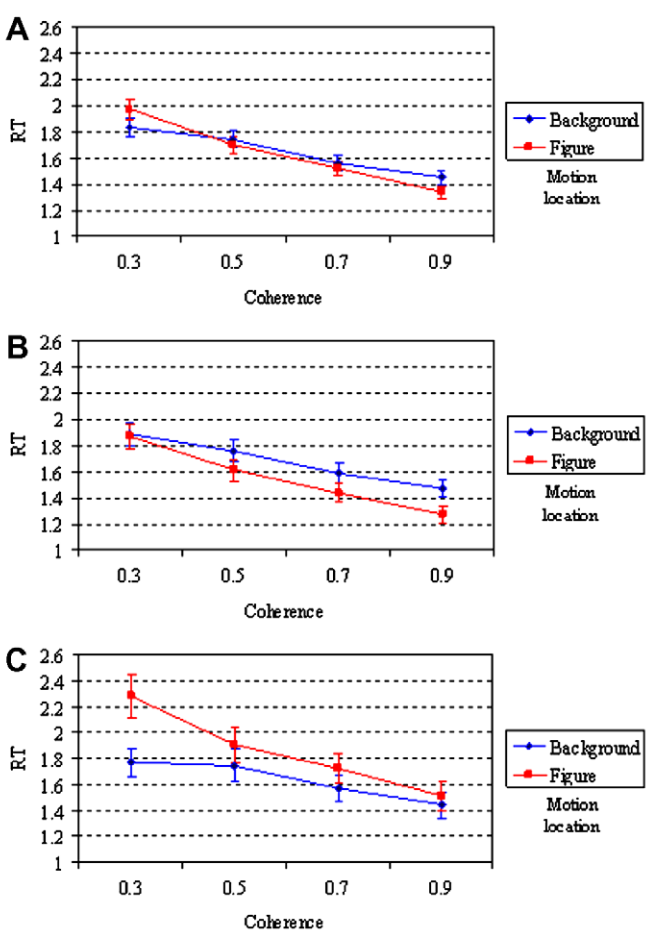

Fig. 2. The effect of motion coherence and motion location (figure or background) on the RTs (s) for correct identification, with the 95% confidence intervals calculated between subjects. (A) For 17 subjects. (B) For a subgroup of 11 subjects who performed better with figural motion. (C) For a subgroup of five subjects who performed better with background motion.

3.3.3. Additional analyses

Taking a closer look at the effect of figural versus background motion, two remarks need to be made. The first remark concerns interindividual differences in the response to figural motion versus background motion and the second remark concerns the distribution of the RTs for figural motion versus background motion.

3.3.3.1. Interindividual differences in response to figural versus background motion. Interestingly, subjects appear to differ in the effect of the location where the dots were moving, figure or background. We divided the subjects into two groups, based on whether they identify the objects faster with figural motion or with background motion. One of the subjects identified the objects in both conditions with the same average RT and thus could not be categorized in one of these two groups.

Eleven subjects identified the object faster with motion in the figure and a static background. An ANOVA within each subject shows that for six of them this difference was significant. A repeated-measures ANOVA on the data of this subgroup of eleven subjects showed that there were four significant predictors of the RT: the location of motion ($F_{1,3141} = 28.90$, $p = .0003$, $\hat{\omega}^2 = .2698$), the motion coherence level ($F_{3,3141} = 121.80$, $p < .0001$, $\hat{\omega}^2 = .6848$), the object compactness ($F_{1,3141} = 31.12$, $p = .0002$, $\hat{\omega}^2 = .3889$) and the interaction between the location of motion and the motion coherence level ($F_{1,3141} = 7.09$, $p = .001$, $\hat{\omega}^2 = .0920$; see Fig. 2b). The correlation between RT and compactness of the object within this subgroup of subjects was .1418 ($p < .0001$).

Five subjects identified the objects faster with motion in the background and a static figure. An ANOVA within each subject found that for three of them this difference was significant. A repeated-measures ANOVA on the data of this subgroup of five subjects showed that again there were four significant predictors of the RT: the location of motion ($F_{1,1447} = 1.30$, $p = .0326$, $\hat{\omega}^2 = .3923$), the motion coherence level ($F_{3,1447} = 44.17$, $p < .0001$, $\hat{\omega}^2 = .8630$), the object compactness ($F_{1,1447} = 14.39$, $p = .0192$, $\hat{\omega}^2 = .7220$) and the interaction between the location of motion and the motion coherence level ($F_{1,1447} = 17.85$, $p < .0001$, $\hat{\omega}^2 = .5261$; see Fig. 2c). The correlation between RT and compactness of the object within this subgroup of subjects was .1678 ($p < .0001$).

The subjects who identified the objects faster with motion in the figure were on average significantly faster than the subjects who identified the objects faster with motion in the background (1551 ms (SD = .77) versus 1733 ms (SD = .87) $t_{4730} = -5.21$, $p < .0001$, two sample t-test). In Fig. 2b and c it can be seen that this differential effect is due to the RTs in conditions with figural motion only. The RTs for conditions with background motion do not differ between the two groups. The comparison between these two panels also suggests that the interaction between location of motion (figure versus background) and coherence level differs between the two subgroups: for subjects who identified objects faster with figural motion (Fig. 2b), the difference with background motion increases with higher coherence, while it increases with lower coherence for subjects who identified objects faster with background motion (Fig. 2c).

