PERGAMON

Vision Research 40 (2000) 601–611

Vision Research

www.elsevier.com/locate/visres

Coarse-grained information dominates fine-grained information in judgments of time-to-contact from retinal flow

Mike G. Harris *, Christos D. Giachritsis

Cognitive Science Research Centre, School of Psychology, The University of Birmingham, Birmingham B15 2TT, UK

Received 17 May 1999; received in revised form 11 October 1999

Abstract

To investigate the relative importance of fine- and coarse-grained structure in the analysis of retinal flow, subjects made estimates of time-to-contact from random dot kinematograms depicting movement towards a flat, sparsely textured surface. Individual display elements moved smoothly away from each other while expanding smoothly in size. By artificially manipulating the rate at which the individual elements expanded we showed that this cue has only a small effect upon performance. When individual elements were replaced by small clusters of dots, expansion of the clusters had a similarly small effect upon performance. However, estimates of time-to-contact were possible when a single expanding cluster was presented in isolation. We conclude that both types of information are available to the subject but that estimates of time-to-contact are based primarily on coarse-grained changes in the position of image elements and that fine-grained changes in element size or position play only a minor role. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Fine-grained structure; Coarse-grained structure; Random dot kinematograms; Retinal flow

1. Introduction

As we approach an object the resulting expansion of its image provides information about how long we will take to reach it. In the natural world, this visual expansion has at least two potentially useful aspects; texture elements get further apart and they get bigger. The first of these is a coarse-grained cue, consisting of changes in the relative positions of different display elements, typically over a relatively large spatial scale. The second aspect is a fine-grained cue, consisting of transformations of individual display elements, typically over a relatively small spatial scale. Knowing the relative importance of these two cues in judgments of time-to-contact may help us to distinguish between two different approaches to the general analysis of retinal flow (Longuet-Higgins & Prazdny, 1980).

The *discrete* approach, developed primarily by Horn and Schunk (1981), uses just a few (minimally six) flow points to recover the observer's movement direction and the times-to-contact with each imaged point. Because this approach requires the solution of a set of simultaneous equations, it will generally be more accurate if the points are widely spaced in the flow, so that the differences between their motions will be as large as possible. Consequently, a system that takes a discrete approach ought to favour the coarse-grained relationships between different display elements, rather than the typically fine-grained expansion of individual elements.

The alternative, *continuous*, approach, developed primarily by Koenderink and his co-workers (e.g. Koenderink & van Doorn, 1976; Koenderink, 1986) measures the 2D transformation of each small region of the flow in terms of its translation, divergence (local expansion), curl (local rotation) and deformation (local change in shape produced by an expansion along one axis and an equivalent compression along the equivalent axis, so that area is preserved). As might be expected, time-to-contact can be recovered from the divergence, although the analysis is quite complex (Tresilian, 1991). In this approach, the transformations need to be measured in small regions of the display because they will be inaccurate if they combine infor-

* Corresponding author. Tel.: +44-121-4144913; fax: +44-121-4144897.

E-mail address: harris@bham.ac.uk (M.G. Harris)

0042-6989/00/$ - see front matter
PII: S0042-6989(99)00209-6

mation from points at different distances, and this danger will be much greater if the display points are widely separated. Consequently, a system that takes a continuous approach ought to favour the fine-grained transformations of individual texture elements, rather than the typically coarse-grained relationships between different texture elements.

Although the evidence is by no means conclusive, both psychophysical and neurophysiological studies of general flow tend to support the discrete, rather than the continuous approach. For example, the continuous approach requires fine-grained flow that is densely defined and that varies smoothly from position to position (which means that the 3D world through which the observer moves must contain regions that vary smoothly in depth). Yet people can judge locomotory heading (e.g. Warren & Hannon, 1990) and time-to-contact (e.g. Harris, 1998) from random dot flow patterns that are very sparse or that depict an incoherent cloud of 3D points. Neither of these conditions presents a problem to the discrete approach, which requires only coarse-grained flow. Similarly, neurophysiological studies of the medial superior temporal area (MST) reveal cells with large receptive fields that respond selectively to translation, expansion or rotation of the image (e.g. Saito, Yukie, Tanaka, Hikosaka, Fukada & Iwai, 1986), or to some particular combination of these motions (e.g. Duffy & Wurtz, 1991). Such cells might provide the basis for a biologically plausible, discrete analysis using simple directional templates (e.g. Perrone & Stone, 1994; Harris, 1998) but their receptive fields are too large — up to 100 deg — to provide the localized measures of expansion, rotation and deformation needed by the continuous approach.

The psychophysical evidence relating specifically to time-to-contact is more equivocal. Conventional random dot kinematograms, in which the dots move but do not change size, demonstrate that coarse-grained information alone is sufficient for accurate judgments of time-to-contact (e.g. Regan & Hamstra, 1993; Freeman, Harris & Tyler, 1994; Regan & Vincent, 1995). The role of fine-grained information is much less clear. Casual observation of 2D displays in which the elements expand but do not change position suggests that fine-grained information, by itself, is not a sufficient cue. On the other hand, Savelsbergh, Whiting & Bootsma (1991) showed, in a real 3D ball-catching task, that changing the ball's size during its flight does have subtle effects upon performance, suggesting that fine-grained information can play some role. Moreover, using a continuously textured, 2D expanding square display, Vincent & Regan (1997) showed that even small changes in the rate of expansion of the fine-grained texture could produce relatively large changes in the judgment of the time-to-contact of the square.

In this paper, we clarify the relative importance of fine- and coarse-grained structure in judgments of time-to-contact by a series of experiments in which we systematically placed the two cues in conflict with each other. Our results provide further indirect support for the discrete approach by showing that, in this situation, judgments of time-to-contact are based almost entirely upon coarse-grained information and that fine-grained information plays only a minor role.

