



# Simulating the Detection of First-order Optical Flow Components

ASTRID M. L. KAPPERS,\*† SUSAN F. TE PAS,\* JAN J. KOENDERINK,\* ANDREA J. van DOORN\*

Received 15 August 1995; in revised form 30 October 1995; in final form 15 January 1996

**Thresholds for the detection of rotation and divergence in the presence of a translational component in sparse random dot patterns are determined for human observers and two computer algorithms. The algorithms only make use of local velocity directions and not of local velocity magnitude (speed). The results show that psychophysical performance in this task can be well described without the need of specialized mechanisms tuned to either rotation or divergence. Possibly, integration of information over more than two frames occurs for low velocities. For high velocities the correspondence problem seems to limit performance. Copyright © 1996 Elsevier Science Ltd.**

Simulation   Human psychophysics   Optical flow   Velocity direction

## INTRODUCTION

Ever since Gibson (1950) introduced the notion “optical flow field”, its importance as a source of information about the geometrical structure of the world around us and our position relative to it has been acknowledged (e.g., Koenderink & van Doorn, 1975, 1992; Rieger, 1983; Warren & Hannon, 1988; Warren *et al.*, 1991; Dijkstra *et al.*, 1994). Much of the information is contained in the local differential structure of the field and not in the speed or average direction of the flow. Information about the slant and tilt of objects and the relative movement of the observer is mainly contained in the first-order structure of the flow field. For tasks like the recognition of objects, higher-order structure is needed (for instance, for curvature up to second order). The first-order flow field can be decomposed into a number of elementary differential invariants, namely curl (vorticity, rotation), divergence (expansion/contraction, looming), and two components of the deformation (dilation and pure shear) (e.g., Koenderink & van Doorn, 1975, 1976; Longuet-Higgins & Prazdny, 1980; Koenderink, 1986).

This theoretical decomposition of the first-order flow field has led to a host of research, both psychophysical and electrophysiological, focussing on the question of whether or not the visual system actually makes use of such a decomposition (e.g., Regan & Beverley, 1978; Saito *et al.*, 1986; Tanaka *et al.*, 1989; Tanaka & Saito, 1989; de Bruyn & Orban, 1990, 1993; Lappin *et al.*, 1991; Freeman & Harris, 1992; Orban, 1992; Orban *et*

*al.*, 1992; Milne & Snowden, 1993; Regan, 1993; Graziano *et al.*, 1994; Kappers *et al.*, 1994; Snowden & Milne, 1994; Kappers *et al.*, 1996; te Pas *et al.*, 1996a). Comparison of the outcome of these experiments is not at all easy since a large variety of different experimental techniques has been exploited. More seriously, different criteria for the existence of specialized mechanisms are being used. As a consequence, some results which are presented as evidence for a decomposition performed by the visual system, should be interpreted as counter-evidence if other criteria are followed.

One such a criterion which should be satisfied, at least from a mathematical point of view, is the independence of the output of one such a specialized “detector” to the presence of other zero or first-order components in the stimulus. Orban (1992) and Orban *et al.* (1992) have shown that MST cells which are selectively tuned to either clockwise or counterclockwise rotation, or, contraction or expansion, are more or less insensitive to a translational component when added to their stimulus. However, the same cells were strongly influenced by the addition of a more complex component (e.g., expansion added to rotation resulting in spiral motion) which pleads against a decomposition into only first-order components. Graziano *et al.* (1994) report the existence of similar cells, but they add the finding of MST cells preferentially tuned to spiral motions. In psychophysical masking experiments, Freeman and Harris (1992) do indeed find that the detection of expansion was unaffected by the presence of rotation and vice versa. Similarly, our own psychophysical experiments have shown that the detection of rotation, divergence or deformation is independent of translational velocity (Kappers *et al.*, 1994, 1996; te Pas *et al.*, 1996a) and that the detection of divergence

\*Helmholtz Instituut, Princetonplein 5, 3584 CC Utrecht, The Netherlands.

†To whom all correspondence should be addressed.

is independent of rotation, and vice versa (te Pas *et al.*, 1996a). Although at first sight these results seem to support a decomposition into first-order flow components, other aspects of the results strongly suggest that such a decomposition is unlikely. te Pas *et al.* (1996a) do not really rule out the existence of different first-order components, but our results indicate that the underlying mechanisms of such detectors should at least be similar for all the components. Moreover, removing the mathematical divergence component from the stimulus (resulting in a stimulus with radial flow lines along which the dots decelerate instead of accelerate for an expansion) did not in any way affect our results (Kappers *et al.*, 1996). This latter finding makes the probability that specialized mechanisms play a role in our experimental paradigm rather small. In addition, the same series of experiments also showed that performance with a divergence stimulus with parallel instead of radial flow lines, deteriorated. This suggests that, at least in our task, the local velocity directions rather than the actual value of the divergence are of major importance. This is in line with conclusions drawn by, for instance, Warren and coworkers from their heading experiments (Warren *et al.*, 1991).

In this paper, we want to attack this problem from a new point of view. We asked ourselves the question whether it is possible to simulate our psychophysical results by means of a simple algorithm based on the use of local velocity directions. By using the coordinates of the random dots on the screen as input to the algorithm, human and algorithm performance can indeed be compared. For this comparison, we took the results of our best human observer (Kappers *et al.*, 1994, 1996).