3.3.3.2. Differences in the distribution of RTs to figural versus background motion. Fig. 3 shows the distribution of RTs in the conditions with figural motion versus background motion and this figure shows an interesting pattern. A two-sample Kolmogorov–Smirnov test showed that the two distributions differ significantly ($p < .0001$). For the faster RTs (i.e., RTs until approximately 2.5 s) the RTs to figural motion were faster than the RTs to background motion. For the slower RTs (i.e., RTs slower than approximately 2.5 s) the pattern changes and the RTs to background motion were faster than the RTs to figural motion. However, 75% of all observations fall below 2 s. Analyzing only this subset of observations, not only coherence ($F_{3,3665} = 52.66$, $p < .0001$) compactness ($F_{1,3665} = 33.30$, $p < .0001$) and the interaction between coherence and motion location ($F_{3,3665} = 9.38$, $p < .0001$) were significant factors, but also location of motion was a significant determinant ($F_{1,3665} = 11.60$, $p < .0001$). As noted before and indicated by the leftward shift in the mean RT for figural versus background motion, for this subgroup of faster RTs it is the case that the RTs for figural motion are faster than to background motion.

3.4. Discussion

In the first experiment we found that the RTs for correct identification of kinetic shapes decrease for higher motion coherence (i.e., a higher proportion of dots moving in the same direction). Also, the RTs were faster for more complex, less compact objects (e.g., plane and bicycle) than for simpler, more compact objects (e.g., apple and pear). In accordance with this, the number of identification mistakes decreases for higher motion coherence and for more complex, less compact objects. The objects were identified nearly equally well as soon as the dots either in the figure or in the background started to move. This pattern of results implies that identification is easier for kinetic shapes with more coherent motion and for kinetic shapes with a higher complexity, irrespective of where the motion is (in the figure or in the background). However, there were interindividual differences in the response to figural versus background motion: several subjects were clearly faster to identify objects defined by figural motion, while others were faster to identify objects when the background was moving. Also, there were differences in the distribution of the RTs for figural versus background motion. The distribution of the RTs for figural motion was shifted leftwards compared to that for background motion, and for the 75% fastest RTs, identification of kinetic contours with figural motion was faster than with background motion.

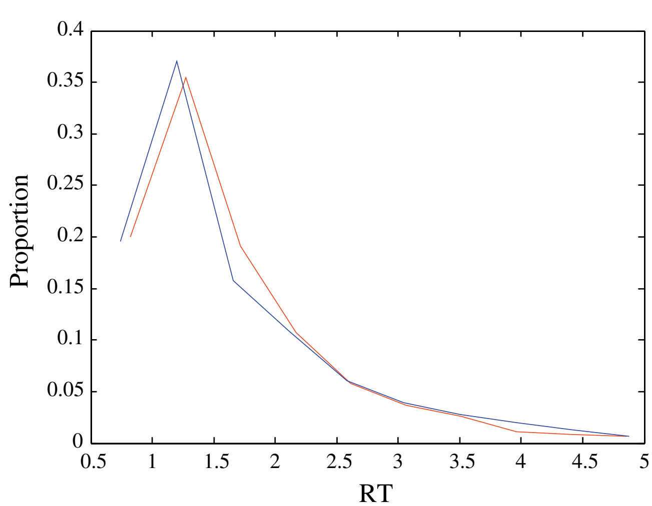

Fig. 3. The distribution of RTs (s) in the conditions with figural motion (blue line) and background motion (red line).

3.4.1. Motion coherence determines the performance in low-level as well as high-level tasks

In the psychophysical literature it has been established that the detection and discrimination of motion-defined boundaries and shapes becomes easier as the proportion of signal dots to noise dots becomes larger. This is because grouping by common fate (or region-based grouping) becomes stronger with higher coherence levels. In the present experiment we manipulated the motion coherence level, using random-direction noise, between .3 and .9. In line with the extensive evidence from experiments using low-level tasks like detection and discrimination, the present experiment using a more high-level task shows that the RTs decrease with higher motion coherence. Our results strongly indicate that motion is a powerful cue for identification, even in the presence of motion noise.

3.4.2. Figural motion versus background motion

In the present experiment, it was found that for figural motion the RTs tend to be somewhat slower than for background motion. This seems to be in line with the results of Poom and Börjesson (2004) who found detection was slightly easier with bar-motion than with flank-motion. In their experiment as well as in the present experiment the overall difference was not statistically significant. However, when focusing on the fastest 75% of the identification responses, it was the case that the RTs for the kinetic contours defined by figural motion were faster than the RTs for the kinetic contours defined by background motion. Asymmetries in figure–ground processing (e.g., Likova & Tyler, 2008) may be quite relevant in understanding the center–surround interactions at the neural level (see Section 5).

Interestingly, there were individual differences in the effect of figural versus background motion. A subgroup of subjects performed better with figural motion, while another subgroup performed better with background motion. This led us to hypothesize that differences in field-independence could have influenced the individual differences observed in our experiment. Subjects who identified the objects faster with figural motion could be field-independent and subjects who identified the objects faster with background motion could be field-dependent. Field-dependence versus field-independence was introduced as a cognitive style by Witkin, Dyk, Faterson, Goodenough, and Karp (1962). It refers to the extent to which someone's perceptual organisation of the visual field depends on the context. People who are field-independent are good at identifying objects in a surrounding that makes the identification difficult or hides the objects. On the other hand, people who are field-dependent are very much influenced by the context and the background. Several authors have argued more recently that field-independence is not to be seen as a broad cognitive style but matches spatial ability (MacLeod & Jackson, 2002; Zhang, 2004).