2. Experiment 1

Previous studies with simple 2D random dot displays have typically used fixed element size and have thus demonstrated that the rate at which texture elements move apart is, by itself, a sufficient basis for estimating time-to-contact (e.g. Lee, 1976; Schiff & Detwiler, 1979; Lee & Reddish, 1981; Todd, 1981; Lee, Lishman & Thomson, 1982; Lee & Thomson, 1982; Wagner, 1982; Lee, Young, Reddish, Lough & Clayton, 1983; McLeod & Ross, 1983; Lee & Young, 1985; Cavallo & Laurent, 1988; Bootsma & van Wieringen, 1990; Regan & Hamstra, 1993; Freeman et al., 1994; Regan & Vincent, 1995). On the other hand, in a real 3D ball-catching task, Savelsbergh et al. (1991) reported that changes in target size can affect performance. When the ball was deflated during its trajectory, thus contracting and signalling a longer time-to-contact, the catcher's initial grasping movements were delayed by a few milliseconds. Since the ball corresponds to a single, small display element, this suggests that fine-grained structure can, in some circumstances, play a role in estimating time-to-contact. This view is supported by Vincent and Regan (1997), who independently manipulated the rates of expansion of the texture and the figure using a randomly chequered 2D square. They found that slowing the expansion rate of the chequered texture by as little as 10% produced reliable 8–9% overestimation of the time-to-contact with the square.

Experiment 1 used a similar technique to Vincent and Regan (1997). The rate of size change of individual elements in a conventional, discretely textured 2D kinematogram was artificially manipulating during estimates of time-to-contact.

2.1. Method

2.1.1. Subjects

The authors acted as subjects. Both are practiced in similar tasks and have normal vision after simple optical correction.

2.1.2. Task

Previous studies of time-to-contact have often used discrimination tasks (e.g. Todd, 1981; Regan & Hamstra, 1993). These provide good scientific reliability by minimizing the subjective component of the task but are consequently rather indirect. Here we required subjects to make a direct estimate of time-to-contact. A brief display depicted an approaching surface and was followed after a variable interval by a beep. The subject simply indicated whether the beep occurred earlier or later than the time at which the surface would had reached him, if the display had continued. The relative importance of the rate at which display elements expand and the rate at which they move apart was investigated by independently controlling these two aspects of the stimulus. On a given trial, for example, the elements might move rapidly apart, signalling short time-to-contact, while expanding slowly, signalling long time-to-contact.

2.1.3. Stimulus

The stimulus was generated by a PC-type computer and displayed via a standard laboratory interface (Cambridge Research Systems, VSG 2/2) upon a monitor screen (Eizo Flexscan F550i-M) with a refresh rate of 90 Hz and a spatial resolution of 60 pixels deg^{-1}. The display was updated every five scans, giving a frame duration of 55.5 ms, and consisted of 15 frames, giving a total duration of 833 ms. Despite the relatively slow frame rate, subjects reported that the display transformation appeared smooth and that both the expansion and translation of the individual elements was clearly discriminable in all conditions.

The stimulus consisted of a circular display with a constant diameter of 7 deg. At the start of each trial, the pattern typically consisted of 18 circular elements. These were initially arranged at the intersections of a regular square grid and their positions were then shifted by up to 0.25 deg in a random direction to produce a pseudorandom pattern. The luminance of each element had a maximum of 43 cd m^{-2} at its centre and declined linearly to the background luminance of 5 cd m^{-2} along each radius. The initial element diameter was either 0.21 or 0.42 deg, depending upon the condition. The elements moved radially, as though placed upon a frontoparallel, virtual surface which moved towards the observer at a constant 3D speed. Although this situation does not capture the full range of naturally occurring visual expansion, it makes the judgment of time-to-contact relatively easy. To eliminate flicker at the display edges, the luminance of elements crossing the outer perimeter was changed linearly between the maximum and background luminance over a distance of 0.7 deg. On each trial, the Focus of Expansion (FoE) of the stimulus was randomly positioned within 1.75 deg of the display centre, so that the stimulus appeared to approach on a slightly tangential trajectory. This discouraged the subject from basing estimates upon the average speed in any fixed region of the display (Freeman et al., 1994).

The room was darkened and the subject viewed the display monocularly from a distance of 114 cm using a chin and forehead rest incorporating an occluder that allowed normal light adaptation but prevented the non-viewing eye from seeing the display.

2.2. Procedure

Each session investigated five times-to-contact (2, 3, 4, 5, or 6 s) signalled by the rate at which elements moved apart. Each time-to-contact condition was controlled by a separate staircase procedure that adjusted the delay between the offset of the stimulus and the signal beep according to whether the subject responded 'too early' or 'too late' using the standard mouse buttons. No feedback was given. The staircases used a simple one up/one down rule, aimed at the point at which the beep coincided with the estimated arrival time of the stimulus. The first delay duration on each staircase was randomly chosen to be between 1 and 10 s and was initially adjusted in steps of 1.424 s. Step size was halved on each staircase reversal until it was reduced to 178 ms. Each staircase ended after ten reversals at this final step size. The five different staircases were randomly interleaved throughout the session.

Each session investigated one of ten combinations of two initial element sizes (0.21 or 0.42 deg) and five rates of relative element expansion (0.25, 0.5, 1, 2 or 4). Each subject completed five sessions, in a different random order, for each of the ten conditions.

Rates of element expansion are expressed as a ratio in terms of their *time-to-contact*, at the offset of the display, relative to that signalled by the rate at which the elements moved apart. Thus, for example, a ratio of 0.25 indicates that the elements expanded relatively *rapidly*, so that when the rate at which elements moved apart signalled a time-to-contact of 2, 3, 4, 5, or 6 s, the elements signalled 0.5, 0.75, 1, 1.25, or 1.5 s, respectively. A ratio of 1 indicates that both aspects of the display signalled the same time-to-contact, whilst a ratio of 4 indicates that the elements expanded relatively *slowly* and signalled time-to-contact of 8, 12, 16, 20, or 24 s.

2.3. Results

Both subjects reported that the stimuli produced a compelling impression of a flat surface approaching smoothly in depth.

