Psychol Res (1990) 52: 22 – 27

Psychological Research

The effects of spatial phase on reaction time to spatially filtered images

James G. May[1, 2, 3], James M. Brown[1, 2], Carolina Gutierrez[1], and Mike Donlon[1]

[1] Department of Psychology, Univerity of New Orleans, Lakefront, New Orleans, LA 70148, USA
[2] Kresge Hearing Research Laboratory of the South, Lousiana State University Medical Center, 2020 Graver Street, New Orleans, LA 70112, USA
[3] LSU Eye Center, Louisiana State University Medical Center, 2020 Gravier Street, New Orleans, LA 70112, USA

Received August 14, 1989/Accepted February 17, 1990

Summary. Many studies of visual perception have used periodic stimuli such as sine-wave gratings and checkerboard patterns. It is well known that reaction time (RT) to such stimuli increases with increasing spatial frequency and decreasing contrast. While this is the case with periodic stimuli it is not clear that these relationships obtain for aperiodic stimuli such as natural scenes. A digitized image of an object (a vase) was submitted to two-dimensional Fourier analysis. Four pairs of spatial frequency band-limited images were created for each image. Each pair consisted of a normal-phase (NP) and a scrambled-phase (SP) version, with the magnitude spectrum and space-averaged luminance the same within each pair. Filter bandwidths were 1 octave wide. Manual RT was measured for onset and offset of each spatially filtered image. Mean RT for SP images increased significantly with increasing spatial frequency, while no other significant differences were found with the NP images. This suggests that the temporal processing of complex, aperiodic images is influenced by the spatial frequency and contrast of local regions within the image, rather than by the space-averaged contrast of the entire image, and cannot be predicted by global estimates of contrast and spatial frequency.

Simple manual reaction time (RT) has been used to assess the temporal processing of various visual stimuli. It has been shown to decrease with stimulus intensity as a simple monotonic function (Stebbins, 1966; Moody, 1970; Mansfield, 1973). RT has been shown to increase as a function of the visibility of the stimulus when sine-wave gratings were used as stimuli. Thus, RT to stimuli of high spatial frequency or low contrast is longer than RT to stimuli of low spatial frequency or high contrast (Lupp, Hauske, & Wolf, 1976; Vassilev & Mitov, 1976; Breitmeyer, 1975). Harwerth and Levi (1978) have shown that contrast-sensitivity curves can be derived from criterion measures of RT to sine-wave gratings of various spatial frequency and contrasts. Gish, Shulman, Sheehy, and Leibowitz (1986) have shown that RT declines with the detectability of sine-wave gratings and that low spatial-frequency gratings are more detectable than high-frequency gratings. They interpret this increased RT as due to an increased perceptual latency for high spatial-frequency stimuli. A study of perceived simultaneity (Parker & Dutch, 1987) indicated that visual probes were adjusted to imply that the onset of high spatial-frequency gratings is perceived later than that of low spatial-frequency stimuli. These studies suggest that the temporal processing of visual images depends upon the contrast and spatial content of the image. If this is strictly true, then any two images with equal contrast and spatial frequency would be expected to produce equal RTs. However, most complex images can also be characterized in terms of the meridional orientation and phase relationships of their sinusoidal components.

Studies of orientational differences with grating stimuli have shown that grating acuity (Higgins & Stultz, 1948; Leibowitz, 1953), contrast sensitivity (Campbell, Kulikowski, & Levinson, 1966; Camisa, Blake, & Lema, 1977), and temporal resolution (Foley, 1962; Schwartz, Winstead, & May, 1982) are higher for gratings oriented in the vertical and horizontal meridians as opposed to those in the oblique meridians. This suggests that two-dimensional patterns containig different power in these meridians might result in significant differences in RT. Studies of two-component sine-wave gratings (Campbell & Robson, 1964; Graham & Nachmias, 1971) indicate that the detection of complex gratings containing different phase relationships do not differ significantly, even if the net contrast is quite different (e.g. peaks-add vs. peaks-subtract phase). This suggests that two images equated for contrast, spatial frequency, and orientation will be detected with the same rapidity, despite any differences in phase relationships in the two images. That is not to say that the two images are metamers – that they are indiscriminable. For as Piotroski and Campbell (1982) have shown, it is the phase spectrum of an image that conveys most of its structure. Yet there are

---

Offprint requests to: J. G. May at his first address other studies that suggest that RTs may differ even when stimuli are equated for all, but their phase relationships.