Field-independence occurs more in men than in women and more in students in natural and biomedical sciences than in humanities and behavioural sciences. A first indication of the role of field-independence could therefore be provided by an unequal division of gender and study program (which were divided equally in the whole sample) in the different subgroups. This was not the case. In addition to this crude indication at the group level, we also tried to test cognitive style differences at the individual level. To this end, we asked all of our 20 subjects to return to the lab and to complete Witkin's (1971) Group Embedded Figures Test). In this test, subjects have to localize a simple figure in the context of a larger and more complex figure that is designed to conceal the simple figure. The simple figure has always been seen before without the context of the more complex figure. The available norms date back to 1971 and are based on American male and female students. In general, our subjects scored very high on the Group Embedded Figures Test. Sixteen out of the 20 subjects scored in the highest quartile and thus were considered field-independent. The fact that this test differentiated so little between the subjects in this study indicates that it might be recommendable to provide and use new norm data. In any case, it can be stated that the differences in the effect of the location of motion are unlikely to be grounded in differences in field-independence, since the correlation of the each subject's score on the test with the signed difference in average RT between the condition with figural motion and the condition with background motion, was virtually zero ($r = .0020$, $p = .9940$). To investigate the factors underlying the interindividual differences between participants who are faster at identifying objects with figural motion than with background motion, requires a separate study with a larger sample and a larger test-battery specifically tailored to addressing perceptual and cognitive abilities.

3.4.3. The role of object properties like compactness

Our previous identification and segmentation studies have found evidence for the influence of factors indexing global outline complexity (De Winter & Wagemans, 2004, 2006; Panis et al., 2008; Panis & Wagemans, in press). The results of the present study are in line with the previous results. The identification is faster for more complex, less compact objects (e.g., plane and bicycle) than for simpler, more compact and thus more circle-like objects (e.g., apple and pear). The influence of curvature of an object contour in identification tasks is thus not the same as the influence of curvature of a path in detection tasks (Bex et al., 2003; Field et al., 1993; Ledgeway et al., 2005). For objects it is the case that when they have low compactness and high part saliency, the global shape information is informative and diagnostic for identification. For objects with low part saliency and high compactness, there are many other candidate objects that look like it, so identification is more difficult (see Panis & Wagemans, in press, for a more detailed study of this effect).

4. Experiment 2

In the second experiment both the figural dots and the background dots were moving. The speed differences and direction differences between figure and background were each manipulated in a linear manner to investigate whether this has a linear effect on the RT for correct identification. This experiment also investigates whether speed and direction cues are combined in an additive manner.

4.1. Introduction

In perceiving the external environment, the brain generally benefits from combining and integrating multiple cues of information. Combining multiple reliable and consistent cues leads to an increase of the information that is available for perception and usually to a better and more reliable performance of the visual system. A striking example is depth perception, which is the result of combining multiple cues like occlusion, relative size, disparity, texture gradient and shading (e.g., Ernst & Bülthoff, 2004; Todd, 2004). Cue combination has been investigated both within a modality (e.g., texture and motion, Rosas et al., 2004, 2007) and between modalities (e.g., vision and haptics, Rosas et al., 2005). Also for identification of kinetic shape stimuli, multiple cues can be used and combined. When there is motion both in the figure and in the background, they can be segregated if there is a variation of the dot velocity between the dots of the figure and the dots of the background. This variation can occur across time as well as space. Thus, speed differences as well as direction differences can be used as cues for object identification.

The results of several path detection studies are informative on this issue. First, Field et al. (1993) looked into orientation or good continuity as path detection cue and they found that path detectability was higher when the orientation of the local elements matched the orientation of the global path than when the orientation was orthogonal to the axis of the global path. Later, Ledgeway et al. (2005) showed that path detectability was higher when the orientation was orthogonal than when the elements were oriented obliquely. Hence, the relationship between relative orientation of the elements and path detection is U-shaped. In line with the work on linking by orientation and the concept of a local association field that integrates local orientation information proposed by Field et al. (1993), more recent studies have investigated linking based on direction (Ledgeway & Hess, 2002) and speed (Hess & Ledgeway, 2003). Ledgeway and Hess (2002) found evidence for a direction-based association field and showed that contours with moderate curvature defined by motion with a direction along the curvature (good continuity) are more detectable than contours defined by motion of any common direction (common fate). Hess and Ledgeway (2003) found evidence for a speed-based association field, thus again linking based on common fate but now over time instead of over space.

It is important to note that the speed-based association field is much weaker than the direction-based association field, which is in turn stronger than the orientation-based association field. Bex, Simmers, and Dakin (2001) found speed-insensitivity for orientation-defined contours and Hess and Ledgeway (2003) found similar speed-insensitivity for direction-defined contours. Ledgeway et al. (2005) showed that detection based on orientation and direction is stronger than detection for static oriented elements. Moreover, Hess and Ledgeway (2003) and Ledgeway et al. (2005) showed that detection based on good continuity and common speed is more effective than detection based on common speed alone. This enhanced performance argues for the existence of two separate visual processes: one mechanism is a specialized contour extraction mechanism that integrates local direction signals that are aligned along the contour regardless of speed, and the other mechanism is a more generalized segmentation process utilizing shared common speed.

In our experiment, kinetic contours are defined by a double cue. Both speed and direction can be used as cues for object identification. There is a shared common speed and a shared common direction within the figure and within the background, while there is a difference in speed and direction between figure and background. Our experiment investigates whether a linear decrease in the direction difference and a linear decrease in the speed difference lead to a linear increase in the time to identify everyday objects based on kinetic contours. In addition, our experiment investigates if speed and direction are combined additively in object identification.

4.2. Methods

4.2.1. Subjects

Five subjects (two male, three female) participated in exchange for a monetary reward. All were undergraduate students (three from humanities and behavioural sciences, two from natural and biomedical sciences).