The data from each session were initially summarized by linear regression. The means of the final ten reversal points for each staircase were plotted against the time-to-contact signalled by the rate at which the elements moved apart. An example of the raw data is shown in Fig. 1. Two aspects of performance are important. *Accuracy* is captured by the slope and the intercept of the regression. In this example, where the relative element time-to-contact is 0.5, the dashed line (slope = 1) indicates perfect performance if judgments were based solely upon the rate at which elements move apart; the dotted line (slope = 0.5) indicates perfect performance if judgments were based solely upon the rate at which individual elements expand. *Consistency* is captured by the R^2 value associated with the regression (perfect consistency would produce an R^2 of 1). Both these aspects of performance are shown, separately for the two subjects, in Fig. 2.

2.4. Accuracy

Initial element size had little systematic effect upon either the slopes or the intercepts of the regression, although MGH tended to have a higher (more accurate) slope with the larger elements ($F(1, 4) = 13.12$, $P = 0.022$). There was, however, no reliable interaction between size and time-to-contact for either subject ($F < 1$).

More importantly, the relative time-to-contact of the elements had very little effect upon accuracy. If judgments are determined entirely by the rate at which elements move apart, then the slope functions in Fig. 2 should be flat (dashed line in Fig. 2). If, on the other hand, they are determined entirely by the rate at which elements expand, then the functions should increase linearly with unit slope (dotted line in Fig. 2). There is a slight but non-reliable downward trend for CDG ($F < 1$, slope = -0.03) and a reliable, though small, upward trend for MGH ($F(4, 16) = 3.18$, $P = 0.042$, slope = 0.02).

Changing the rate of element expansion by a constant factor should have no effect upon the regression intercept. There is no reliable effect for MGH ($F(4, 16) = 1.26$, $P = 0.325$) but a reliable effect for CDG ($F(4, 16) = 8.03$, $P = 0.001$). For CDG, at least, it seems that rapidly expanding elements generally make the stimulus appear more immediate, even though they have little or no effect upon the regression slopes. The effect is very small, however, in that increasing the rate of element expansion by a factor of 16 only halves the intercept.

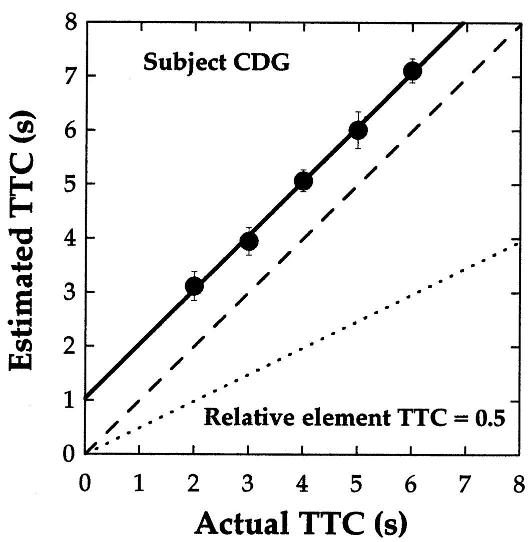

Fig. 1. Typical raw data, for subject CDG, showing the regression of estimated time-to-contact onto actual time-to-contact signalled by the rate at which elements move apart (solid line). In this instance, slope = 1.0062, intercept = 1.024, $R^2 = 0.998$. Error bars show the 95% confidence limits derived from the last ten reversal points of each staircase. In most cases, they are smaller than the symbols. The dashed line (slope = 1, intercept = 0) shows the predicted performance if judgments are based entirely upon the rate at which elements move apart; the dotted line (slope = 0.5, intercept = 0) shows the predicted performance if judgments are based entirely upon the rate at which individual elements expand.

2.5. Consistency

If judgments depend upon both the rate at which elements expand and the rate at which they move apart, we might expect performance to be most consistent when both these aspects of the stimulus agree (i.e. when relative element immediacy is 1). As conflict between the two aspects increases (i.e. as relative element time-to-contact tends to 0.25 or 4), consistency should decrease. There is no evidence of such an effect as measured by the R^2 values in Fig. 2 ($F < 1$ for both subjects).

2.6. Summary

The data confirm the casual observations that originally provoked the study. Judgments of time-to-contact seem to be determined almost entirely by the (coarse-grained) rate at which elements move apart and are almost unaffected by the (fine-grained) rate at which individual elements expand.

3. Experiment 2

Experiment 2 extended the results of experiment 1 to less practiced subjects, using a simpler procedure.

3.1. Method

3.1.1. Subjects

Subjects were eight students from the Psychology School, Birmingham University, and one of the authors (CDG). All had normal or corrected vision. None of the students was practiced with this type of stimulus or knew the purpose of the experiment.