We used two different versions of the algorithm. The first, termed Auto, is given information about which dots correspond to other dots from one frame to the next. This is evidently not realistic, but it provides a convenient base line. The second version, termed Scott, uses an algorithm proposed by Scott and Longuet-Higgins (1991) to find corresponding dots in successive frames. From the pairs of dots in successive frames we compute local velocity vectors. These vectors "vote" for either clockwise or counterclockwise rotation (in the rotation experiment) or, in the divergence experiment, for expansion or contraction. Combined evidence from all vectors of all pairs of successive frames (winner takes all) leads to a final decision (clockwise/counterclockwise or expansion/contraction). Adding noise to the stimulus gives us the possibility of determining thresholds in terms of noise levels.

With the choice of these two algorithms we certainly did not aim to construct an "ideal detector" for this particular task. Both algorithms discard intentionally some information like, for instance, correspondences between dots over more than two frames, the lengths of the vectors (that is, speed), and speed gradients. Our standpoint is that we use these simulation studies to full advantage if we incorporate in the algorithms only those

stimulus features which we think are of importance to the human observer.

We ran our simulations over the same wide range of conditions as we did in our psychophysical experiments. The values of both rotation and divergence, which have to be detected in the presence of a translational velocity, are varied between the minimum and maximum values as determined by screen resolution and stimulus size. Comparing the results obtained with Auto, Scott and our human observer, gives us the opportunity to study the influence of the correspondence problem in this experimental task. In order to gain an impression of the influence of the information content of the stimulus, the number of dots is varied, although the maximum of 64 dots which we used for our human observers was beyond the computational powers of our computer. The lifetime of the dots was varied in order to manipulate the local acceleration information. This latter parameter could only be of influence to performance of the human subject, since the algorithms do not make use of correspondence over more than two frames. Finally, we also tried our algorithms on some of the radial stimuli which had mathematically zero divergence.

## METHODS

### *Apparatus*

The stimuli for the psychophysical experiments were generated on an Atari MEGA ST4 computer and shown on an Atari SM125 high resolution monochrome monitor. Dark dots were shown on a light background. The monitor was viewed monocularly from a distance of 34 cm. The resolution of the display was  $400 \times 640$  pixels, corresponding to a field of view of  $21.1 \times 33.7$  deg of visual angle (pixel separation was 3.2 min arc). The simulation experiments were run on various types of Macintosh computers. The software was written in Mathematica.

### *Stimulus*

Stimuli consist of pseudorandom dot patterns. The spatial configuration of the dots is based on a regular hexagonal grid which is slightly perturbed with a two-dimensional Gaussian perturbation vector whose spread is  $\frac{1}{8}$  of the grid spacing. The number of dots per frame and the lifetime of the dots determine the grid spacing. Each dot consists of  $3 \times 3$  pixels. The diameter of the circular stimulus is always 380 pixels (20 deg). The upper left panel of Fig. 1 shows one frame of a 4-dot stimulus.

The total presentation time is kept constant at 16 frames (228 msec). The number of dots per frame is varied over the experimental sessions. Due to the statistical nature of the distribution of the dots over the stimulus, we can only give an approximate value of the maximum number of dots per frame. We use values of 1, 4, 16 or 64 dots; the latter value was only used in the psychophysical experiments. The lifetime of the dots is either 2, 3, 4 or 16 (maximum) frames.

In the rotation experiments, the deterministic compo-

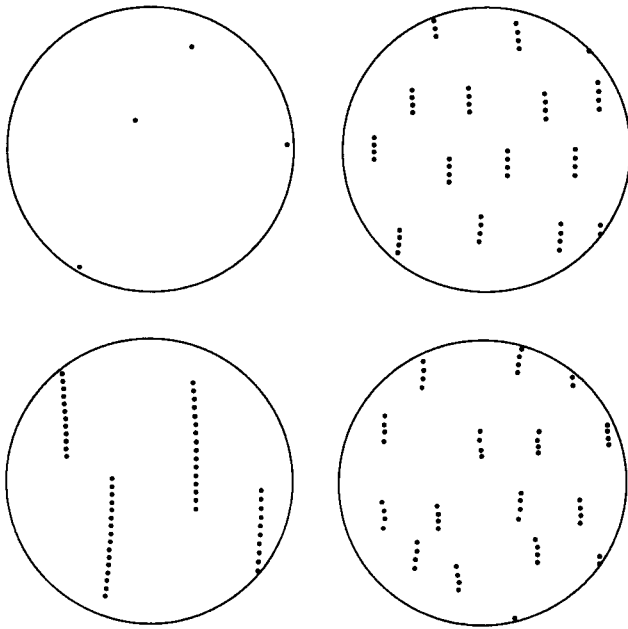


FIGURE 1. Examples of a rotation stimulus. The upper left panel shows one frame consisting of four dots. Of course, information about the flow type or direction is not available. In the other three panels, the 16 frames of the stimulus are superimposed. The lower left panel shows a rotation stimulus consisting of four dots per frame and a lifetime of 16 frames. The upper right panel shows again a 4-dot rotation stimulus, but this time the lifetime of the dots is only four frames. A similar stimulus with noise added is shown in the lower right panel. In all panels the curl and translational velocity are the same, namely 1 rad/sec and 40 deg/sec, respectively. The circular border of the stimulus is not actually shown to the subject.

nent of the stimulus consists of a curl (vorticity), either clockwise or counterclockwise, added to a downward directed translational velocity (henceforth called translation). As a consequence, the centre of rotation lies outside the stimulus region on a horizontal axis through the centre of the stimulus window. Rotation and translation can be varied independently and both cover a wide range, 0.03–128 rad/sec and 0.15–1280 deg/sec, respectively. Examples of this stimulus can be seen in Fig. 1.