4.2.2. Stimuli

Both the dots in the background and in the figure were moving. The speed and motion differences between figure and background were manipulated. The speed difference was varied at three levels (i.e., 35°, 55° and 75°), as was the direction difference (i.e., .8, 1.6, 2.4 pixels per stimulus frame), and these were fully crossed leading to nine conditions for each object. We did not include conditions in which only speed or only direction varied because extensive pilot experiments showed that these conditions produced kinetic contours that were not clear enough to be identified. The dots in the figure were moving with a direction of 221° while the dots in the background were moving with a direction of 186°, 166° or 146°, all expressed as clockwise angles from horizontal with 0° on the right. The dots in the background were moving with a speed of 1.5 pixels per stimulus frame while the dots in the background were moving with a speed of 2.3, 3.1 or 3.9 pixels per stimulus frame. These nine conditions were fully crossed with 10 objects, which are a subset of the objects depicted in Fig. 1 (i.e., airplane, apple, arrow, bear, bicycle, butterfly, heart, pear, star and tree). The subset of objects used in Experiment 2 has a similar variation in compactness and object category as the objects used in Experiment 1, containing both objects that in Experiment 1 evoked few missers and mistakes as well as objects that evoked a lot of missers and mistakes. The 90 stimuli were presented 10 times in a random order, making a total of 900 experimental trials presented to each subject. The coherence level was fixed at .9 and the lifetime of the dots was 20 stimulus frames.

4.2.3. Procedure

The experiment consisted of three sessions each lasting approximately 1 h and 45 min. Like in the first study, there was a preparation and training phase before the actual experiment started. There was a preparation phase in each of the three sessions during which the subjects saw the 10 objects accompanied by their name in the easiest condition, i.e. with a speed difference of 2.4 pixels per stimulus frame and a direction difference of 75°. Only in the first session of the study there was a training phase during which the subjects had to correctly identify each of the 10 objects in each of four conditions (i.e., the combination of minimal and maximal speed and direction differences). So, in this training phase there were at least 80 trials that were presented in a random order. After this, the experimental trials followed. In the first session there were 180 experimental trials, in the second and the third session there were 360 trials, making a total of 900 experimental trials.

4.3. Results

4.3.1. Preliminary analyses

The performance of the five subjects was very high. On average only 21.6 (SD = 12.42) mistakes were made on 900 experimental trials. 48.15% of all mistakes were made in the most difficult condition with a speed difference of .8 pixels per stimulus frame and a direction difference of 35°. The higher performance in this experiment compared to the first experiment could be due to many factors (longer lifetime of the dots, longer training phase, many more experimental trials).

4.3.2. Analyses

There were three significant predictors of the logarithm of RTs for correct identification: the speed differences ($F_{2,4382} = 346.18$, $p < .0001$, $R^2_{\text{speed.partial}} = .04$), the direction differences ($F_{2,4382} = 95.42$, $p < .0001$, $R^2_{\text{dir.partial}} = .14$) and the interaction between those two factors ($F_{4,4382} = 8.51$, $p < .0001$, $R^2_{\text{speedxdir.partial}} = .01$). In contrast to the first experiment, object compactness was no longer significant ($F_{1,4382} < 1$).

As was shown in a test of all pairwise differences and as can be seen in Fig. 4, both within the conditions with a speed difference of 1.6 pixels per stimulus frame and within the conditions with a speed difference of 2.4 pixels per stimulus frame, there was no significant difference between the condition with a direction difference of 55° and the condition with a direction difference of 75°. This could indicate a ceiling effect.

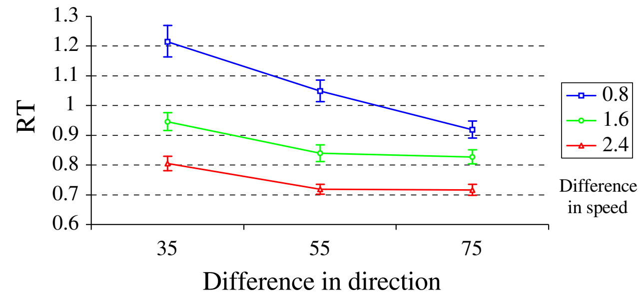

Fig. 4. The RTs (s) and 95% confidence intervals (calculated between subjects) as a function of direction and speed differences.

4.3.2.1. Linearity of single cues. As noted, our first research question was whether a linear decrease of (1) the difference along the dimension of direction and (2) the difference along the dimension of speed would lead to a linear increase in the RT for correct identification.

Three interaction contrasts, one within each level of the factor speed difference with the middle direction difference and the middle speed difference as reference group, could answer the first part of the question. The three contrasts were significant (for speed difference .8 pixels per stimulus frame: $F_{1,4383} = 144.15$, $p < .0001$, for speed difference 1.6 pixels per stimulus frame: $F_{1,4383} = 35.25$, $p < .0001$, for speed difference 2.4 pixels per stimulus frame: $F_{1,4383} = 26.81$, $p < .0001$; compared to $\alpha = .05/3$, Bonferroni correction). This means that within each of the three speed differences tested, there was a significant linear relation between the direction differences and the RTs.

Three other interaction contrasts, one within each level of the factor direction difference with the middle speed difference and the middle direction difference as reference group, could answer the second part of the question. The three contrasts were significant (for direction difference 35°: $F_{1,4383} = 315.77$, $p < .0001$, for direction difference 55°: $F_{1,4383} = 279.62$, $p < .0001$, for direction difference 75°: $F_{1,4383} = 120.85$, $p < .0001$; compared to $\alpha = .05/3$, Bonferroni correction). This means that within each of the three direction differences tested, there was a significant linear relation between the speed differences and the RTs.

4.3.2.2. Cue combination. The second question we addressed was whether speed and direction differences were combined additively in object identification. The method of analysis used to answer this question, is based in part on the method used by Meinhardt, Persike, Mesenholl, and Hagemann (2006) and by Persike and Meinhardt (2008) in studies on cue combination of contrast in spatial frequency and orientation.