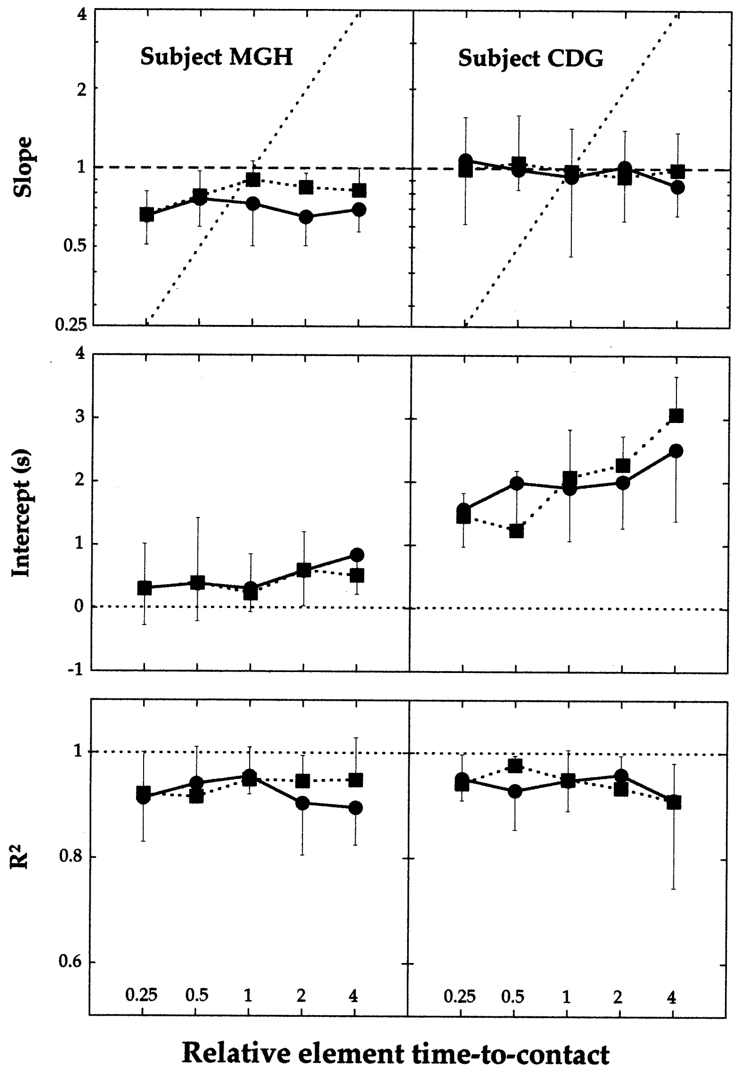

Fig. 2. Summary of the slope (top panels), intercept (middle) and R^2 (bottom) regression results from experiment 1 as functions of the relative time-to-contact of the elements. Subject MGH (left panels); CDG (right panels). Initial element size 0.21 deg (circles and solid line); 0.42 deg (squares and dotted line). The error bars show the 95% confidence interval associated with each individual point. For the slope data, the fine dashed line shows the predicted performance if judgments are based entirely upon the rate at which elements move apart; the fine dotted line shows the predicted performance if judgments are based entirely upon the rate at which individual elements expand. For the intercept and R^2 data, the fine dotted lines show ideal performance.

3.2. Procedure

The display and stimuli were those used in experiment 1. The procedure was greatly simplified so that, on each trial, after the brief (833 ms) stimulus presentation, subjects simply pressed a button to indicate when the stimulus would have reached them, if the display had continued. An additional stimulus condition was included in which the individual texture elements remained fixed in size throughout the display (∞ relative element time-to-contact). Each session consisted of 60 conditions (five times-to-contact $\times$ two initial element sizes $\times$ six rates of element expansion) presented in a different random order. Subjects completed a practice session followed immediately by a single experimental session, from which the data were obtained.

3.3. Results

As in experiment 1, all subjects reported that the stimulus produced a compelling impression of a flat surface approaching smoothly in depth.

The raw data were initially summarized, as before, by linear regression of the estimated times-to-contact onto the actual times-to-contact signalled by the rate at which the texture elements moved apart. The results, pooled across the nine subjects, are shown in Fig. 3.

The results are broadly similar to those of experiment 1. The initial size of the elements had no consistent effect upon accuracy or consistency, and there was no interaction between initial element size and relative element time-to-contact for any of the measures. Changing the relative element time-to-contact had no effect upon accuracy, as measured by slope ($F<1$), and only a small, though reliable, effect upon accuracy as measured by intercept ($F(5, 40)=3.62$, $P=0.009$). As for CDG in experiment 1, changing the time-to-contact signalled by the elements by a factor of 16 altered the intercept only by a factor of 2 or 3. Not surprisingly, given predominantly unpracticed subjects, the simplified task and the small number of data obtained from each subject, consistency is much lower than in experiment 1. For the larger texture elements, there is some suggestion that consistency, as measured by R^2, improves when the cues do not conflict (i.e. when relative element time-to-contact = 1), though this is not reflected in the data for the smaller elements. Indeed, though relative element time-to-contact has a fairly reliable main effect upon the R^2 values ($F(5, 40)=2.59$, $P=0.041$), there is no substantial overall trend.