A large body of evidence suggests that the more global aspects of a scene are processed more rapidly than the fine detail in the image. The prototypical paradigm for quantifying this phenomenon involves constructing stimuli such that local and global aspects can be judged independently. Navon (1977) constructed large letters (H and S), using small letters of the same optotype (H or S). The stimuli could be congruent (H made of Hs or S made of Ss) or incongruent (H made of Ss or S made of Hs), and the subjects were asked to indicate, as quickly as possible, what the large letters were under one condition or what the small letters were under another conditon. The results indicated that large letters were identified more rapidly than small, and that for small letter identification congruent conditions resulted in faster response times than incongruent conditions did. This phenomenon has been observed by many experimenters (Miller, 1981; Navon, 1981a,b; Navon & Norman, 1983; Pomerantz, 1983) and implies that the global aspects of a stimulus are processed more rapidly than the fine detail in the stimulus. Antes and Mann (1984) have recently reported global and local precedence effects with natural scenes. Global precedence was found for small display size, while local precedence was found with large display size. They conclude that the effects may depend on a critical spatial-frequency bandwidth.

Burr, Morrone, and Ross (1986) have shown that the global aspects of a checkerboard (its obliquely oriented fundamental components) are seen if the fifth harmonic is undetectable, but the local aspects (horizontally and vertically oriented check edges) are perceived if the fifth harmonic is visible. If this global-precedence effect obtained with schematic stimuli occurs with more natural scenes as well, then the more global aspects of images would be processed more rapidly than the local features defined by fine detail. If an image is described in terms of its Fourier components, then the hypothesis can be restated: that RT would be expected to be faster to the low spatial-frequency components of an image than it would be to the high spatial-frequency components of the image. Another reason for expecting this is that most images contain more power in the low-frequency components than in the high-frequency components. Thus, two lines of evidence suggest that the spatial-frequency content and contrast of an image determine the speed of processing an image. Previous RT studies have shown that RT is faster for low-frequency gratings, and global-precedence experiments have indicated that global aspects, which contain more low-spatial frequency content, are processed more rapidly than fine detail depicted in high spatial-frequency content. However, there are other reasons to think that images of objects might be processed more rapidly than images of non-objects. Global form may be inferred from the arrangement of detail without low-frequency spatial information. Fiorentini, Maffei, and Sandini (1983) have shown that faces are discriminated more readily when images are composed of high, as opposed to low, spatial-frequency information.

In the present study we asked whether RT measures differed when an object-like image and a non-object-like control image were used as stimuli. We created images with identical contrast spatial-frequency content and orientational arrangement, but different in structure, to determine whether it is the object-related, configural aspects of the image or the spatial-frequency aspects of the image that predict how it will be temporally processed. We examined the contribution of the spatial-frequency content of the images by using a range of band-passed spatial filters to create pairs of object-like and non-object-like images.