As a start, we assumed that speed differences and direction differences are two orthogonal linear dimensions in a vector space. The assumption of the speed and direction cue being two separate dimensions is supported by empirical work using low-level tasks by Hess and Ledgeway (2003) and Durant and Zanker (2008). We have shown empirically in Section 4.3.2.1 that the dimensions are linear: linear steps in increasing direction or speed differences were found to have linear effects on the RT for identification. In this part of the analysis we addressed the question whether these two cues or dimensions combine in a linear additive fashion or a non-linear super-additive or sub-additive fashion. Under the above mentioned assumptions, the predicted RT in case of independent additive cues could be predicted in each condition. Comparing the predicted RT with the observed RTs made it possible to assess whether the combination of the speed and direction cue was additive, super-additive or sub-additive. If the average observed RT does not differ significantly from the predicted RT, this would indicate additive cues. If the average observed RT is faster than the predicted RT, this would be an indication of super-additive cues. If the average observed RT is slower than the predicted RT, this would indicate sub-additive cues.

In each condition of interest the RT was predicted in the case of independent additive cues using the following formula: $\sqrt{(d'_{\text{direction}})^2 + (d'_{\text{speed}})^2}$. With this formula the vector distance between the observed RTs in a baseline condition and the predicted RT in the condition of interest could be calculated. In this formula $d'_{\text{direction}}$ represents the vector distance between the observed RTs in the baseline condition and the observed RTs in the condition with the same speed difference as the one in the baseline condition but the same direction difference as the one in the condition of interest. Likewise, d'_{speed} represents the vector distance between the observed RTs in the baseline condition and the observed RTs in the condition with the same direction difference as the one in the baseline condition but the same speed difference as the one in the condition of interest. The predicted RT in the condition of interest resulted from adding the vector distance calculated with the above mentioned formula to the observed RTs in the baseline condition. The predicted RT could then be compared to the observed RTs and this difference could be tested for significance with a one-sample *t*-test.

Using different baseline conditions and conditions of interest, five comparisons could be tested. For each of these, the calculations and analyses are listed in Table 1. The results for the first three conditions of interest showed that the observed RTs in these conditions were significantly shorter than the predicted RT. This indicates super-additivity of the speed and direction cue in these conditions. In the last two conditions of interest (marked by $*$ in Table 1) there was no evidence for a significant difference between the observed RTs and the predicted RT, providing no evidence for super-additivity but instead for additivity. In the discussion we will argue that these two cases are due to a ceiling effect in the manipulation of the direction differences.

4.4. Discussion

In the second experiment a difference between speed and direction between figural and background motion defined the kinetic contours. In this experiment we found a linear relation between an increase in speed or direction differences and the resulting decrease in RTs for identification. In addition, in some conditions there is evidence that speed and direction differences were combined in a super-additive way in object identification: in those conditions, there was a relative advantage of presenting a double cue compared to what could be predicted under assumption of two independent cues. However, not in all the conditions evidence for super-additivity was found.

4.4.1. Super-additive combination of the speed and direction cue

In the condition with a speed difference of 2.4 pixels per stimulus frame and a direction difference of 75° there was evidence for super-additivity taking one baseline (comparison 1), but taking a different baseline (comparison 5) no evidence for additivity was found. In the condition with a speed difference of 1.6 pixels per stimulus frame and a direction difference of 55° (comparison 2) as well as in the condition with a speed difference of 2.4 pixels per stimulus frame and a direction difference of 55° (comparison 3), evidence for super-additivity was found. No evidence for super-additivity was found in the condition with a speed difference of 1.6 pixels per stimulus frame and a direction difference of 75° (comparison 4). The two cases where no super-additivity was found are cases with a direction difference of 75° that take a baseline with a direction difference of 55° and thus can most probably be explained by a ceiling effect in the manipulation of the direction differences. As revealed by a test of all pairwise differences and as noted earlier, within the conditions with the two highest speed differences, there was no difference in the RTs between a direction difference of 55° and 75°. Thus, our second experiment demonstrates that there is super-additivity of the speed and direction cue unless a ceiling effect in one of both cues is reached. This super-additivity means that when there are motion differences between background and figure, combining both the speed and the direction differences yields a kinetic contour that can be identified faster than the simple additive combination of those cues.

4.4.2. Edge-based and region-based object identification

In our first experiment there was motion in either the figure or the background and the coherence of this motion varied across conditions. As noted earlier, areal dynamics and/or dynamic borders are present at all coherence levels. Although it is not a clean case of region-based grouping, the strength of region-based identification will certainly vary with the coherence level because grouping of the elements within the motion region will become more difficult with lower coherence levels. In other words, the finding that identification becomes more difficult with decreasing coherence reflects the role of region-based grouping more than edge-based segregation. In our second experiment on the other hand, the effects on the RTs for identification tap into edge-based grouping more than into region-based grouping. Within the figure and within the background, there is a shared common speed and a shared common direction, while between figure and background there is a difference in speed and direction. It is not a clean case of edge-based grouping because areal dynamics are present in all

Table 1
For each of five calculations and conditions of interest, the baseline condition and the predicted reaction time are listed together with the vector distance between the reaction time in the baseline condition and the predicted reaction time. Also listed are the observed reaction time and the results of a one-sample *t*-test comparing the observed reaction time and the predicted reaction time.

	Condition of interest	Baseline condition	$\sqrt{(d'\text{direction})^2 + (d'\text{speed})^2}$	Predicted RT (ms)	Observed RT (ms)	One-sample *t*-test
1	75° 2.4 pix/sf	35° .8 pix/sf	0.4977	7385	7161	$t_{497} = 2.38$ $p = .0179$
2	55° 1.6 pix/sf	35° .8 pix/sf	0.2903	9087	8395	$t_{494} = 4.62$ $p < .0001$
3	55° 2.4 pix/sf	35° .8 pix/sf	0.2009	7738	7184	$t_{497} = 6.25$ $p < .0001$
4*	75° 1.6 pix/sf	35° .8 pix/sf	0.2589	8095	8272	$t_{492} = 1.48$ $p = .1408$
5*	75° 2.4 pix/sf	55° 1.6 pix/sf	0.1564	7179	7161	$t_{497} = .20$ $p = .8393$

conditions, but the strength of the edge-based grouping will certainly vary across the conditions of the second experiment. Taking the results of the two experiments together, we can say that region-based grouping as well as edge-based grouping influences the identification of motion-defined objects.

