High temporal frequency synchrony is insufficient for perceptual grouping

Michael Morgan[1*] and Eric Castet[2]

[1]*Applied Vision Research Centre, The City University, Northampton Square, London EC1V 0HB, UK*
[2]*Centre de Recherche en Neurosciences Cognitives (CRNC), FRE 2098 du CNRS, 31 chemin J. Aiguier, 13402 Marseille Cedex 20 France (castet@lnf.cnrs-mrs.fr)*

We used textures of randomly moving grating patches to assess the role of fine-grain temporal synchrony in texture segregation. In the target area, patches reversed direction simultaneously. In the surround, patches changed direction at random times. Thus, phase changes in the target area were precisely synchronous, whereas those in the surround were not. In agreement with work carried out by Lee and Blake, we found that the target area was frequently visible, and that observers could discriminate its shape (horizontal versus vertical) at frame rates of 100 Hz in brief exposures (200 ms). Further experiments suggested that the length of unidirectional motion sequences in the target area, rather than synchrony, determined its visibility. To eliminate completely contrast and motion cues, we made all the background elements identical to the target elements, but with a random starting phase. Despite the presence of synchrony in the target area but not the background, the target was generally very hard to see. Targets that remained visible contained low temporal frequency modulations of direction. We conclude that the human observer can detect synchrony, but only at modest temporal frequencies once motion and contrast artefacts have been eliminated.

Keywords: vision; motion; texture; synchrony

1. INTRODUCTION

Physiologists and psychophysicists have become interested in the possibility that precise temporal synchrony of nerve impulses in different parts of the brain is the solution to the 'binding problem' by which different attributes, such as shape and colour, are seen to belong to the same object (Singer 1999a,b; Castelo-Branco et al. 2000). This suggestion has been controversial (Lee & Blake 1999; Shadlen & Movshon 1999). On general grounds, it is evident that the temporal dimension is important in object recognition: we would not perceive the Cheshire Cat (Carroll 1865) if its grin, body and colour appeared in appropriately adjacent locations on different days. What is at issue is the fineness of temporal synchrony used for grouping.

There have been claims that very fine-grain synchrony in the millisecond region facilitates object segregation (Fahle 1993; Usher & Donnelly 1998; Lee & Blake 1999). Fahle (1993) found that an area of synchronously flickering dots segregated perceptually from a surround with delays of 5 ms. To overcome the problem of motion cues between target and surround inherent in such displays, Lee & Blake (1999) used fields of grating patches (figure 1) within which the phase of the grating changed on every frame at a 100 Hz refresh rate. The direction of the phase change in relation to the change in the previous frame was random, but was the same for all elements in the target area. Elements in the background changed their phase according to independent motion-reversal sequences (point processes). Thus, all the elements in the target area changed their motion direction synchronously, whereas those in the background changed direction asyn-chronously. As the orientation of all the elements was random, and the direction of the first phase displacement was also random, it was argued that segregation of the target area could not be based upon motion direction cues. However, Farid & Adelson (2000, 2001) pointed out that low-pass temporal filtering might suffice to explain segregation in the Lee and Blake stimulus. If a grating changes its phase between 0 and 180° at a sufficiently high rate, its apparent contrast is reduced because the two phases effectively cancel out. In general, any point process that generates phase change will produce a time-averaged contrast that is particular to that point process. As all the elements in the target area are generated by the same point process, they will have the same apparent contrast, whereas those in the background will produce a patchy contrast. Segregation could therefore be based upon contrast-variance discrimination following temporal filtering.