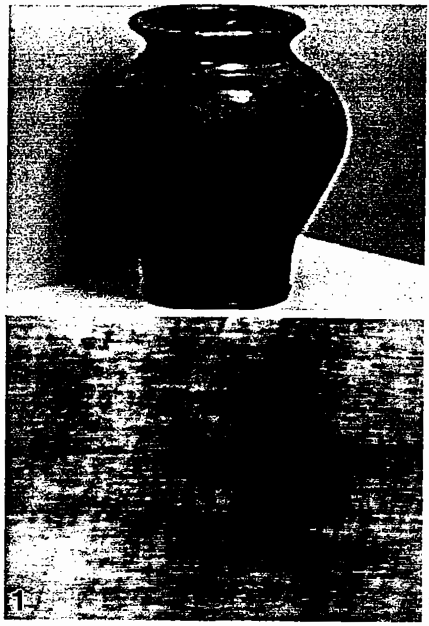
Experiment 1

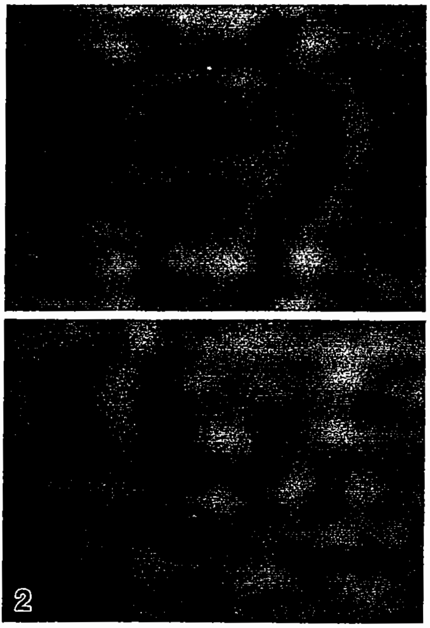
Method

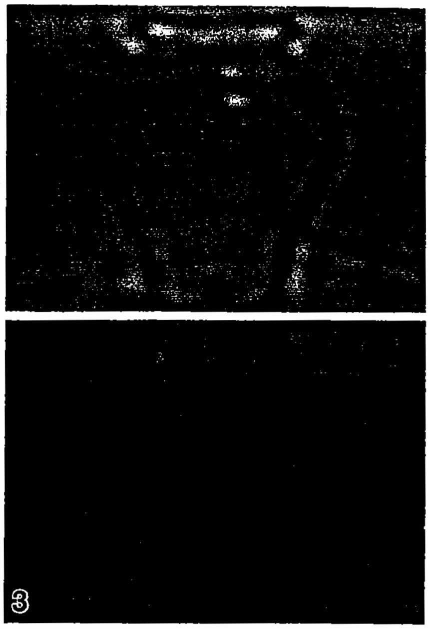
Subjects. Five observers participated in this experiment: three male and two female. Their ages ranged from 20 to 48 years. Two subjects were naive as to the hypotheses under investigation. All subjects had visual acuity of 20/20 with correction and were free from visual abnormalities.

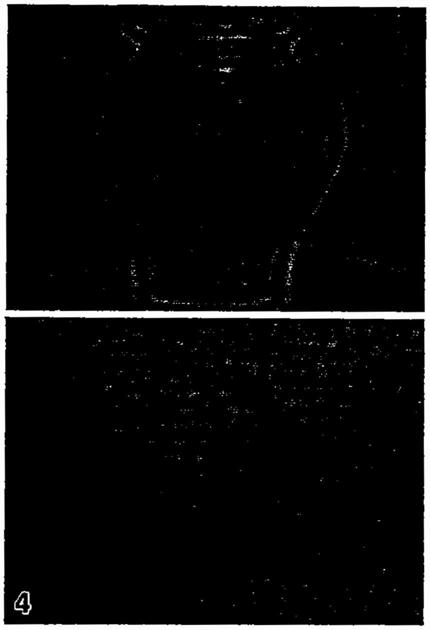
Stimuli. An image of a vase was digitized by means of a video camera (Panasonic, WV-F2) and an image processor (Data Translation frame grabber, Model 2851). With the aid of an array processor (Data Translation, Model 7020) the image was submitted to two-dimensional Fourier analysis. It was then filtered trough a rectangular filter to create four band-passed images with bandwidths of 4.0–8.0, 8.1–16.0, 16.1–32.0, and 32.1–64.0 cycles/image (c/i). A scrambled-phase (SP) version of each normal-phase (NP) filtered vase and of the full-spectrum vase was created by randomizing the imaginary components of the Fast Fourier Transforms (FFTs) of these images. Reverse FFTs of the NP and SP images produced five pairs of object-like and non-object-like images with identical Fourier magnitude spectra. All images were stored in a computer, and the image processor was able to access and display them on a video monitor (Tektronix Model 690SR). At a viewing distance of 170 cm the images subtended 5.11 (heigth) × 5.11 (width) degrees of visual angle (VA), the monitor subtended 9.35 (heigth) × 11 (width) degrees VA, and the spatial frequency content of the four pairs of images was 0.75–1.5, 1.5–3.0, 3.0–6.0, and 6.0–12.0 c/d. The images were presented centered on the monitor with the surround luminance equal to the mean luminance of the image. The stimulus duration was 1,667 ms and the interstimulus interval varied between 1,800 and 3,200 ms. For the onset RT an image was present during the stimulus period and a blank field of the same space-averaged luminance occurred during the interstimulus period. For the offset RT the image field occurred for durations varied between 1,800 and 3,200 ms and was replaced by the blank field for 1,667 ms.