4.4.3. The role of object properties like compactness

In our second experiment the RTs for object identification were determined by the speed difference, direction difference and their interaction, but not by the object compactness. In our first experiment on the other hand, object compactness did influence the RTs: identification was faster for more complex, less compact objects than for simpler, more compact and thus more circle-like objects. A first possible reason why object properties play less of a role in the second experiment could be that there is a longer training phase and many more experimental trials with fewer objects (only half of the objects used in Experiment 1 were used in Experiment 2). Another explanation could be based on the reasoning outlined in Section 4.4.2 that in the second experiment edge-based grouping is probably more relevant than region-based grouping, while in the first experiment region-based grouping was probably more crucial. In cases where region-based grouping plays more of a role, it could be that also global object properties like compactness play more of a role.

4.4.4. Weighing of the speed and direction cue

As noted earlier, previous studies have demonstrated that the direction association field is stronger than the speed association field (Ledgeway & Hess, 2003). However, in this experiment, at first sight the speed differences seem to be a stronger cue than the direction differences. Taking the condition with a speed difference of .8 pixels per stimulus frame and a direction difference of 35° as a reference, the decrease in RT was 24.38% if the direction difference increased with 40° and was 33.76% if the speed difference increased with 1.6 pixels per stimulus frame. However, it is impossible to determine if this is caused by speed differences being a stronger cue or by the gradual differences along the speed dimension possibly being perceptually bigger than the differences along the direction dimension. In this design it is in fact not possible to compare the strength of the differences along the speed dimension with the strength of the differences along the direction dimension. Future studies with a different design can possibly clarify this issue.

5. General discussion

In nature, there are many examples that show the power of motion as a cue for object identification. A well-camouflaged animal, for instance, can immediately be detected and recognized when it moves. There is a large collection of psychophysical studies using kinetic contours that have investigated the role of motion as a figure–ground cue. However, the role of motion as a cue for object identification has been investigated to a lesser extent. We developed stimuli that are very suitable for this purpose. Using random dot displays, we made kinetic contour derived from familiar real-world objects. Thus, these stimuli can be used in a high-level identification task and at the same time they are defined by the same motion cues that were used in previous literature. This way, we could confirm that the motion perception principles discovered in earlier psychophysical work are generalizable to a more high-level identification task.

The results of the first experiment suggested that motion coherence determines the performance in a high-level identification task like in does in low-level detection and discrimination tasks. There was a tendency for faster identification with figural motion and the RT distribution of figural motion clearly differed from that of background motion. Using kinetic contours from real-world objects instead of simple geometric shapes, it is important to take into account general object properties like compactness, since they also have an influence on the RTs. This is likely especially the case when identification depends more on region-based grouping than on edge-based grouping. In the second experiment we found that speed differences and direction differences were combined super-additively in object identification, unless a ceiling effect is reached in one of both cues.

Electrophysiological studies of single-cell responses to oriented bars surrounded by a texture or by noise showed either suppression of the response of neurons when there was a stimulus in the surround of its receptive field or facilitation when center and surround stimuli moved in opposite direction (Allman, Miezin, & McGuinness, 1985). These results were quite important because they provided the first evidence that the direction and speed of background movement outside the classical receptive field exert an influence on the responses of MT neurons to stimuli presented within the classical receptive field. This has been interpreted as evidence that MT neurons are capable of integrating local stimulus conditions within a global context. Psychophysical studies by Lorenceau and colleagues (Castet, Lorenceau, Shiffrar, & Bonnet, 1993; Lorenceau & Boucart, 1995; Lorenceau & Shifrar, 1992) subsequently suggested that the effect of background texture on motion integration of the elements composing the figure reflects lateral interactions that modulate the response of two types of units (motion-sensitive units and end-stopped dot-responsive units).

If one wants to connect the psychophysical literature more closely to the literature on the neural mechanisms, neuroimaging is a viable tool because the same species and tasks can be used. In this respect, an additional and important advantage of our stimuli (kinetic contours derived from everyday objects) and task (object identification) is that they would be highly suitable to be used in future neuroimaging studies investigating the relative role of the kinetic occipital region and the lateral occipital complex. Both areas are likely candidates for interactions between shape processing and motion processing to take place. The lateral occipital complex (LOC) is involved in object recognition. LOC responds stronger to intact objects than to their scrambled versions and the object recognition performance modulates the strength of the signal (Grill-Spector, Kourtzi, & Kanwisher, 2001). Moreover, studies have shown that LOC is involved in processing interactions between different figure–ground segregation cues. Altmann, Deubelius, and Kourtzi (2004) showed the activation in LOC to be modulated by figure–ground segregation and form salience. Contextual information was coded by LOC when this information was relevant to perceiving the target shapes (i.e., when the background elements were presented in the same plane as, and interfered with the target elements). However, there were no contextual effects in LOC when the target form salience was increased by bottom–up information (a disparity or motion cue) or top–down information (i.e., priming the target shapes). Ferber, Humphrey, and Vilis (2003, 2005) used shape-from-motion displays and showed that, although this phenomenon is dependent on motion analysis, it is not the motion selective MT but the LOC that shows persistent activation after the motion stops and that is involved in retention of the percept.