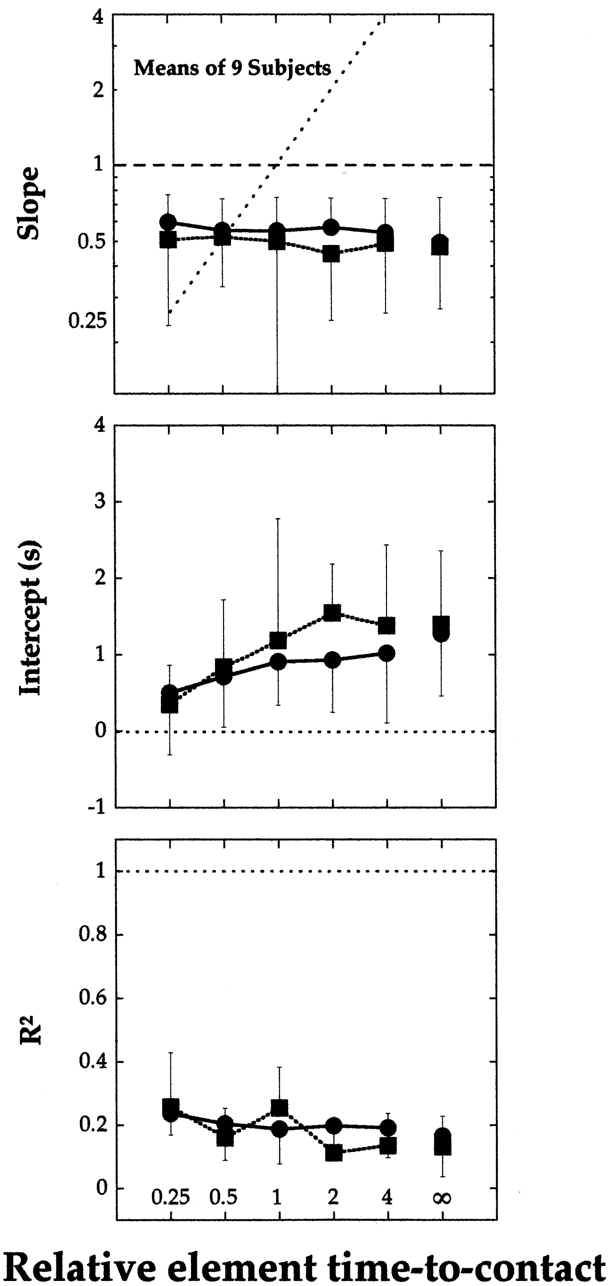

Fig. 3. Average slope (top), intercept (middle) and R^2 (bottom) regression results for nine subjects in experiment 2 as functions of the relative time-to-contact of the elements. ∞ represents the condition in which the element size remained fixed in size. Initial element size 0.21 deg (circles and solid line); 0.42 deg (squares and dotted line). The error bars represent the 95% confidence interval associated with each individual point. For the slope data, the fine dashed line shows the predicted performance if judgments are based entirely upon the rate at which elements move apart; the fine dotted line shows the predicted performance if judgments are based entirely upon the rate at which individual elements expand. For the intercept and R^2 data, the fine dotted lines show ideal performance.

3.4. Summary

The results support those of experiment 1 in demonstrating that the rate of expansion of individual display elements plays little role in judgments of time-to-contact. Large variations in the rate of element expansion produce only small effects that are confined to accuracy as estimated by the regression intercept; very rapid rates of element expansion tend to make the stimulus appear more immediate by a fixed amount. Reassuringly, the results justify the conventional usage of random dot kinematograms in which the dots are fixed in size. There is very little difference, in accuracy or consistency, between such displays (Fig. 3, relative element time-to-contact = ∞) and more 'realistic' stimuli in which the elements expand at the appropriate rate (Fig. 3, relative element time-to-contact = 1).

4. Experiment 3

It is not clear from experiments 1 and 2 whether the expansion of individual display elements is ineffectual because it represents a different *type* of information (transformation of an individual element rather than changes in the spatial relationships between different elements) or because it represents a different *spatial scale* of information (fine-grained rather than coarse-grained). Experiment 3 addressed this question by replacing the individual display elements with small clusters of dots. The clusters were of the same initial size and underwent the same expansion as the original single elements but, within them, the individual dots remained fixed in size and moved apart. Here, then, the conflict arises not because of a difference in the *type* of information but only because of a difference in *spatial scale*.

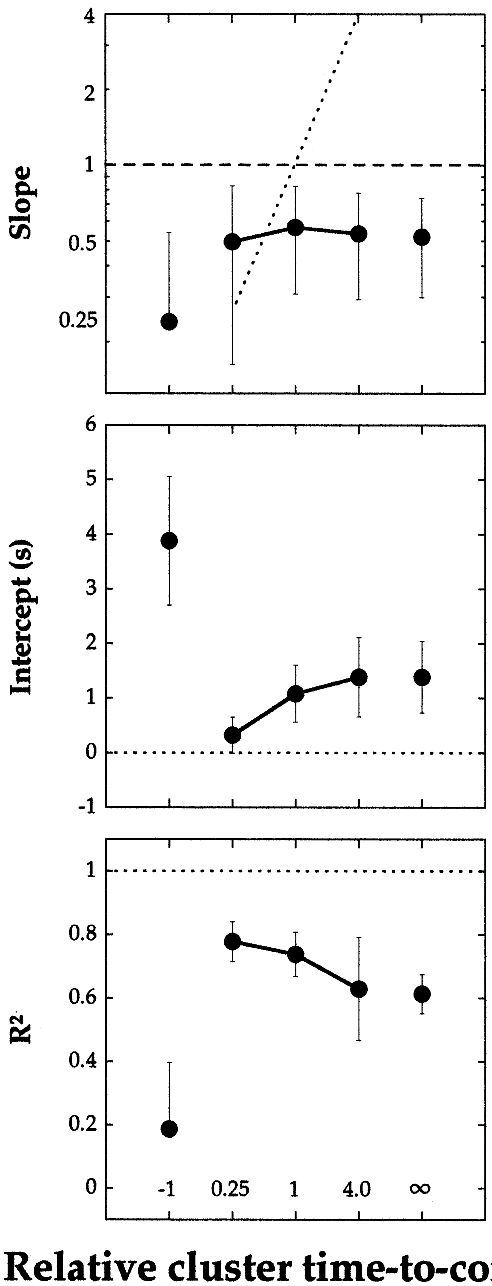

Fig. 4. Average slope (top), intercept (middle) and R^2 (bottom) regression results for two subjects in experiment 3, pooled across initial cluster sizes (0.21 and 0.42 deg), as functions of the relative time-to-contact of the clusters. -1 represents the condition in which clusters contracted, ∞ represents the condition in which they remained fixed in size. Error bars show the 95% confidence interval associated with each individual point. For the slope data, the fine dashed line shows the predicted performance if judgments are based entirely upon the rate at which elements move apart; the fine dotted line shows the predicted performance if judgments are based entirely upon the rate at which individual clusters expand. For the intercept and R^2 data, the fine dotted lines show ideal performance.

4.1. Method

Two subjects, one of the authors (CDG) and an unpracticed postgraduate student who did not know the purpose of work, performed the experiment. Both had normal or corrected vision.

The task was the same as in experiment 2 in which subjects simply estimated the stimulus time-to-contact by pressing a button. The stimulus was similar to that used in experiments 1 and 2 with the important modification that the individual solid elements were replaced by clusters of ten randomly positioned dots of fixed size and luminance. The individual dots were quadrelized to give accurate control of position (Georgeson, Freeman & Scott-Samuel, 1996) and appeared as bright points from the viewing distance of 114 cm.