Procedure. RT was measured, with 1-ms resolution, to the onset and offset of each presentation. Subjects were instructed to fixate upon the center of the screen and respond as quickly as possible to the onset or the offset of the entire image, RTs were accumulated in blocks of 10 for each image and all images were presented in each session. Each subject had one practice session for onsets and offsets, and then completed four sessions for onset RT and four ses-









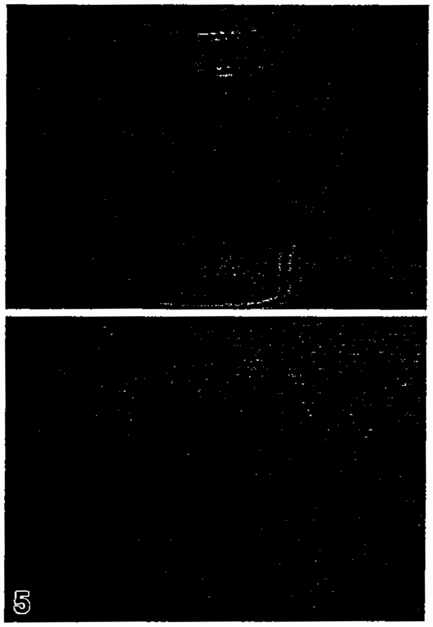


Fig. 1. The unfiltered, normal-phase image (top panel) and the unfiltered, scrambled-phase image (bottom panel) used in Experiment 1

Fig. 2. The low-passed filtered, normal-phase image (top panel) and the low-passed filtered, scrambled-phase image (bottom panel) used in Experiment 1

Fig. 3. The low-medium-passed filtered, normal-phase image (top panel) and the low-medium-passed filtered, scrambled-phase image (bottom panel) used in Experiment 1

Fig. 4. The medium-high-passed filtered, normal-phase image (top panel) and the medium-high-passed filtered, scrambled-phase image (bottom panel) used in Experiment 1

Fig. 5. The high-passed filtered, normal-phase image (top panel) and the high-passed filtered, scrambled-phase image (bottom panel) used in Experiment 1

sions for offset RT, yielding mean RTs based on 40 trials for each stimulus. The order in which images were presented was randomized.

Results

The mean RT for the onset of each image is presented in Figure 6 (top panel). The horizontal lines indicate the RT measurements for unfiltered normal and SP images. Increased RT is seen only with the high-pass SP image. Analysis of variance revealed significant main effects for filters, $F\ (4,16) = 23.06, p < .0001$, and image type, $F\ (1,4) = 27.05, p < .01$, and a significant interaction between these two factors, $F\ (4,16) = 4.76, p < .01$. Subsequent paired comparisons (Newman-Keuls) indicated that the mean RT for the high-passed vase was significantly increased in comparison with that of the nonfiltered vase ($p < .02$) and mean RT for the high-pass SP image was significantly increased in comparison with all other means ($p < .001$).

The mean RT for the offset of each image is presented in Figure 6 (bottom panel). The horizontal lines indicate RT for the unfiltered images. Increased RT is again only seen with the high-pass SP image. Analysis of variance revealed a significant main effect for filters, $F\ (4,16) = 9.63$, $p < .001$, and a significant interaction between filters and image type, $F\ (4,16) = 5.42, p < .01$. Subsequent paired comparisons indicated that the mean RT for the high-passed vase was significantly increased when compared with that of the nonfiltered vase ($p < .01$) and the mean RT for the high-passed SP image was significantly elevated in comparison with all other means ($p < .01$)