Another possible area in which the interaction between shape and motion processing may take place is the kinetic occipital region (KO). This region is specialized for the processing of motion-defined objects and contours and it is anatomically distinct from LOC, MT, V3 and V3A (Dupont et al., 1997; Orban et al., 1995; Van Oostende, Sunaert, & Van Hecke, 1997). There has been some disagreement in the literature if this area is uniquely specialized for the processing of motion-defined contours. Zeki, Perry, and Bartels (2003) argued it is specialized in the processing of contours in

Table 2
The object number, object name and the compactness value as determined by Snodgrass and Vanderwart (1980) of the objects used in Experiment 1 and/or Experiment 2.

Object number	Object name	Compactness value
002	Airplane	.16
006	Apple	.46
008	Arrow	.45
013	Baby carriage	.18
016	Banana	.34
021	Bear	.35
027	Bicycle	.19
032	Bottle	.48
040	Butterfly	.31
081	Duck	.23
090	Flag	.24
111	Guitar	.31
119	Heart	.66
132	Lamp	.42
162	Pants	.25
166	Pear	.62
217	Star	.29
241	Tree	.32
251	Watering can	.43
258	Wine glass	.35

general, while and Tyler, Likova, Kontsevich, and Wade (2006) argued that KO is specialized in the processing of depth structure, although it should be noted that the region of interests in both studies did not coincide. However, even if KO is important for processing other cues than motion only, this area is an obvious candidate for the processing of the interaction between shape and motion.

A recent fMRI study by Likova and Tyler (2008), using dynamic noise stimuli, compared activations for an experimental condition with a clear figure–ground organisation (a single geometric figure surrounded by a larger region) with a control condition without a clear figure–ground organisation (a noise field segmented into multiple parallel stripes). They showed that the figure–ground configuration generated suppression of the ground representation (limited to early retinotopic visual cortex, V1 and V2) and strong activation in the motion complex hMT+/V5+. Conversely, both responses were abolished when the figure–ground organisation was eliminated, suggesting that figure–ground processing is mediated by top–down suppression of the ground representation in the earliest visual areas V1/V2 through a signal arising in the motion complex. This finding may be related to the center–surround interactions revealed in the single-cell studies discussed above (e.g., Allman et al., 1985). Our stimuli and task could also be quite interesting to test related ideas in a context where top–down representations of real objects come into play.

6. Conclusions

In our first experiment, we showed that the RTs for correct identification decrease for higher motion coherence, in line with results from detection and discrimination tasks. We also established that there is a role for object properties like compactness in identification tasks. The RTs are faster for more complex, less compact objects than for simpler, more compact and thus more circle-like objects. For figural motion, the RTs tend to be somewhat faster than for background motion. Although this is not a significant difference over all coherence levels, the RTs clearly follow a different distributions in the condition with figural motion and the condition with background motion. Interestingly, we found that there are individual differences in the effect of figural motion and background motion. We looked into the possibility that these individual differences relate to differences in field-independence but this was not confirmed by our assessment of this cognitive style. The factors underlying these individual differences can be subject of future research with a larger sample and a larger test-battery that is designed to address perceptual and cognitive abilities.

In our second experiment, we demonstrated that there is a linear relation between an increase in speed or direction differences between figure and background and the resulting decrease in RTs. In addition, we found evidence that speed and direction differences were combined super-additively in object identification, unless a ceiling effect in one of both cues is reached. In this experiment, object properties like compactness did not play the same role as in the first experiment. Possibly, object properties like compactness determine identification speed mainly when identification depends more on region-based grouping than on edge-based grouping. In our experiment, it was not possible to compare the strength of direction differences and speed differences as identification cues. Future studies with a different design could try to clarify this issue.

Acknowledgments

This research was supported by Methusalem funding to the last author (METH/08/02). The second author was funded by a PhD fellowship from the European Research and Training Network: "Perception for Recognition and Action" (HPRN-CT-2002-00226). Thanks are due to Rob Stroobants for advice with data analysis and to two anonymous reviewers for their useful comments.

Appendix A

See Table 2.

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.

Altmann, A. F., Deubelius, A., & Kourtzi, Z. (2004). Shape saliency modulates contextual processing in the human lateral occipital complex. *Journal of Cognitive Neuroscience, 16*, 794–804.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2001). Snakes and ladders: The role of temporal modulation in visual contour integration. *Vision Research, 41*, 3775–3782.

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

Blake, R., & Lee, S.-H. (2005). The role of temporal structure in human vision. *Behavioral and Cognitive Neuroscience Reviews, 4*, 21–42.

Bleumers, L., Ceux, T., Buekers, M., & Wagemans, J. (2006). Synchronising horizontal arm movement with transparent motion. *Behavioural Brain Research, 173*, 199–204.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *The Journal of Neuroscience, 12*, 4745–4765.

Castet, E., Lorenceau, J., Shiffrar, M., & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research, 33*, 1921–1936.

Ceux, T., Wagemans, J., Rosas, P., Montagne, G., & Buekers, M. (2005). The influence of motion coherence manipulations on the synchronization level of a perception–action task. *Behavioural Brain Research, 162*, 83–89.

Ceux, T., Wagemans, J., Rosas, P., Montagne, G., & Buekers, M. (2006). Perceptual-motor adaptations in a synchronization task: The joint effects of frequency and motion coherence manipulations. *Behavioural Brain Research, 168*(2), 226–235.

Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. *Vision Research, 37*, 1415–1427.

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.

De Winter, J., & Wagemans, J. (2006). Segmentation of object outlines into parts: A large-scale, integrative study. *Cognition, 99*, 275–325.

Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A. M., Michiels, J., Marchal, G., et al. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex, 7*, 283–292.