A subset of the relative time-to-contact conditions from experiment 2 was used (∞, 0.25, 1 and 4) though, in this case, the elements are the clusters rather than the individual dots. A new condition (-1), in which the clusters *contracted* at the rate at which they should have expanded, was also introduced. Each session consisted of 25 trials (five times-to-contact $\times$ five rates of cluster expansion) in random order. Each subject completed a practice session followed by two experimental sessions at each of two initial cluster sizes (0.21 and 0.42 deg).

4.2. Results

The data were first summarized by linear regressions as before. The data from the two initial cluster sizes were very similar for both subjects. The initial size of the cluster had no reliable main effect and did not interact with the relative time-to-contact of the cluster on any of the measures ($F < 1$ in all cases). Fig. 4 shows the mean results, pooled across the two initial cluster sizes and across the two subjects.

For the previously used conditions, the results are very similar to those from experiments 1 and 2. Most notably, changing the rate of cluster expansion by a factor of 16 has no systematic effect on regression slope. Indeed, the only effect upon accuracy occurs in the regression intercepts and is similar to the small effects seen in experiment 2 and, for CDG, in experiment 1. There is a suggestion of an increase in consistency, as measured by R^2, at higher cluster expansion rates, but this effect is small (about 30% over a 16-fold range).

Only the new condition (-1), in which clusters contracted rather than expanded, shows any deviation from the usual pattern. Here both accuracy and consistency are severely impaired. This result is in keeping with the subjects' reports of the stimulus; unlike the other conditions, this stimulus did not appear as a surface approaching in depth. Rather, it either provoked no impression of depth or was seen as a set of distinct objects that appeared to recede. In either case, judgments of time-to-contact were difficult.

4.3. Summary

These results suggest that it is the spatial scale of the stimulus, rather than the type of stimulus, that is important in judgments of time-to-contact. The rate at which stimulus elements move apart is a useful stimulus cue at a coarse-grained spatial scale but it is no more disruptive, when placed in conflict at a fine-grained scale, than the rate of expansion of individual elements. Performance is only disrupted when the conflict is extreme, when the coarse-grained and fine-grained cues signal 3D movement in opposite directions.

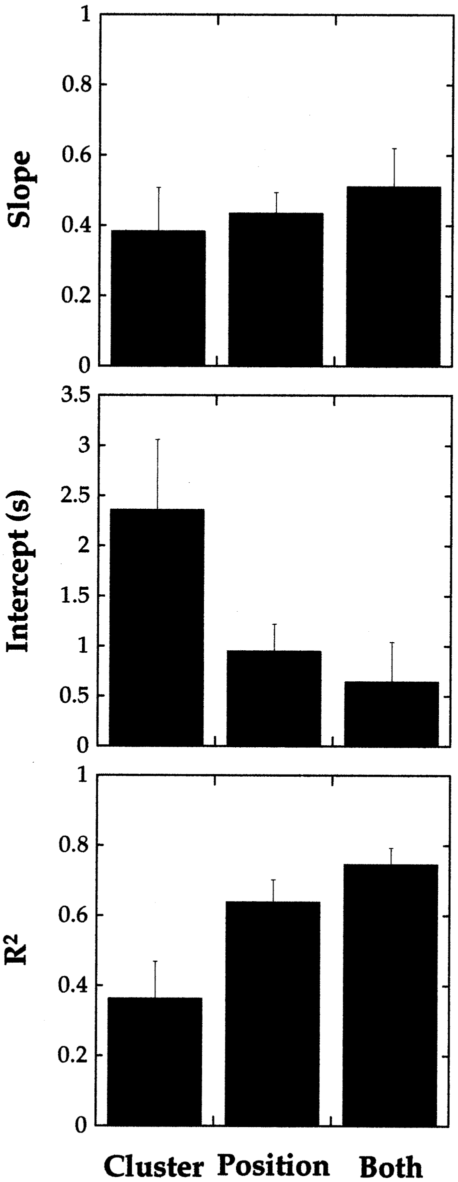

Fig. 5. Average slope (top), intercept (middle) and R^2 (bottom) regression results for two subjects, pooled across two initial cluster sizes (0.21 and 0.42 deg) for the single cluster display used in experiment 4 (Cluster condition). For comparison, corresponding data for the same two subjects are shown for the condition in experiment 3 (relative time-to-contact = ∞) in which multiple clusters moved apart but remained fixed in size (Position condition); and for the condition in experiment 3 (relative time-to-contact = 1) in which multiple clusters moved apart and expanded at the appropriate rate (Both condition). Error bars show the positive half of the 95% confidence interval associated with each average.

5. Experiment 4

Is information about fine-grained expansion unavailable to the subject or is it available but ignored in estimating time-to-contact? The extreme conflict (-1) condition in experiment 3 suggests that small scale expansion is not completely ineffectual since it can, under these extreme conditions, substantially alter performance. It seems, rather, that less extreme conflicts are resolved by basing estimates almost entirely upon the coarse-grained information and largely ignoring conflicting fine-grained information, even though it may well be available. Experiment 4 investigated whether fine-grained information is ignored, or simply unavailable, by removing conflicts and examining judgments of time-to-contact when a single cluster of dots is presented in isolation.

5.1. Method

The two authors acted as subjects. The basic procedure was the same as that used in experiments 2 and 3, in which subjects directly estimated time-to-contact by pressing a button. The stimulus used the same type of dot cluster as in experiment 3 but, on each trial, only one cluster was randomly chosen from the array used in experiment 3 and was presented by itself at a fixed position. Thus the stimulus was like a small conventional random dot kinematogram, in which individual dots moved apart but the mean position of the cluster remained the same.

Each session consisted of 25 trials (five replications of five times-to-contact) in a different random order. Each subject completed two sessions at each of two initial cluster sizes (0.21 and 0.42 deg). MGH, who was not a subject in experiment 3, also did appropriate comparison sessions based on that experiment.