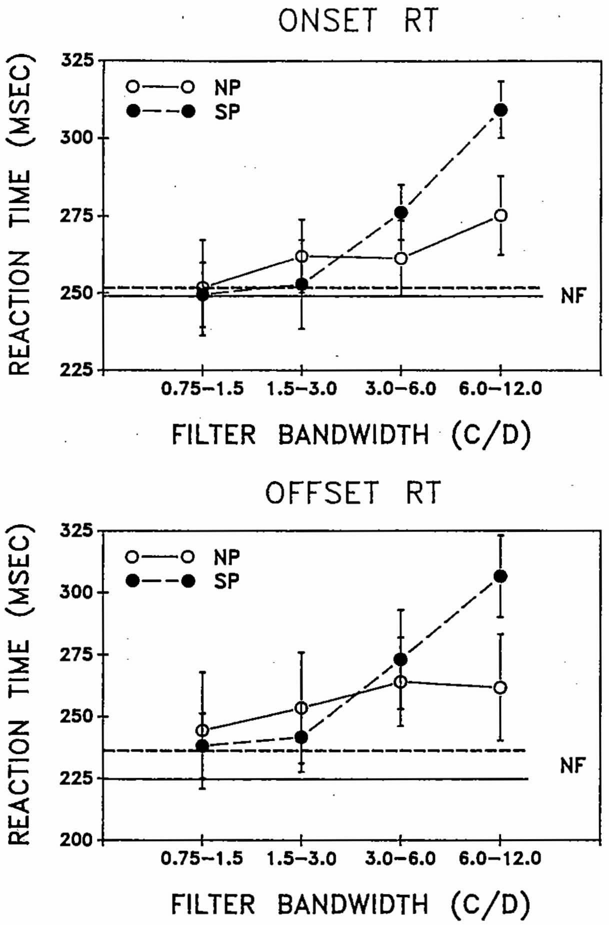


Fig. 6. Mean RT as a function of filter bandpass and image type for image onset (top panel) and image offset (bottom panel). Solid and dotted horizontal lines indicate the mean RT for the unfiltered images

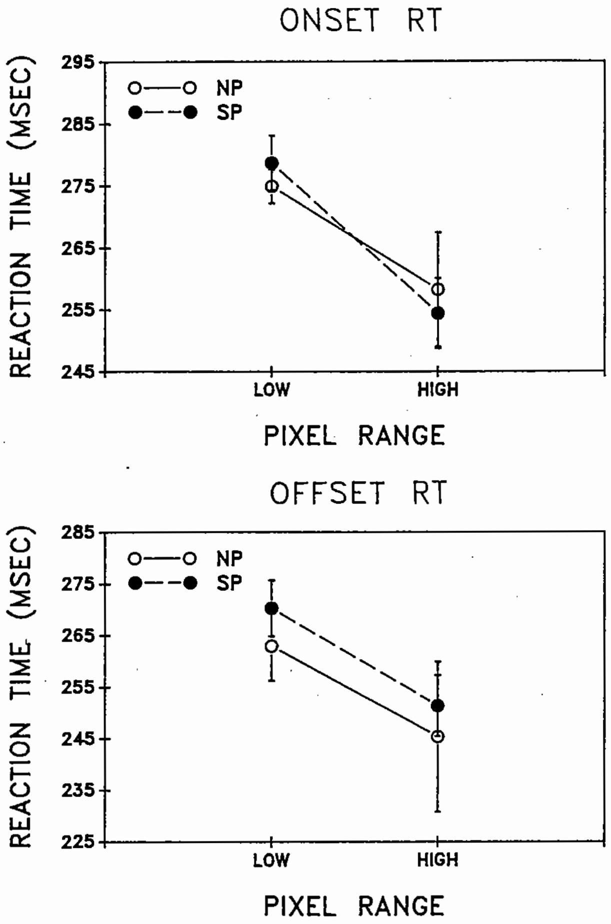


Fig. 7. Mean RT as a function of the range of pixel brightness and image type for image onset (top panel) and image offset (bottom panel)

Discussion

These results indicate that the RT to the three lowest-passed objects is the same for NP images and SP control stimuli and that the RT to these images does not differ from the RT to the unfiltered images. This suggests that the low spatial-frequency content of unfiltered images mediates the rapid RT to them, partly because these images contain more power than the high-passed images. More interestingly, the RT to the high-passed SP control images was significantly greater than the RT to the object-like image, despite the fact that both had the same space-averaged contrast and spatial-frequency content. The results of sine-wave grating studies predict that RT would increase with increasing spatial frequency and decreasing contrast. This was the case with the SP image, but not with the object-like NP image. Thus it appears that images with equal power, as defined by two-dimensional Fourier analysis, will produce equivalent RT measures if the spatial-frequency content is low, but result in faster RT for object-like images if the spatial-frequency content is high.