In the divergence experiments, the deterministic component of the stimulus consists of a contraction or

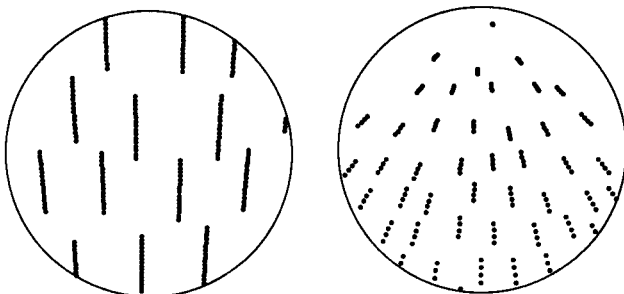


FIGURE 2. Examples of a divergence stimulus. Both stimuli contain 16 dots per frame. Left panel: divergence 0.5/sec, translation 20 deg/sec, lifetime 16 frames, contraction; right panel: divergence 4/sec, translation 20 deg/sec, lifetime four frames, expansion.

an expansion, again added to a downward directed translation. Due to the addition of the translational component, the centre of divergence lies outside the visible region on a vertical axis through the centre of the stimulus window. This results in stimulus patterns like in Fig. 2.

In addition to the regular divergence stimulus (termed Div), we also used a radial stimulus from which mathematically the divergence component was removed. The difference between this latter stimulus (termed Div\_no) and the Div stimulus lies in the velocities of the dots along the flow lines. For an expanding stimulus, the dots accelerate in the Div case, whereas they decelerate in the Div\_no case (and vice versa for contraction). Examples and a more detailed description can be found in Kappers *et al.* (1996). Subjects measured the whole range of conditions, but of the algorithms only Auto was tested for the 4-dot stimuli.

### Experimental procedure

After presentation of each stimulus, subjects had to decide whether a clockwise or counterclockwise rotation (or, a contraction or an expansion) was shown. Psychophysical thresholds were measured by jittering the deterministic positions of the dots. This was done by means of adding a two-dimensional Gaussian perturbation vector to the dot positions. The lower right panel of Fig. 1 shows an example of such a perturbed stimulus. A 2AFC-paradigm was used to determine threshold signal to noise ratios defined as the 75% correct noise levels. This procedure is described in much more detail in Kappers *et al.* (1994) and te Pas *et al.* (1996a,b).

The simulation experiments were run in parallel on all the Macintosh computers we had available in our laboratory. Even so, it took the evenings and weekends of almost half a year to finish all the simulations.

### Subjects

A number of subjects participated in our psychophysical experiments, but for comparison with our algorithms we only use the data of our “best” human observer, that is, the subject who reached the highest noise levels (that is, the lowest signal to noise thresholds) and had the widest range of measurable conditions. Other subjects differ in a quantitative and certainly not a qualitative way from this subject.

### Algorithms

Both algorithms use the integer values of the screen coordinates of the dots specified per frame as input. In addition, Auto is given information about corresponding dots in successive frames. For two different values of the curl and the translational velocity, examples are shown in the upper panels of Fig. 3. All corresponding pairs of dots in a sequence of 16 frames (the whole stimulus duration) are shown superimposed.

Since one of our aims was to learn something about the influence of the correspondence problem on performance in our task, we searched for a reasonable algorithm that is

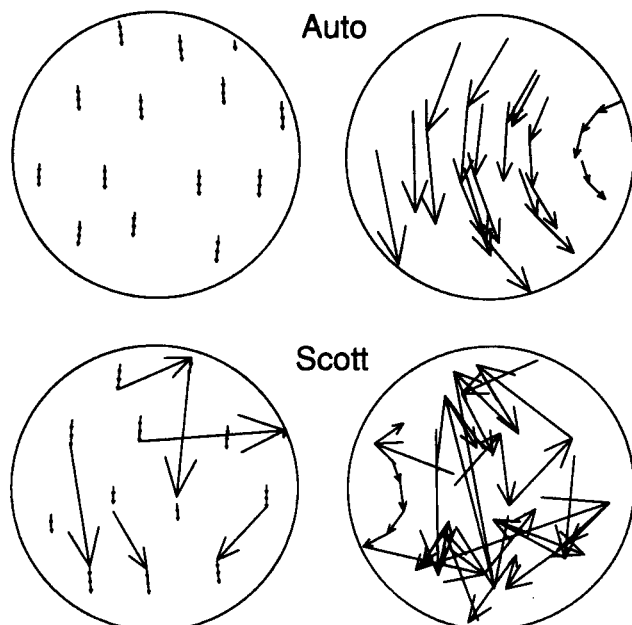


FIGURE 3. Examples of the solution of the correspondence problem. Again the 16 stimulus frames are superimposed. The upper panels show the results for Auto, the lower panels for Scott. Auto always has the correct solution, whereas Scott sometimes makes mismatches. Left panels, curl 1 rad/sec, translation 40 deg/sec; right panels, curl 64 rad/sec, translation 320 deg/sec. In all panels the lifetime is four frames and the number of dots per frame is four.