5.2. Results

The data were summarized by linear regression, as before. Again, the initial size of the cluster had no reliable effect upon performance. Fig. 5 ('Cluster') shows the results, pooled across initial cluster sizes and across the two subjects. For comparison, Fig. 5 also shows relevant results from the same subjects from experiment 3. 'Position' shows results from the condition in which clusters moved apart but did not expand (experiment 3, relative cluster time-to-contact = ∞), and 'Both' shows results from the condition in which the clusters moved apart while expanding at the correct rate (experiment 3, relative cluster time-to-contact = 1).

Accuracy generally improves across the conditions, as shown by a small increase in regression slope and a more substantial decrease in regression intercept. For the slope data, none of the differences between conditions is reliable. For the intercept data, the difference between 'Cluster' and 'Position' is reliable ($t(8) = 4.49$, $P < 0.01$) but the difference between 'Position' and 'Both' is not ($t(12) = 1.52$, $P > 0.05$). The results for consistency follow an equivalent pattern to those for the intercept. The R^2 value for fine-grained expansion ('Cluster') is only about half that of the other two conditions. Here both the differences between 'Cluster' and 'Position' ($t(10) = 5.42$, $P < 0.001$) and between 'Position' and 'Both' ($t(13) = 3.37$, $P < 0.01$) are reliable.

Even when there is no conflicting information, fine-grained expansion ('Cluster') provides a fairly poor cue to time-to-contact, as shown by a relatively large intercept and a relatively low R^2. However, the slope data do suggest that fine-grained information is, at least, available to the observer. This confirms, amongst other things, that the lack of effect found in the previous experiments is not due to the spatial or temporal limitations of the display. The data shown in Fig. 5 indicate that, for the 'Cluster' condition, a 1 s time-to-contact would produce an estimate of about 2.75 s whereas a 5 s time-to-contact would produce an estimate of around 4.3 s. For the 'Both' condition, the equivalent estimates are 1.16 and 3.21 s.

5.3. Summary

The results support the suggestion that the visual system can use fine-grained expansion in estimating stimulus time-to-contact but that performance is relatively inaccurate and unreliable when based on this cue alone. Perhaps not surprisingly, when placed in conflict with coarse-grained expansion, as in experiments 1–3, fine-grained expansion plays little role in making an overall judgment.

6. General discussion

When coarse-grained and fine-grained cues conflict, judgments of time-to-contact are based predominantly upon the coarse-grained cues. Experiments 1 and 2 showed that, when the two cues conflict, judgments of time-to-contact are determined almost entirely by the rate at which display elements move apart. The rate of expansion of individual texture elements plays very little role; faster rates of element expansion may make the stimulus look a little more immediate, as demonstrated by a slight decrease in regression intercept, but they have remarkably little effect upon regression slopes or upon the consistency of judgments. Experiment 3 confirmed that the important factor is spatial scale, rather than the type of stimulus information available. Judgments were no more influenced by small, expanding clusters of dots than by small expanding blobs, even though coarse-grained and fine-grained cues were both carried by changes in the relative positions of display elements. Finally, experiment 4 demonstrated that fine-grained changes in element position can act as an effectual cue when presented in isolation, though judgments are relatively inaccurate and inconsistent under these conditions. It seems that fine-grained information can be extracted, but that it is largely ignored when it conflicts with coarse-grained cues.

We cannot say on the basis of our results, of course, that coarser-grained information will always dominate finer-grained information, irrespective of the absolute scale of the display. It may be, instead, that there is some optimal absolute spatial scale of analysis. Our findings nonetheless have obvious practical significance in confirming that traditional random dot kinematograms, in which dots move apart but do not change size, are valid stimuli in the study of time-to-contact. They also have more general and theoretical implications in providing, at least indirect, support for discrete approaches to the analysis of retinal flow. Such approaches rely naturally upon coarse-grained changes between a few discrete display elements, which is the type of cue that proved to dominate judgments of time-to-contact here. In contrast, continuous approaches rely more naturally upon finer-grained transformations in the stimulus, which is the type of cue that proved largely ineffectual here.

Two obvious arguments might be raised in defence of the continuous approach. Firstly, it requires a fine-grained analysis only in the initial encoding of flow information, whereas our results reflect the more complex problem of resolving conflict. This is especially pertinent in the light of experiment 4, which showed that fine-grained information can act as a cue when no conflict exists. However, the conflict in our displays arises largely from the comparison of translation (the way that elements change position) and divergence (the way that elements expand). Continuous models make use primarily of divergence and deformation, whilst translation is used only to scale this information in the analysis of surface layout (see, for example Koenderink, 1986). If a purely continuous approach recognized the conflict present in our stimuli, it would surely not ignore the consistent divergence measures provided by a fine-grained analysis. Secondly, it might be argued that the spatial scale of our stimuli is inappropriate for a practicable continuous approach; that while analysis is relatively fine-grained, it uses a larger scale than that provided by our individual blobs or clusters. There is even some suggestion, in the slopes of one of the experienced subjects in experiment 1, that the larger blobs do exert the expected influence on time-to-contact judgments, at least for fast expansion rates (Fig. 2, top left panel). However, this effect is not statistically reliable and is not repeated in any of the other experiments. Moreover, though the larger blobs and clusters initially subtended 0.42 deg, they expanded during the display so that, for example, at the highest rate of element expansion and the shortest time-to-contact, they subtended 1.12 deg at offset. As argued in the Introduction, useful estimates of divergence are limited in spatial scale by the need to avoid combining measures from points at different distances. In most natural scenes, even those containing only flat surfaces, depth often varies substantially over a degree of visual angle. We would therefore argue that the scale of our stimuli is appropriate for any practicable continuous approach. Indeed, this may well be one reason for the human visual system to favour a discrete approach. In a system limited by additive noise, the fine-grained measures needed for the continuous approach will inevitably have a rather poor signal-to-noise ratio. Perhaps this underpins the inconsistent performance noted in experiment 4. On the other hand, because flow velocity scales with size, the discrete approach can improve the signal-to-noise ratio by taking measures at a coarser spatial scale.