Another potential cue in the Lee & Blake (1999) display is motion itself. If, by chance, one element in the target area has a point process that generates a long run of phase shifts in the same direction, it will appear to be in smooth movement; and because all the elements in the target area have the same point process, they will all appear to move. The direction of movement will be random, but by chance may be biased in a particular direction, and it is established that coherent overall motion may be seen, despite large variation in the direction of individual elements. To examine the effects of contrast and motion cues on target visibility in displays like those used by Lee and Blake, we fixed the point process for the target area, leaving the sequences in the background random. To investigate the effects of contrast, the elements either had all the same contrast (1.0) or their contrast was randomly sampled from a uniform distribution with mean of 0.5 and range of 1.0. The aim of this 'perturbation' method (Creelman &
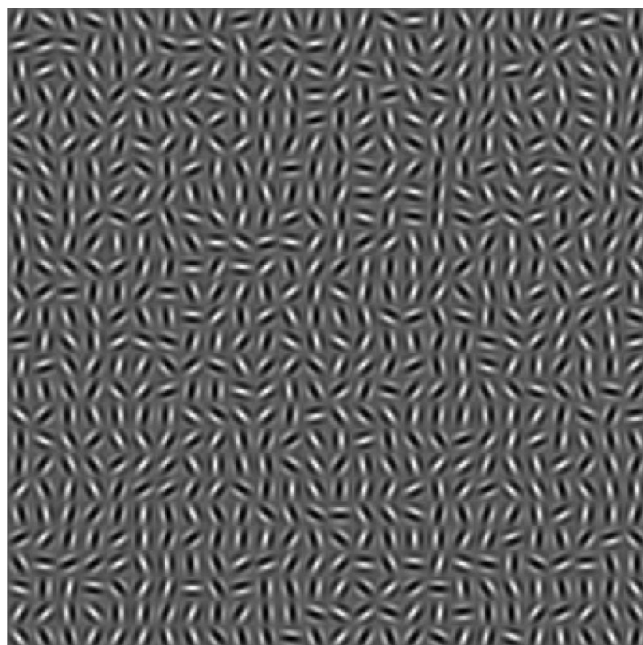
*Author for correspondence (m.morgan@city.ac.uk).

Proc. R. Soc. Lond. B (2002) **269**, 513–516
DOI 10.1098/rspb.2001.1920

513

© 2002 The Royal Society

Figure 1. An actual example of one frame of the stimulus used in our experiments. In the next frame, all the Gabor elements would phase shift by ±72°, the direction being random over elements. In following frames, changes in direction would be synchronous for elements in the target area, but asynchronous in the background. The target area (invisible here) was an 8 × 4 or 4 × 8 array of patches in the centre of one the four quadrants of the display.

Macmillan 1979) was to camouflage any subtle differences in contrast between target and surround by introducing large random-contrast 'patchiness' over the whole array. (The logic is identical to that of the pseudo-isochromatic colour vision plates, in which random luminance variation camouflages luminance but not colour differences between target and surround.) To investigate motion, we devised a method to equate the motion energy between the target and surround. This was done by sampling the background elements from the target area, but randomizing their spatio-temporal phase.

2. METHODS

An example of one frame of our movies is shown in figure 1. The orientation distribution of the individual patches was Gaussian with a mean of 90° (vertical) and a standard deviation of 45°. Stimuli were presented on a Sony Trinitron RGB monitor (GDM-F500T9) under control of a Cambridge Research Systems VSG 2/3 board in a host PC. The monitor refresh rate was 100 Hz. Programs for generating the stimuli were written in MATLAB. Each frame of the display consisted of a 26 × 26 array of Gabor patches, each containing a 5.8 cycles deg^{-1} sinusoidal grating windowed with a Gaussian envelope. The mean luminance of the patches was 36 cd m^{-2}, and unless otherwise stated the Michelson contrast was 1.0. The overall dimension of each frame was 7.55° × 7.55°, viewed from 2 m. The phase of the gratings was randomized within patches in the range 0–2π. Sequences of 20 frames were stored in the VSG memory in which the phase of the grating was changed on each frame by ±1/5 of the grating cycle (72°, as in Lee & Blake (1999)). This resulted in an irregular series of motion steps in which the

grating either moved in the same (+) or opposite (−) direction to the previous step. The probability of + and − steps were equal. We refer to the resulting series of steps as the point process (cf. Lee & Blake 1999) for that element.

All the elements in the target area had the same point process; those in the background were randomly selected, unless otherwise stated. The direction of the first phase shift was randomized over elements, unless otherwise stated. The target area was either horizontal (8 × 4 patches) or vertical (4 × 8 patches), and was placed in the centre of one of three screen quadrants: top left, bottom right or bottom left (shortage of memory prevented use of all the quadrants, as all the possible stimulus sequences were computed and stored before each experiment). The target position was randomized over trials and the observer's task was to report whether target orientation was horizontal or vertical (not its position). Exposure duration was 200, 600, 1000, 1400 or 1800 ms, unless stated otherwise. A full 20-frame sequence took 200 ms. Longer durations were obtained by wrap-around, which must be taken into account when assessing the point process. Each combination of exposure, screen position and target shape (horizontal versus vertical) was repeated 10 times inside a block of 300 trials. At the end of each block the screen went blank and the observer took a rest. There was no fixation point and the observer was free to make eye movements to search the display. There was no error feedback, unless stated otherwise.