able to find the correspondences between dots in successive frames. An algorithm that fulfilled our requirements is the one proposed by Scott and Longuet-Higgins (1991). This algorithm, henceforth called Scott, operates on the distances between features in two related (in our case successive) images. Two principles underlie this algorithm: the principle of proximity and the principle of exclusion. The first principle requires that matches across shorter distances are to be favoured. This is established by choosing a suitable distance measure,  $\exp(-r_{ij}^2/2\sigma^2)$ , where  $r_{ij}$  is the distance between feature  $i$  (in our case dot  $i$ ) in the first image and feature  $j$  in the second image;  $\sigma$  can be considered as an adjustable scale parameter which gives an indication over what distance correspondences can be expected. The principle of exclusion prevents many-to-one feature correspondences. The algorithm maximizes the inner product of a proximity matrix (the elements of this matrix are the distances between all possible pairs of features in the two images using the above distance measure) and a pairing matrix. The elements of this latter matrix indicate the extent of pairing between features in the two images. From this pairing matrix follows the solution of the correspondence problem. For more details we refer to the paper of Scott and Longuet-Higgins (1991). Pilot experiments showed that the results do not depend critically on the value of  $\sigma$ , as long as it gives a rough estimate of the distances between corresponding features.

In our simulation experiments we made  $\sigma$  depend on the value of the translation ( $\sigma = \text{translation}/100$ ).

The two lower panels of Fig. 3 give examples of the correspondences obtained by Scott. In the left panel the values of the curl and translation are small (1 rad/sec and 40 deg/sec, respectively) and most often the correct correspondences are found. However, since the lifetime of the dots is limited (four frames), dots sometimes simply do not have a corresponding dot. Of course, this is information unknown to the algorithm and such dots are usually paired with newborn dots elsewhere in the stimulus. For higher values of the curl and translation other mismatches also occur, as can be seen in the lower right panel of Fig. 3 (curl 64 rad/sec; translation 320 deg/sec).

#### Determination of flow direction

Given the pairs of corresponding dots, determined either by Auto or by Scott, the next step is to determine the direction of the flow (that is, clockwise or counter-clockwise rotation, or, expansion or contraction). For the rotation and divergence stimuli similar but slightly different procedures are used. For both types of flow the direction of the moving dots is always downwards. This is known to the subjects, so we decided to give this information also to the algorithms. As a consequence, all arrows pointing upwards (see Fig. 3) are simply discarded. Although on first sight it may seem that this only is of influence to Scott, one should realize that when

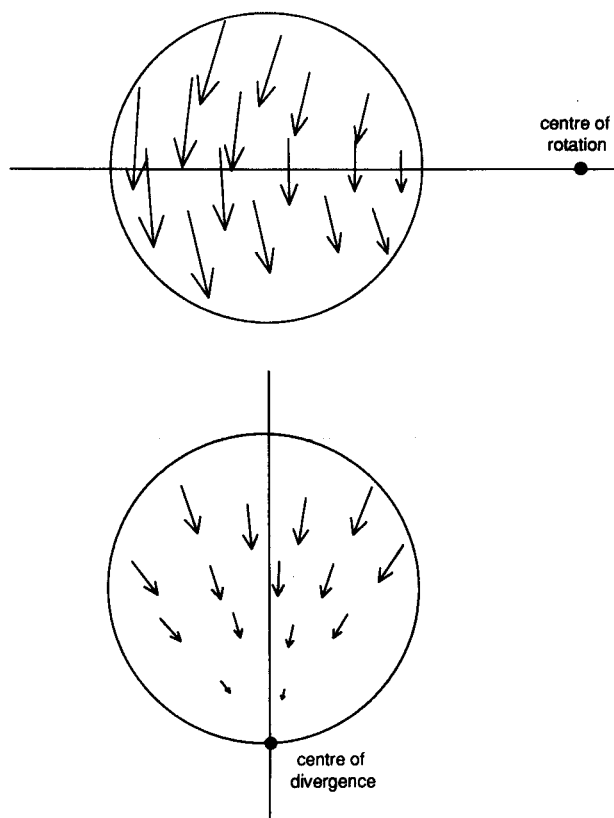


FIGURE 4. Schematic illustration of the flow directions in different parts of the stimulus. The upper panel shows a rotation, the lower panel a divergence. Further explanations can be found in the text.