Our results are in keeping with those of Vincent and Regan (1997), although the emphasis of our interpretation is different. Vincent and Regan used a 2D expanding square stimulus depicting a single time-to-contact of 2 s and a staircase task similar to that of our experiment 1. They found that independently changing the rate of expansion of the internal, randomly chequered texture of the square, from no expansion to twice the normal rate, altered judgments of time-to-contact by up to 10%. They rightly point out that effects of this size would be crucial in real tasks such as landing a helicopter and so stress the importance of providing appropriate texture expansion in virtual training simulators. However, like us, Vincent and Regan also conclude that the weighting given to the texture cue is relatively small. Moreover, they showed that although absolute judgment of time-to-contact was affected by conflicting texture expansion, the ability to discriminate time-to-contact, as measure by the slope of the psychometric function, was not. Since their absolute judgments and discrimination measures are respectively equivalent to our intercept and slope measures, this fits the pattern of results reported here. We give greater emphasis to the lack of effect upon slope because we have measured this directly over a wider range of conditions and because this is the measure that we would expect to change if analysis were based upon the fine-grained structure of the stimulus.

Finally, our findings are easily reconciled with those of Savelsbergh et al. (1991), who found that changing the size of an individual stimulus element, by deflating the ball during its flight, did affect performance in a real 3D catching task. While it might be tempting to ascribe any apparent difference in findings to a difference in task demands and thus even to different functional processing streams (e.g. Milner & Goodale, 1995), there is no need to do so. The effects reported by Savelsbergh et al. were small — initial grasping movements were delayed by about 5 ms — and entirely in keeping with the effects on intercept reported here and with the results of Vincent and Regan (1997). Moreover, though Savelsbergh's manipulation would have generated conflicts with other binocular cues, there was no conflicting motion information at a coarse spatial scale. We, too, have shown, in experiment 4, that fine-grained motion can determine performance when there is no conflict with coarse-grained motion.

Acknowledgements

This work was partly supported by EPSRC Grant GR/H/52986 and EC Grant ERB-FMB1-CT96-1251. Part of this work has previously been presented at ECVP, Oxford, 1998.

References

Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 21–29.

Cavallo, V., & Laurent, M. (1988). Visual information and skill level in time-to-collision estimates. *Perception*, *17*, 623–632.

Duffy, C. J., & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*, 1329–1345.

Freeman, T. C. A., Harris, M. G., & Tyler, P. A. (1994). Human sensitivity to temporal proximity: the role of spatial and temporal speed gradients. *Perception and Psychophysics*, *55*, 689–699.

Georgeson, M. A., Freeman, T. C. A., & Scott-Samuel, N. E. (1996). Sub-pixel accuracy: psychophysical validation of an algorithm for fine positioning and movement of dots on visual displays. *Vision Research*, *36*, 605–612.

Harris, M. G. (1998). Time-to-contact and heading direction for sparse non-evolving optic flow: data and models. *Perception*, *27*(suppl.), 141.

Horn, B. K. P., & Schunk, B. G. (1981). Determining optic flow. *Artifical Intelligence*, *17*, 185–204.

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*, 161–180.

Koenderink, J. J., & van Doorn, A. J. (1976). Local structure of movement parallax of the plane. *Journal of the Optical Society of America*, *66*, 717–723.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*, 437–459.

Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets: a paradigm of ecological optics. *Nature*, *293*, 293–294.

Lee, D. N., & Thomson, J. A. (1982). Vision in action: the control of locomotion. In D. T. Ingle, M. Goodale, & R. Mansfield, *Analysis of visual behaviour*. Boston: MIT Press.

Lee, D. N., Lishman, J. R., & Thomson, J. A. (1982). Regulation of gait in long jumping. *Journal of Experimental Psychology*: *Human Perception and Performance*, *8*, 448–459.

Lee, D. N., & Young, D. S. (1985). Visual timing of interceptive action. In D. T. Ingle, M. Jeanerod, & D. N. Lee, *Brain mechanisms and spatial vision*. Dordrecht: Martinus Nijhoff.

Lee, D. N., Young, D. S., Reddish, P. E., Lough, S., & Clayton, T. M. H. (1983). Visual timing in hitting an accelerating ball. *Quarterly Journal of Experimental Psychology*, *35A*, 333–346.

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London*, *Series B*, *208*, 385–397.

McLeod, R. W., & Ross, H. E. (1983). Optic-flow and cognitive factors in time-to-collision estimates. *Perception*, *12*, 417–423.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Perrone, J. A., & Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate cortex. *Vision Research*, *34*, 2917–2938.

Regan, D., & Hamstra, S. J. (1993). Dissociation of discrimination thresholds for time to contact and for rate of angular expansion. *Vision Research*, *33*, 447–462.

Regan, D., & Vincent, A. (1995). Visual processing of looming and time to contact throughout the visual field. *Vision Research*, *35*, 1845–1857.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of motion signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience*, *6*, 145–157.

Savelsbergh, G. J. P., Whiting, H. T. A., & Bootsma, R. (1991). Grasping tau. *Journal of Experimental Psychology*: *Human Perception and Performance*, *17*, 315–322.

Schiff, W., & Detwiler, M. L. (1979). Information used in judging impending collision. *Perception*, *8*, 647–658.

Todd, J. T. (1981). Visual information about moving objects. *Journal of Experimental Psychology*: *Human Perception and Performance*, *7*, 795–810.

Tresilian, J. R. (1991). Empirical and theoretical issues in the perception of time to contact). *Journal of Experimental Psychology*: *Human Perception and Performance*, *17*, 865–876.

Vincent, A., & Regan, D. (1997). Judging the time to collision with a simulated textured object: effect of mismatching rate of expansion of object size and texture element size. *Perception and Psychophysics*, *59*, 32–36.

Wagner, H. (1982). Flow-field variables trigger landing in flies. *Nature*, *297*, 147–148.

Warren, W. H., & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A*, *7*, 160–169.