The stimuli were precomputed and stored before each block of 300 trials. Each of the six combinations of target position and shape had a different random number seed, and thus a different set of random Gabor orientations and point processes (except in the target area, where the point processes, but not the orientations, were always the same). Thus, observers could not identify the target shape by searching for a patch with a particular orientation and point process. The absence of error feedback meant that they could not learn to associate the presence of a particular Gabor patch with a correct response.

Careful note was taken of all the seeds used to start the random number generators so that every experimental condition could be reproduced.

3. RESULTS

(a) *Experiment 1: effects of contrast randomization and motion sequence length on visibility*

The first experiment was essentially a replication of Lee & Blake (1999), except that we controlled the point processes rather than allowing them to vary randomly, and also compared fixed-contrast and random-contrast conditions. In the former, all the elements had the same contrast (1.0); in the latter, each grating patch had its contrast randomly chosen from a uniform distribution in the range 0–1. Contrast did not vary within a sequence. When we examined stimuli generated by the methods described by Lee and Blake, it was obvious that they were very different in difficulty, depending on the seed given to the random number generator. Easy sequences were those in which there were long run lengths in the same direction in the target area (e.g. LLLLLL ...) or those in which there was rapid alternation (e.g. LRLRLR ...). There was thus a U-shaped relationship between visibility and temporal frequency of alternation (figure 2). By inspection, long run sequences were visible because they were of uniformly low contrast and because they presented a strong motion
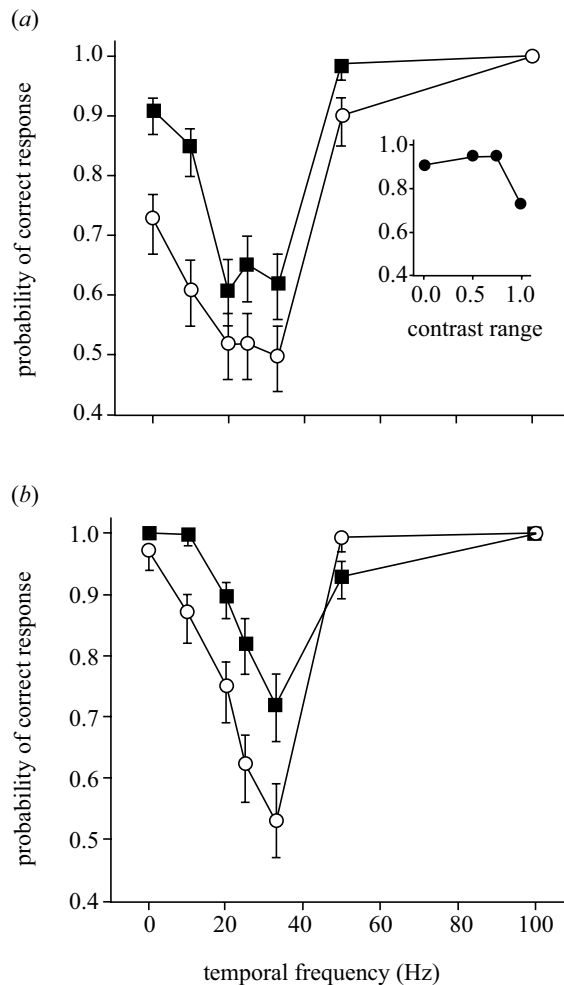
(a)

(b)

Figure 2. Results from the first experiment, examining the visibility of temporal frequency of reversal (horizontal axis) in fixed (squares) versus random (circles) contrast conditions. Results are shown separately for two observers (*a*) M.M. and (*b*) E.C. Because there was little effect of exposure duration on accuracy, the results from the different exposures have been combined.