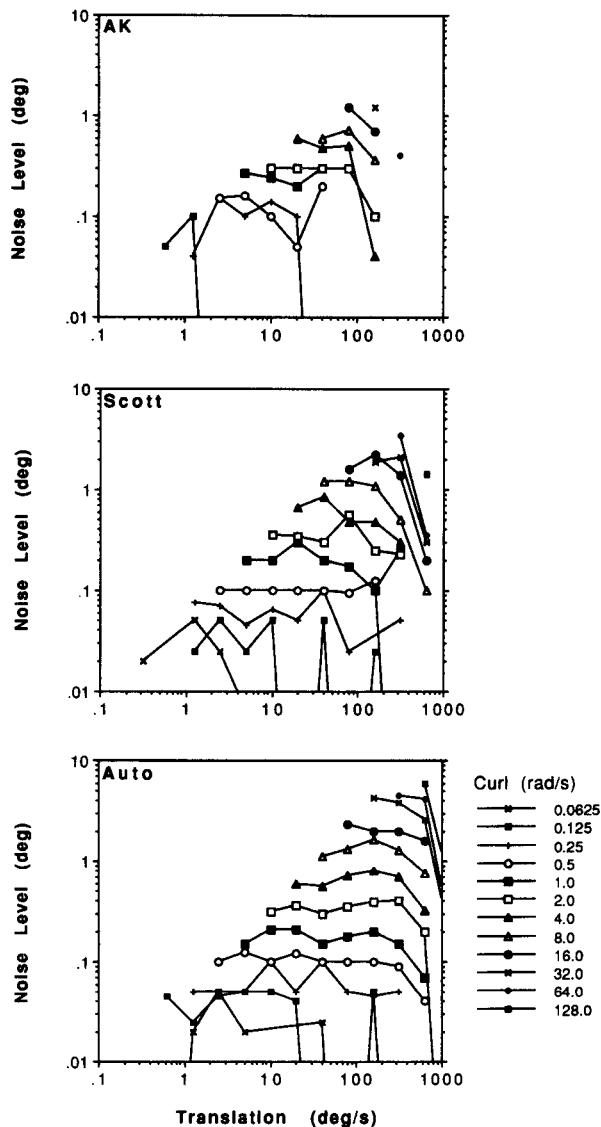


FIGURE 5. The three graphs present noise levels as a function of the translational velocity for the human observer (AK) and algorithms Scott and Auto. The number of dots was four and the lifetime of the dots was four frames. Curves represent different values of the curl (see legend).

noise is added to the stimuli, Auto will also have arrows pointing upwards. Next, arrows pointing exactly in the vertical direction are discarded, since they contain no information at all about the flow direction. The remaining arrows will be used to make a decision.

For rotation stimuli the centre of rotation always lies on a horizontal axis through the centre of the stimulus window (this is known to the subject). The consequence of this can be seen in the upper panel of Fig. 4. When the centre of rotation lies on the right side of the stimulus window (counterclockwise rotation) the arrows in the upper half of the stimulus point leftwards, whereas the arrows in the bottom part of the stimulus point rightwards (all relative to the vertical direction). The opposite is true for clockwise rotation. To determine the flow direction, we use the following procedure: arrows in the upper part

pointing leftwards and arrows in the lower part pointing rightwards "vote" for counterclockwise rotation. Similarly, arrows in the upper part pointing rightwards and arrows in the lower part pointing leftwards vote for clockwise rotation. The flow direction which collects most votes wins and determines the "answer" of the algorithm.

A similar procedure is followed for the divergence stimuli (see lower panel of Fig. 4). In that case the centre of flow always lies on a vertical line through the centre of the stimulus window. The voting procedure goes as follows: arrows in the left part pointing rightwards and arrows in the right part pointing leftwards vote for contraction; arrows in the left part pointing leftwards and arrows in the right part pointing rightwards vote for expansion. Again, the winner decides the flow direction.

## RESULTS

A representative illustration of our data is given in Fig. 5 where results measured using a rotating 4-dot stimulus with a lifetime of four frames are shown. Each curve gives the noise level as a function of the translation. Different curves belong to different values of the curl (see legend). The upper panel shows the results of subject AK; the middle and lower panels show results of algorithms Scott and Auto, respectively. The graphs contain data points for all conditions for which a threshold could be measured. Thus, differences in the number of data points in the three graphs reflect differences in performance. The translation value of the beginning of each curve is determined by our experimental paradigm: in order to locate the centre of rotation outside the visible stimulus window, a certain minimum translation is required. The value of this minimum translation depends on the value of the curl.

It can be seen that higher values of the curl result in higher noise levels. For a given value of the curl, performance (that is, noise level) is fairly independent of the translation until, for a certain maximum translation, performance drops steeply. Although always present, this sudden drop in performance is not always visible in the curves. Very often it was impossible to measure a threshold for the next value of the translation, even at the zero noise level. As conditions where performance is at chance level are not shown in the graphs, steep slopes do not always show up.

The most apparent difference between the graphs of AK, Scott and Auto is the number of data points which is highest for Auto and lowest for AK. For AK it becomes almost impossible to measure thresholds if the translational component lies above 100 deg/sec. The drop in performance occurs for much higher translations for Scott and even more so for Auto. Also, Auto and Scott can perform the task for higher and lower values of the curl than the human observer. The noise levels reached for each value of the curl are not too different for AK, Scott and Auto. Later in this paper, they will be compared in more detail.

As stated earlier, the results for AK are very similar to

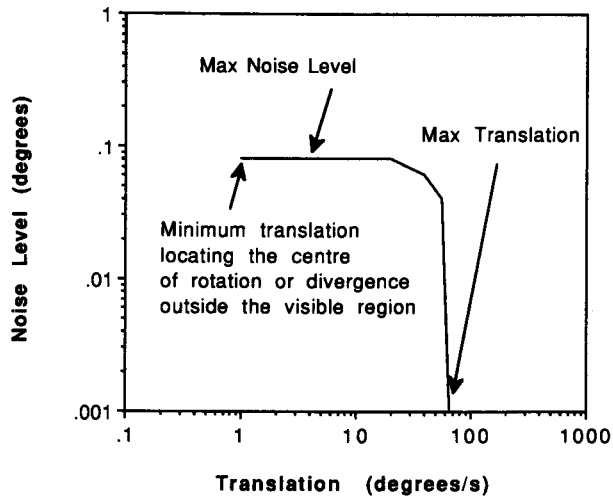


FIGURE 6. Schematic drawing of one of the curves of which Fig. 5 consists. Each curve can be characterized by two features: a region of a rather constant noise level (for reasons explained in the text this level is termed Max noise level) and a sudden drop in performance at a certain maximum translation. Also indicated in this figure is the fact that there is a relationship between the starting point of the curve belonging to a certain rotation or divergence, and the value of the translation.

those of other subjects. The above general description of the results is true for all of them. Differences show up in somewhat lower noise levels and a smaller range of measurable conditions (that is, the number of data points in the graph). Also, the results obtained with the divergence stimuli under the same condition (4 dots per frame and lifetime of four frames), do not in any way deviate from those obtained with the rotation stimuli. This latter finding also holds for both algorithms.