Durant, S., & Zanker, J. M. (2008). Combining direction and speed for the localisation of visual motion-defined contours. *Vision Research, 48*, 1053–1060.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*, 162–169.
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.
Gerlach, C., Law, I., & Paulson, O. B. (2006). Structural similarity and category-specificity: A refined account. *Neuropsychologia, 42*, 1543–1553.
Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics, 5*, 113–116.
Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*, 1409–1422.
Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron, 21*, 191–202.
Grzywacz, N. M., Watamaniuk, S. N., & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Research, 35*, 3183–3203.
Hess, R. F., & Ledgeway, T. (2003). The detection of direction-defined and speed-defined spatial contours: One mechanism of two? *Vision Research, 43*, 597–606.
Kandil, F. I., & Fahle, M. (2004). Figure–ground segregation can rely on differences in motion direction. *Vision Research, 44*, 3177–3182.
Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics, 6*, 193–198.
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., & 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.
Likova, L. T., & Tyler, C. W. (2008). Occipital network for figure/ground organization. *Experimental Brain Research, 189*, 257–267.
Lorenceau, J., & Boucart, M. (1995). Effects of a static textured background on motion integration. *Vision Research, 35*, 2303–2316.
Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research, 32*, 263–273.
MacLeod, C. M., & Jackson, R. A. (2002). On the relation between spatial ability and field dependence. *Intelligence, 10*, 141–151.
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.
Møller, P., & Hurlbert, A. C. (1996). Psychophysical evidence for fast region-based segmentation processes in motion and color. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 7421–7426.
Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research, 25*, 625–660.
Orban, G. A., Dupont, P., De Bruyn, B., Vogels, R., & Vandenberghe, R. (1995). A motion area in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 993–997.
Palmer, S. E. (2003). Perceptual organization and grouping. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 3–43). Mahwah, NJ: Erlbaum.
Panis, S., De Winter, J., Vandekerckhove, J., & Wagemans, J. (2008). Identification of everyday objects on the basis of fragmented versions of outlines. *Perception, 37*, 271–289.
Panis, S., Wagemans, J. (in press). Time-course contingencies in perceptual organization and identification of fragmented object outlines. *Journal of Experimental Psychology: Human Perception and Performance*.
Persike, M., & Meinhardt, G. (2008). Cue summation enables perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1–26.
Peterson, M. A. (2003). On figures, grounds and varieties of surface completion. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 87–116). Mahwah, NJ: Erlbaum.
Poom, L., & Börjesson, E. (2004). Good continuation with kinetic edges. *Vision Research, 44*, 2101–2108.
Regan, D. (2000). *Human perception of objects: Early visual processing of spatial form defined by luminance, color, texture, motion, and binocular disparity*. Sunderland, MA: Sinauer Associates.
Rosas, P., Wagemans, J., Ernst, M. O., & Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: Reliability-based cue weighting without statistically optimal cue combination. *Journal of the Optical Society of America, A. Optics, Image Science and Vision, 22*(5), 801–809.
Rosas, P., Wichmann, F. A., & Wagemans, J. (2004). Some observations on the effects of slant and texture type on slant-from-texture. *Vision Research, 44*, 1511–1535.
Rosas, P., Wichmann, F. A., & Wagemans, J. (2007). Texture and object motion in slant discrimination: Failure of reliability-based weighting of cues may be evidence for strong fusion. *Journal of Vision, 7*, 1–21.
Scase, M. O., Braddick, O. J., & Raymond, J. E. (1996). What is noise for the motion system? *Vision Research, 36*, 2579–2586.
Sekuler, R., Watamaniuk, S. N. J., & Blake, R. (2002). Perception of visual motion. In S. Yantis & H. Pashler (Eds.). *Stevens' handbook of experimental psychology Sensation and perception* (Vol. 1, 3rd ed., pp. 121–176). New York: Wiley.
Smith, A. T., & Curran, W. (2000). Continuity-based and discontinuity-based segmentation in transparent and spatially segregated global motion. *Vision Research, 40*, 1115–1123.
Snodgrass, J. G., & Vanderwart, M. (1980). Standardized set of 260 pictures – norms for name agreement,image agreement,familiarity,and visual complexity. *Journal of Experimental Psychology – Human Learning and Memory, 6*, 174–215.
Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences, 8*, 115–121.
Tyler, C. W., Likova, L. T., Kontsevich, L. L., & Wade, A. R. (2006). The specificity of cortical region KO to depth structure. *Neuroimage, 30*, 228–238.
Uttal, W. R., Spillmann, L., Stürzel, F., & Sekuler, A. B. (2000). Motion and shape in common fate. *Vision Research, 40*, 301–310.
Van Oostende, S., Sunaert, S., & Van Hecke, P. (1997). The kinetic occipital (KO) region in man: An fMRI study. *Cerebral Cortex, 7*, 690–701.
Von der Heydt, R., Zhou, H., & Friedman, H. S. (2003). Neural coding of border ownership: Implications for the theory of figure–ground perception. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 281–304). Mahwah, NJ: Erlbaum.
Wagemans, J., De Winter, J., Op de Beeck, H. P., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of everyday objects on the basis of silhouette and outline versions. *Perception, 37*, 207–244.
Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. *Vision Research, 35*, 65–77.
Watamaniuk, S. N., & Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research, 32*, 2341–2347.
Wertheimer, M. (1938). Gestalt theory. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 1–11). New York: Harcourt (Original Work Published in 1924).
Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. (1962). *Psychological differentiation*. New York: Wiley.
Witkin, H. A. (1971). *Group embedded figures test*. Palo Alto, California: Consulting Psychologists Press, Inc..
Zeki, S., Perry, R. J., & Bartels, A. (2003). The processing of kinetic contours in the brain. *Cerebral Cortex, 13*, 189–202.
Zhang, L. (2004). Field-dependence/independence. cognitive style or perceptual ability?–validating against thinking styles and academic achievement. *Personality and Individual Differences, 37*, 1295–1311.