Perception & Psychophysics
1990, 47 (6), 563-567

Visual persistence of spatially filtered images

JAMES G. MAY and JAMES M. BROWN
University of New Orleans, Lakefront, Louisiana
and Louisiana State University Medical Center, New Orleans, Louisiana

and

SHERRYL SCOTT and MICHAEL DONLON
University of New Orleans, Lakefront, Louisiana

Periodic stimuli such as sine-wave gratings and checkerboard patterns have been used in many studies of visual perception. It is well known that with such stimuli, visual persistence increases as spatial frequency increases and as contrast decreases. It is not clear, however, that similar relationships obtain for aperiodic stimuli such as natural images. Digitized images of objects (a face and a vase) were 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 (NP) and a scrambled (SP) phase version, with the magnitude spectrum and space-averaged luminance the same within each pair. Filter bandwidths were one octave wide. Threshold persistence was measured for each spatially filtered image. Visual persistence for SP images increased significantly as spatial frequency increased, whereas no significant differences were found for NP images. This suggests that the temporal processing of complex, aperiodic visual images is influenced by the spatial frequency and contrast of local features within the image and cannot be predicted by space-averaged estimates of contrast and spatial frequency.

Visual persistence is the neural perseveration of a visual stimulus beyond its physical termination (Bowen, Pola, & Martin, 1974). Using gratings and checkerboards as stimuli, persistence has been shown to increase as a function of the visibility of the stimulus. Thus, stimuli of high spatial frequency or low contrast persist longer than those of low spatial frequency or high contrast (Badcock & Lovegrove, 1981; Bowling & Lovegrove, 1980, 1981; Bowling, Lovegrove, & Mapperson, 1979; Marx & May, 1983; Meyer & Maguire, 1977, 1979, 1981; Schwartz, Winstead, & May, 1982). This suggests that the temporal processing of visual images depends on the contrast and spatial frequency content of the image. If this were strictly true, any two images with equal contrast and spatial frequency content would be expected to produce equal persistence. There are some reasons, however, to believe that other factors in addition to spatial frequency and contrast may underlie visual persistence.

A number of reports have indicated that targets embedded in a meaningful context are processed more accurately than targets embedded in less meaningful contexts. These phenomena constitute a class of superiority effects which suggest that a simple description of the contrast and spatial frequency content of a stimulus cannot predict how the visual system will process the image. Numerous experiments (Weisstein, M. C. Williams, & Harris, 1979; A. Williams & Weisstein, 1978; M. C. Williams & Weisstein, 1981) have indicated that targets viewed in the context of a three-dimensional figure are detected more quickly than targets viewed in a two-dimensional pattern of equivalent spatial frequency content and contrast. This phenomenon has been called the object-superiority effect. An equally impressive set of studies (Carr, Lehmkuhle, Kottas, Astor-Stetson, & Arnold, 1976; Johnson, 1978; Krueger & Shapiro, 1979; McClelland, 1976; Solman, May, & Schwartz, 1981; Spector & Purcell, 1977; Staller & Lappin, 1979) has indicated that letters are detected more accurately when they are parts of words than when they are parts of nonwords. This is termed the word-superiority effect. Finally, in some studies (Homa, Haver, & Schwartz, 1976; Purcell & Stewart, 1986, 1988; van Santen & Jonides, 1978), it has been demonstrated that facial features are more accurately detected when they are part of a face than when they are part of a nonface stimulus. This has been described as the face-superiority effect.

In the present study, we asked whether persistence differed when an object-like image and a non-object-like image were used as targets. We created images with identical space-averaged contrast and spatial frequency content, but with different structures, to determine whether the spatial aspects of an image or the object-like aspects of an image predict how that image will be temporally processed. In addition, we examined the contribution of the spatial frequency content of the images by using a range of band-passed spatial filters to create the paired images.

This research was supported by the Louisiana Board of Regents through the Louisiana Educational Support Fund, with a grant to James G. May.
Correspondence may be addressed to James G. May, Department of Psychology, University of New Orleans, New Orleans, LA 70148.

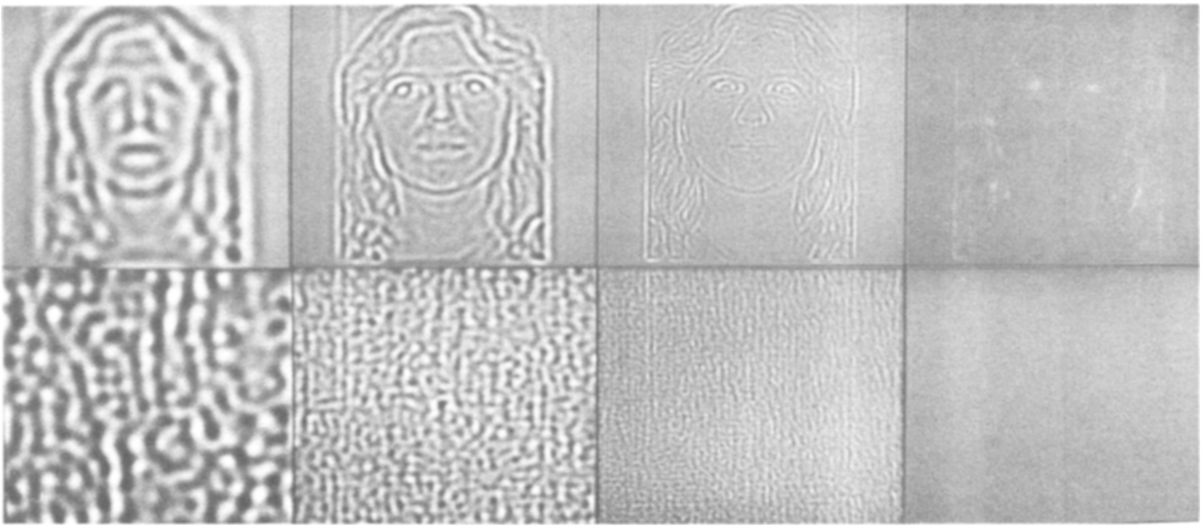

Figure 1. Examples of the filtered images used in Experiment 1.