signal. Alternating sequences were visible because they appeared high contrast, stationary, or to shiver slightly. The high visibility of long runs and alternations agrees with the suggestion that contrast is an important cue for segregation (Adelson & Farid 1999), as these sequences would be expected to have low and high time-averaged contrast, respectively. To control for the effects of contrast, we investigated the effects of contrast randomization of the individual grating patches (figure 2). Lee & Blake (1999) informally reported that contrast randomization over elements had little effect. Our findings agree broadly with theirs. Even the extreme contrast randomization we used, which made some elements invisible, failed to perturb performance with the easiest sequence (alternation at 100 Hz). Nor was there a significant effect of contrast randomization in the continuous motion case (0 Hz) for observer E.C. (figure 2*b*). Additional observations, shown in the inset to figure 2*a*, revealed that the performance drop for M.M. in the latter condition occurred only with extreme randomization. There was no effect when the range was 0.25–1.0. It can be plausibly argued that extreme randomization had its effect by making a pro-

portion of the elements invisible. The first experiment has therefore failed to confirm the hypothesis that synchrony is detected entirely by a contrast-sensitive mechanism, although contrast may be one of the cues involved. We therefore suspected that motion might be involved, and the second experiment was devised to provide a stringent test of that possibility.

(b) *Experiment 2: equating the point processes between target and surround*

We conclude from the first experiment that both contrast and motion are potential cues for object segregation. We could not rule out temporal synchrony as an additional cue. To isolate synchrony from all other cues, we devised a new way to generate the display. First, a series of frames defining the target area was generated, exactly as before. Then the background was generated, selecting every background patch from the set of target patches, but with a random starting frame. For example, if all the target patches had the sequence LLLLRRLLLLRR ..., a particular background patch might have the sequence LLRRLLLLLRR ... and another background patch the sequence RRLLLLRRLLLL .... Thus, all the target patches changed direction synchronously, whereas those in the background had exactly the same temporal frequencies of reversal, but changed asynchronously. (Note that the correlation between a random binary sequence and a phase-shifted version of the same sequence is zero.) This procedure equates the first-order spatio-temporal statistics between target and surround, leaving only second-order differences (phase). There was no possibility of a contrast cue or a differential motion cue in these stimuli, and none was ever seen. The results are easily summarized without the aid of a figure. None of the sequences visible in the first experiment (figure 2) could be seen at all in this experiment by either observer. Most randomly determined reversal sequences were also invisible, with rare exceptions. In these exceptional cases, the target area seemed to undergo a low temporal frequency pulsation. Inspection revealed that these cases had runs alternating with stationary flicker. To check this, we examined the visibility of the sequence LLLLLLLLLLLRLRLRLRLRL. The Fourier spectrum of this stimulus shows a strong periodicity at 5 Hz (note that the full sequence length is $20 \times 10 = 200$ ms). The target was highly visible to observer E.C. (96% correct across durations and across three different seeds) and could be seen by observer M.M. in long exposure, although it failed to 'pop out' convincingly.

4. CONCLUSIONS

Our experiments suggest that the synchronous target area in the original study by Lee & Blake (1999) stands out perceptually from the background for a complex variety of reasons, none of them necessarily based on synchrony itself. The easiest cases to see, as conjectured by Farid & Adelson (2000), are those in which the target area has either very short or very long run lengths of motion: these are exactly the conditions that would be expected to produce contrast differences through a low-pass temporal filter. Our attempt to destroy the visibility of structure with these sequences by randomizing element contrast was

only partly successful, and extreme randomization was needed to have any effect (see figure 2). Even if an alternating point process changed effective contrast by as little as 0.1, this may still have been visible against the variation of contrast ($\sigma = 0.288$) in our display. Contrast randomization had a much clearer effect on the visibility of point processes such as double alternation, which were less visible to start with (figure 2). The reason this evidence is not conclusive is that ordinary static texture segmentation can sometimes be disrupted by variation in an irrelevant stimulus dimension. For example, segmentation based on orientation can be disrupted by randomly colouring the texture elements red and green (Morgan *et al.* 1992). Thus, it would be possible to maintain that segmentation in the Lee & Blake (1999) display is based on synchrony, but is disrupted by contrast variation.