One version of the global-precedence point of view might suggest that images of objects usually contain a cluster of local features and that such features are defined mainly by fine spatial detail in the image. Furthermore, it is the unique arrangement of these features that conveys the information necessary to recognize the object (Fiorentini et al., 1983) and to infer its global form. Since the SP image is devoid of such structure, it is processed on the bases of spatial frequency and contrast alone. However, the NP image contained local features, especially obvious in the high-passed image. The arrangement of these features could be used to infer global form, resulting in a global-precedence effect, which facilitated the speed of processing these images.

Another view, however, is that the local features in the high-passed NP image contain more local contrast than any comparable local areas in the SP image. If RT is mediated by contrast of local areas as opposed to the average power in the whole image, then RT to the NP image would be expected to be less than that to the SP image. If this is the case, then it should be possible to adjust the contrast of the high-passed images to equate or to reverse the differences in RT. In Experiment 2 we constructed two more images of the high-passed stimuli and scaled them to reverse the differences in peak contrast.

Experiment 2

Method

The high-passed images used in Experiment 1 were analyzed to determine the range of pixel brightness in each image. The vase was found to have a greater range (61–196) than the SP image (102–161), although the means (M) and standard deviations (SD) of pixel brightness were identical ($M = 132.75$, $SD = 7.08$). Two new versions of each image were created. In one the range of pixel brightness for the new NP image was scaled to match that of the original SP image, and in the other the range of pixel values for the new SP image were scaled to have the range of the original NP image. These two new images were quite similar in average pixel brightness (NP, $M = 131.38$, $SD = 16.24$; SP, $M = 133.33$, $SD = 3.10$). The same five subjects that participated in Experiment 1 served as subjects in this experiment. Using the same methods as in Experiment 1, onset and offset RTs were measured for the two new and the two old high-passed images.

Results

The mean RTs for onset and offsets are presented in Figure 7. It is quite clear that RT to both onsets and offsets was faster for the image with the higher pixel brightness range, regardless of image type. This observation was supported by analyses of variance which revealed a significant main effect for pixel range for both onsets, $F(1,4) = 11.11$, $p < .05$ and offsets, $F(1,4) = 9.28$, $p < .05$. No other main effects or interactions achieved significance.

Discussion

The results of Experiment 2 indicate that RTs to the onset and the offset of the high-passed images were determined by the range of pixel brightness in the images. RT was not influenced by image type. This suggests that the results of Experiment 1 can be interpreted in favor of the local-contrast explanation. The reason that RT was shorter for the high-pass NP image was that the image contained local areas of higher contrast than any corresponding areas in the high-pass SP image. Although the procedure used to produce the SP images resulted in image pairs with the same space-averaged magnitude spectra, this technique created SP images devoid of local areas of high contrast which characterize high-pass NP images.

General Discussion

Previous experiments with gratings indicated that RT increased with spatial frequency and decreased with contrast. The present investigation sought to determine whether these relationships could be generalized to a more spatially complex natural scene. In Experiment 1 we found that filtered versions of a vase resulted in onset and offset RTs that did not differ significantly from those to the unfiltered image, even though the RT to non-vase-like comparison images increased as the spatial-frequency content of the stimulus increased. One explanation for this finding might be that the global form of the object was easily inferred from the high spatial-frequency detail in the high-pass NP image. Thus some sort of global precedence effect might explain these results. There are several reasons for rejecting such an explanation, however. First, the global-precedence effect is observed in situations in which both low- and high-frequency components are present in the stimulus and subjects are asked to make a choice reaction time concerning one aspect while ignoring another. In our experiment only some spatial components were present and simple reaction time was required. If the global-precedence effect derives from a competition between mechanisms that process local and global features, then one would not expect such an effect with a paradigm such as ours, since no choice was required and only local or global information was present. Second, the global-precedence effect has been shown to depend on the existence of low spatial-frequency content and is not obtained with high-passed stimuli (Badcock, Whitwort, Badcock & Lovegrove, 1989). Third, the results of Experiment 2 suggest that the decreased RT to the high-passed NP image is mediated by the local areas of high contrast in that image, and not by the total contrast of the scene.