Like in our previous studies (Kappers *et al.*, 1994, 1996; te Pas *et al.*, 1996a,b), we will characterize the curves in Fig. 5 by means of two parameters, as is illustrated schematically in Fig. 6. As a rough but adequate estimate of the more or less constant noise level of each curve we take the maximum value of the noise level. The maximum translation is defined as the highest translation for which a threshold could be measured. In this way, we can easily compare performance for AK, Scott and Auto, and also for different conditions (number of dots per frame, lifetime of the dots).

In Fig. 7 the maximum noise level is shown as a function of the curl for three different lifetimes of the dots. The stimulus was again a 4-dot rotation. Different curves show results for AK, Scott and Auto. Clearly, all maximum noise levels increase with curl, much the same for AK as for the algorithms. For high values of the curl, there is a distinct order in performance. Auto always performs best closely followed by Scott. In that range, performance of the human observer lies significantly below that of the two algorithms. The situation is different for the smaller values of the curl. Auto and Scott are not really different from each other, but both perform less than subject AK. Here it should be

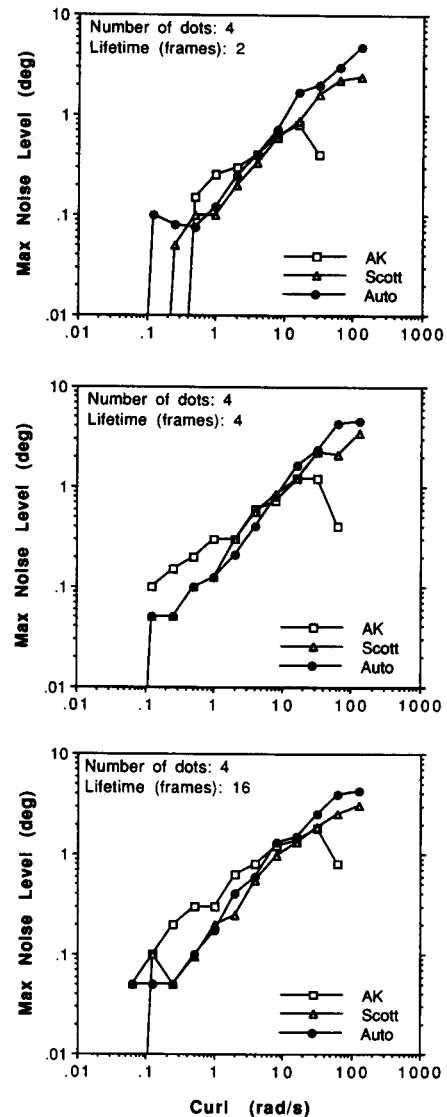


FIGURE 7. Maximum noise level as a function of the curl. Graphs show data for different lifetimes of the dots. The number of dots is always four per frame. Curves show performances of AK, Scott and Auto.

mentioned, that performance of the other subjects lies below that of the algorithms. Comparing the three graphs, it can be seen that the influence of lifetime of the dots is at most minor. Performance very slightly decreases for shorter lifetimes. This is true for the subject as well as for the two algorithms.

Once again, the results look very similar to those obtained with the divergence stimuli (not shown here). In addition, maximum noise levels obtained with the Div\_no stimulus are almost indistinguishable from those of the actual divergence stimuli. For the subjects, there really is no difference (Kappers *et al.*, 1996). For Auto, the only algorithm tested with the Div\_no stimulus, the maximum noise levels obtained with the Div\_no stimulus differ slightly from those of the Div stimulus, but only for the highest values of the divergence. This difference does not seem to be systematic; for a lifetime of two frames