A stronger test is to eliminate the contrast cue entirely while maintaining the synchrony cue, as we did in experiment 2. In that experiment, all the elements in the display had the same point process, and thus there was no question of contrast or motion cues leading to segregation. All the elements in the target area were synchronous, whereas those in the surround had random phase relations. We found identification of the target area impossible in most displays of this kind, arguing that synchrony is at best a very weak cue to segregation when it is not accompanied by temporal frequency or contrast cues. A possible rejoinder is that making the point processes in the surround identical increased the degree of their synchronization. This would not be true if the point processes had been infinitely long, as the correlation of a random binary sequence with its phase-shifted version is by definition zero. However, our sequences were only 20 frames long, so there was a 1 in 20 chance that two randomly-selected versions would be synchronous. It is hard to believe that such a low probability would have made the surround sufficiently synchronous to be indistinguishable from the target, where the probability of synchrony was unity.

We conclude that synchrony is unlikely to be the explanation of texture segmentation in the Lee & Blake (1999) display. Known mechanisms such as temporal filtering are equally plausible and are consistent with the evidence presented in this paper. It does not follow that synchrony can be rejected as a solution to the 'binding problem' in normal perception, only that this attempt to reveal it as a necessary and sufficient condition has failed. The Lee and Blake task is an unusual and difficult one. Target elements have as much in common with non-target ones in terms of their orientation and spatial frequency structure, and synchrony on its own may be insufficient to work in the absence of these additional cues to segregation.

We agree (as is obvious on logical grounds) that synchrony of relatively low temporal frequency modulations of motion and contrast can be the basis for segregation, in spite of some differences between observers, but we argue that the temporal grain is not at the 1000 Hz rate that would be required for synchrony of individual neural spikes. A recent psychophysical study also suggests a modest effect of synchrony. Kandil & Fahle (2001) found that synchronously rotating elements could be segregated from an asynchronous background at a rate of 8 Hz but not at 32 Hz. There is no evidence in our data for a role of synchrony when the first and second moments of the contrast and motion statistics are controlled. In the case of segregation of static stimuli by orientation, we have a good understanding of the statistical regularities that can and cannot be extracted by the observer (Dakin & Watt 1997). An equivalent understanding of spatio-temporal statistics is now needed, rather than the inspection of uncontrolled randomness.

REFERENCES

Adelson, E. & Farid, H. 1999 Filtering reveals form in temporally-structured displays. *Science* **286**, 2331.

Carroll, L. 1865 *Alice's adventures in Wonderland*. London: The Bodley Head.

Castelo-Branco, M., Goebel, R., Neuenschwander, S. & Singer, W. 2000 Neural synchrony correlates with surface segregation rules. *Nature* **405**, 685–689.

Creelman, C. D. & Macmillan, N. A. 1979 Auditory phase and frequency discrimination: a comparison of nine procedures. *J. Exp. Psychol. Hum. Percept. Perform.* **5**, 146–156.

Dakin, S. C. & Watt, R. J. 1997 The computation of orientation statistics from visual texture. *Vis. Res.* **37**, 3181–3192.

Fahle, M. 1993 Figure-ground discrimination from temporal information. *Proc. R. Soc. Lond.* B **254**, 199–203.

Farid, H. & Adelson, E. H. 2000 Standard mechanisms can explain grouping in temporal synchronous displays. *Investigative Ophthalmol. Vis. Sci.* **4**, S438.

Farid, H. & Adelson, E. H. 2001 Synchrony does not promote grouping in temporally structured displays. *Nat. Neurosci.* **4**, 875–876.

Kandil, F. & Fahle, M. 2001 Purely temporal figure-ground segregation. *Euro. J. Neurosci.* **13**, 2004–2008.

Lee, S.-H. & Blake, R. 1999 Visual form created solely from temporal structure. *Science* **284**, 1165–1168.

Morgan, M. J., Adam, A. & Mollon, J. D. 1992 Dichromats break colour-camouflage of textural boundaries. *Proc. R. Soc. Lond.* B **248**, 291–295.

Shadlen, M. N. & Movshon, J. A. 1999 Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* **24**, 67–77.

Singer, W. 1999a Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49–65.

Singer, W. 1999b Time as coding space? *Curr. Opin. Neurobiol.* **9**, 189–194.

Usher, M. & Donnelly, N. 1998 Visual synchrony affects binding and segmentation in perception. *Nature* **394**, 179–182.