These findings suggest caution in extrapolating from previous research with simple grating stimuli. While it is possible to characterize any visual scene in Fourier-analytic terms, one assumption of that approach is that space-averaged contrast in the image is distributed evenly throughout the image. With most complex natural images, that is certainly not the case. Rather, local areas of high contrast constitute the distinctive features of the object. This is especially true for features defined by high-spatial frequencies (Fiorentini et al., 1983). From the present investigation we must conclude that the temporal processing of complex scenes does not depend on the space-averaged contrast of the total scene, but is determined by the contrast of local features within the scene. In addition, it is clear that, while speed of processing differs for sinusoidal components, this simple rule does not provide a unitary construct that can be employed to understand the global-precedence effect and the simple detection of a natural object.

Acknowledgement. This research was supported by a grant to the first author from the Louisiana Board of Regents through the Louisiana Education Quality Support Fund.

References

Antes, J. R., & Mann, S. W. (1984). Global-local precedence in picture processing. *Psychological Research, 46*, 247–259.

Badcock, J. C., Whitworth, F. A., Badcock, D. R., & Lovegrove, W. J. (1989). Low frequency filtering and the processing of local-global stimuli. Submitted to *Perception*.

Burr, D. C., Morrone, M. C., & Ross, J. (1986). Local and global visual processing. *Vision Research, 26*, 749–757.

Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal properties of sustained and transient channels. *Vision Research, 15*, 1411–1412.

Camisa, J. M., Blake, R., & Lema, S. (1977). The effects of temporal modulation on the oblique effect in humans. *Perception, 6,* 165–171.

Campbell, F. W., & Robson, J. G. (1964). Application of Fourier analysis to the modulation response of the eye. *Journal of the Optical Society of America, 54,* 581.

Campbell, F. W., Kulikowski, J. J., & Levinson, J. (1966). The effect of orientation on the visual resolution of a grating. *Journal of Physiology (London), 187,* 427–436.

Fiorentini, A., Maffei, L., Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception, 12,* 195–201.

Foley, P. J. (1962). Stimulus orientation and retinal summation. *Journal of the Optical Society of America, 52,* 474–475.

Gish, K., Shulman, G. L., Sheehy, J. B., & Leibowitz, H. W. (1986). Reaction times to different spatial frequencies as a function of detectability. *Vision Research, 26,* 745–747.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11,* 251–259.

Harwerth, R., & Levi, D. (1978). Reaction time as a measure of suprathreshold grating detection. *Vision Research, 18,* 1579–1586.

Higgins, G. C., & Stultz, K. (1948). Visual acuity as measured with various orientations of a parallel-line test object. *Journal of the Optical Society of America, 38,* 756–758.

Leibowitz, H. (1953). Some observations and theory on the variation of visual acuity with the orientation of the test object. *Journal of the Optical Society of America, 43,* 902–905.

Lupp, U., Hauske, G., & Wolf, W. (1976). Perceptual latencies to sinusoidal gratings. *Vision Research, 16,* 969–972.

Mansfield, R. J. W. (1973). Latency functions in human vision. *Vision Research, 13,* 2219–2234.

Miller, J. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1161–1174.

Moody, D. B. (1970). Reaction time as an index of sensory function. In: W. C. Stebbins (Ed.), *Animal psychophysics.* New York: Appleton-Century-Crofts.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Navon, D. (1981a). The forest revisited: More on global precedence. *Psychological Research, 43,* 1–32.

Navon, D. (1981b). Do attention and decision follow perception? Comment on Miller. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1175–1182.

Navon, D., & Norman, J. (1983). Does global precedence really depend on visual angle? *Journal of Experimental Psychology: Human Perception and Performance, 9,* 955–965.

Parker, D. M., & Dutch, S. (1987). Perceptual latency and spatial frequency. *Vision Research, 27,* 1279–1283.

Piotroski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency, amplitude, and phase. *Perception, 11,* 337–346.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General, 112,* 516–540.

Schwartz, B. D., Winstead, D. K., & May, J. G. (1982). Meridional differences in temporal resolution. *Perception, 11,* 25–34.

Stebbins, W. C. (1966). Auditory reaction time and the derivation of equal loudness contours for the monkey. *Journal of the Experimental Analysis of Behavior, 9,* 135–142.

Vassilev, A., & Mitov, D. (1976). Perception time and spatial frequency. *Vision Research, 16,* 89–92.