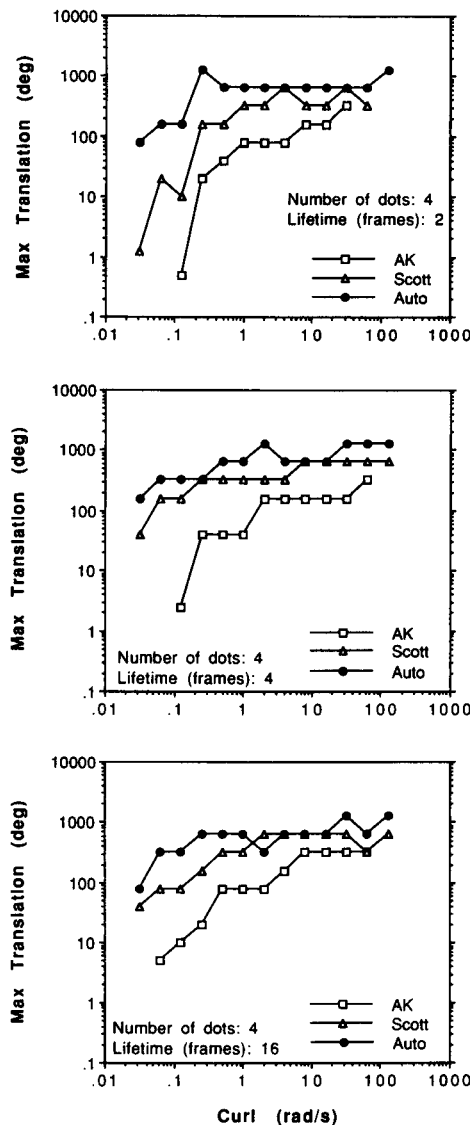


FIGURE 8. Maximum translation as a function of the curl. Graphs show data for different lifetimes of the dots. The number of dots is always four per frame. Curves show performance of AK, Scott and Auto.

Div\_no performs better, but the opposite is true for a lifetime of 16 frames. There is no difference for a lifetime of four frames.

Roughly, the maximum noise levels obtained with stimuli containing a different number of dots per frame look very much like those presented in Fig. 7. As we found in our previous studies, the actual number of dots hardly has any influence on the maximum noise level, as long as there are sufficient dots (for most subjects this is four). For 16-dot rotation stimuli, the performance of AK is nowhere better than that of Auto, not even for the smallest values of the curl. For 1-dot stimuli, Scott's performance usually lies below that of Auto and AK.

In Fig. 8, the maximum translation is shown as a function of the curl for the same conditions as in Fig. 7. In all three graphs, it can clearly be seen that Auto lies above

Scott, and both lie above the human observer. The maximum translation increases with curl, although this is more evident for AK than for the two algorithms. It seems as if performance saturates for higher values of the curl, that is, translations higher than a certain maximum cannot be reached. Most probably it indicates at what point relations between dots in succeeding frames vanish. Comparing the three graphs, it can be seen that the lifetime of the dots does not have any influence on the maximum translation.

The results obtained with either 1 or 16-dot stimuli, are much the same as those shown in Fig. 8. Levels of performance of Auto, Scott and AK always occur in the same order. There is only a minor effect of the number of dots: the actual values of the maximum translation very slightly increase with the number of dots. Results obtained with the Div and Div\_no stimuli are identical to those of the rotation stimuli.

## DISCUSSION

The experiments described in this paper show convincingly that our psychophysical results could be simulated by means of simple algorithms which only make use of local velocity directions. Clearly, there was no need to implement specialized mechanisms selectively sensitive to either rotation or divergence. All important features of the psychophysical results, such as the constant noise level as a function of translation, the maximum noise level as a function of curl or divergence, the maximum translation, the influence of lifetime and number of dots, and the similarity of rotation and divergence results, are captured by the algorithms. Although simulation experiments cannot, of course, be decisive, they add to the already existing evidence (e.g., Warren *et al.*, 1991; Kappers *et al.*, 1996) that in psychophysical tasks like ours, the local velocity direction is of major importance for performance.

Since the results of the simulation experiments are so similar to those of the psychophysical experiments, it is of interest to study the deviations in more detail. The most apparent difference was shown in Fig. 8, where the maximum translation was plotted as a function of the curl. The level reached by Auto was higher than that of Scott, and the levels of both Auto and Scott were higher than that of the human observer. Performance of Auto is clearly determined by stimulus characteristics. The stimulus window was 20 deg and combined with a frame rate of 70 Hz, the maximum translation which could be presented was 1400 deg/sec. Of course, the actual maximum translation must lie below this value, because in the exceptional situation of 1400 deg/sec at most one dot contains information; all other dots must be considered as noise. Sometimes, Auto manages to perform the task at a translational velocity of 1280 deg/sec, but apparently this value still lies too close to 1400 deg/sec since otherwise Auto should have been able to do the task for the zero noise level. For a translation of 640 deg/sec, the stimulus contains sufficient information for Auto. For Scott, however, most of the time this

velocity is still too high. The cause must be the correspondence problem, since the only difference between Scott and Auto lies in knowledge about corresponding dots. As the difference between the human observer and Scott is similar to that of Scott and Auto, it is reasonable to assume that the comparatively bad performance of the human subject is due to difficulties in matching corresponding dots.

Human performance was also worse than that of the algorithms for the highest values of the curl (and divergence) as was shown in Fig. 7. In those situations, again high local velocities occur within the stimulus window. Thus, most probably, the correspondence problem explains this difference in performance.

In Fig. 7 it was also shown that for the lowest values of the curl, subject AK's performance was better than that of the two algorithms. Although this effect was not found for our other human observers, we think it is important enough to pay some attention to it. Better performance of the human observer suggests that she makes use of information discarded by the algorithms. An obvious possibility is that the subject correlates dots over more than two frames. A prerequisite for this explanation is that the advantage for the subject disappears for stimuli in which the lifetime of the dots is only two frames. In such stimuli, no information can be gained by correlation over more than two frames. For the 1 and 16-dot rotation stimuli, it is indeed the case that performance of the subject equals that of Auto, but it is only partially true for the 4-dot rotation stimulus shown in Fig. 7. Although this remaining difference might be due to statistical fluctuations (for instance, in this case the maximum noise level seems to be a too high estimate for the constant noise level), other possibilities should not be excluded. An alternative explanation is that the subject somehow incorporates the length of the velocity vectors, that is, the velocity magnitude or speed. As can be seen in Fig. 4, there exists a speed gradient in both the rotation and the divergence stimulus. In a previous study (Kappers *et al.*, 1996), however, we performed similar experiments with a radial stimulus in which the dots moved with a constant speed, thus effectively eliminating the gradient. The results obtained with this stimulus (termed Div const) were indistinguishable from those using the Div and Div\_no stimuli, which clearly argues against an explanation using speed gradients. It is also feasible that in stimuli with small values of the curl, the human subject simply discards the smaller velocity vectors since their directions are more severely perturbed by noise than the directions of the larger ones. Taking all evidence together, we opt for the integration of information over more than two frames possibly in combination with selective attention to the higher velocities.

In conclusion, we think that simulation experiments are very helpful in gaining insight into our understanding of how human observers use the information contained in the optical flow. Although the voting procedure, in particular, is fairly specific to our psychophysical task, we think that, for instance, the heading experiments of

Warren *et al.* (1991) and the experiments of de Bruyn & Orban (1990, 1993) might be simulated with slightly adapted algorithms. Unfortunately, further investigations are time-consuming and they necessarily must lie outside the scope of this paper. It remains an interesting question for future research whether indeed other psychophysical results could be simulated with algorithms similar to ours. Combined evidence of various psychophysical and simulation experiments will eventually lead to more definitive conclusions about the importance of the local velocity directions.

## REFERENCES

- de Bruyn, B. & Orban, G. A. (1990). The role of direction information in the perception of geometric optic flow components. *Perception & Psychophysics*, 47, 433–438.
- de Bruyn, B. & Orban, G. A. (1993). Segregation of spatially superimposed optic flow components. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 1014–1027.
- Dijkstra, T. M. H., Snoeren, P. R. & Gielen, C. C. A. M. (1994). Extraction of three-dimensional shape from optic flow: A geometric approach. *Journal of the Optical Society of America A*, 11, 2184–2196.
- Freeman, T. C. A. & Harris, M. G. (1992). Human sensitivity to expanding and rotating motion: Effects of complementary masking and directional structure. *Vision Research*, 32, 81–87.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Graziano, M. S. A., Andersen, R. A. & Snowden, R. J. (1994). Tuning of MST Neurons to spiral motions. *Journal of Neuroscience*, 14, 54–67.
- Kappers, A. M. L., van Doorn, A. J. & Koenderink, J. J. (1994). Detection of vorticity in optical flow fields. *Journal of the Optical Society of America A*, 11, 48–54.
- Kappers, A. M. L., te Pas, S. F. & Koenderink, J. J. (1996). Detection of divergence in optical flow fields. *Journal of the Optical Society of America A*, 13, 227–235.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26, 161–170.
- Koenderink, J. J. & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22, 773–791.
- 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.
- Koenderink, J. J. & van Doorn, A. J. (1992). Second-order optic flow. *Journal of the Optical Society of America*, 9, 530–538.
- Lappin, J. S., Norman, J. F. & Mowafy, L. (1991). The detectability of geometric structure in rapidly changing optical patterns. *Perception*, 20, 513–528.
- Longuet-Higgins, H. C. & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society London B*, 208, 385–397.
- Milne, A. B. & Snowden, R. J. (1993). Is there anything special about expansion and rotational flow fields? *Perception*, 22(Suppl.), 95.
- Orban, G. A. (1992). The analysis of motion signals and the question of the nature of processing in the primate visual system. In Orban, G. A. and Nagel, H.-H. (Eds), *Artificial and biological vision systems* (pp. 24–56). Berlin: Springer.
- Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H. & Torre, V. (1992). First-order analysis of optical flow in monkey brain. *Proceedings of the National Academy of Science U.S.A.*, 89, 2595–2599.
- te Pas, S. F., Kappers, A. M. L. & Koenderink, J. J. (1996a). Detection of first-order structure in optic flow fields. *Vision Research*, 36, 259–270.
- te Pas, S. F., Kappers, A. M. L. & Koenderink, J. J. (1996b). Detection of optical expansion as a function of field size and eccentricity. *Perception & Psychophysics*, 58, 401–408.

- Regan, D. (1993). The divergence of velocity and visual processing. *Perception*, 22, 497–500.
- Regan, D. & Beverley, K. I. (1978). Looming detectors in the human visual pathway. *Vision Research*, 18, 415–421.
- Rieger, J. H. (1983). Information in optical flows induced by curved paths of observation. *Journal of the Optical Society of America*, 73, 339–344.
- Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience*, 6, 145–157.
- Scott, G. L. & Longuet-Higgins, H. C. (1991). An algorithm for associating the features of two images. *Proceedings of the Royal Society London B*, 244, 21–26.
- Snowden, R. J. & Milne, A. B. (1994). Evidence for the existence of mechanisms tuned for complex motions in the human visual system. *Investigative Ophthalmology & Visual Science*, 35, 1726.
- Tanaka, K., Fukada, Y. & Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction, and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, 62, 643–656.
- Tanaka, K. & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, 62, 626–641.
- Warren, W. H. Jr, Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G. & Kalish, M. L. (1991). On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics*, 65, 311–320.
- Warren, W. H., Jr & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, 336, 162–163.

---

**Acknowledgements**—This research is supported by the Insight II project of the ESPRIT Basic Research Action of the European Commission and the Netherlands Organization for Scientific Research (NWO).