EXPERIMENT 1

Method

Subjects. Six observers participated as subjects. All were naive as to the hypotheses under investigation. All had corrected 20/20 visual acuity and were free from other visual abnormalities. They ranged in age from 19 to 33 years.

Stimuli. An image of a female face was digitized and submitted to two-dimensional Fourier analysis. The image was then filtered with a Gaussian envelope to create four normal-phase (NP) images with bandwidths of 8.0-16.0, 16.0-32.0, 32.0-64.0, and 64.0-128.0 cycles/image (cpi). A scrambled phase (SP) version of each NP image was created by randomizing the imaginary components of the FFTs. Reverse FFTs were displayed and photographed to produce slides (2 in. × 2 in.). The slides were presented tachistoscopically (Scientific-Prototype Model N-1100) such that the spatial frequency content of each pair of images was 1.5-3.0, 3.0-6.0, 6.0-12.0, or 12.0-24.0 cycles/degree (cpd). The stimuli used in this experiment are presented in Figure 1.

Procedure. Each trial began with the observer fixating a point centered on a blank field. When the fixation point was turned off, an image was then alternately presented with a blank field, 12 times in rapid succession. Image duration was fixed at 100 msec, while blank frame duration or interstimulus interval (ISI) could vary from 5-995 msec in 5-msec steps. Image and blank field luminances were always equated. Persistence was measured, using blockwise tracking to determine the ISI necessary for perception of the image as having temporal continuity on 75% of the trials. Order of presentation of image pairs was randomized across observers. The NP image of a pair was always tested first, followed by the SP version.

Results

The mean ISI for each filtered image is presented in Figure 2. It is apparent that persistence for the SP control images increased as the spatial frequency of the image increased, while the persistence of the NP images remained constant. Analysis of variance revealed significant main effects for filters [$F(3,9) = 17.79, p < .0004$] and image type [$F(1,3) = 19.29, p < .0219$], and a significant interaction between image type and filters [$F(3,9) = 4.66, p < .0314$]. Subsequent paired comparisons (Newman-Keuls) indicated a significant difference in persistence between the high-pass SP image and all other conditions ($p < .0003$).

Discussion

The results of this experiment indicate that temporal processing for filtered faces is significantly different from that for a nonface control stimulus containing the same Fourier magnitude spectra. The face stimuli did not differ in persistence, whereas persistence increased for the SP versions as the spatial frequency content increased. To determine if this result was reliable and not due to a subtle artifact introduced by the photographic process, we carried out a second experiment with images presented on a display monitor, using an image-processing system.

EXPERIMENT 2

Method

Subjects. Five subjects participated in this experiment. Two had participated in Experiment 1. One new subject was recruited and was naive as to the hypotheses under investigation. Two of the authors also participated as observers. All subjects had corrected 20/20 visual acuity and were free of other visual abnormalities. They ranged in age from 19 to 48 years.

Stimuli. An image of a vase was digitized, using a video camera (Panasonic, WV-F2) and an image processor (Data Translation frame grabber, Model 2851). Using an array processor (Data Translation, Model 7020), the image was submitted to two-dimensional Fourier analysis. It was then filtered with 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 cpi. An SP version of each NP filtered vase and the full-spectrum, unfiltered vase was created by randomizing the imaginary components of the FFTs. Reverse FFTs on 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 on a computer so that the image processor could access them and then display them on a video monitor (Tek-

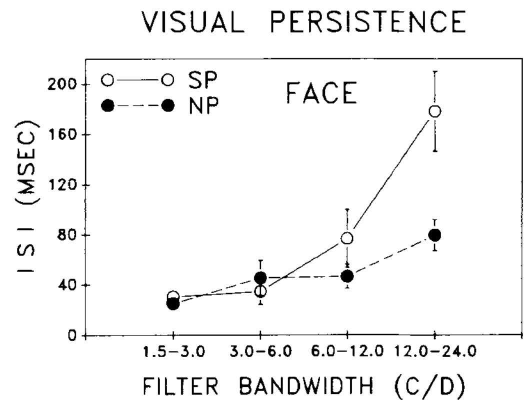

Figure 2. Mean persistence as a function of filter bandwidth and image type. NP = normal phase; SP = scrambled phase.

tronix, Model 690SR). At a viewing distance of 170 cm, the images subtended 5.11° × 5.11° visual angle (VA), the monitor subtended 9.35° × 11° VA (height × width), 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 cpd. The images were presented centered on the monitor, with the surround luminance equal to the mean luminance of the image. During the experiment, the images were alternated with a blank field with luminance equal to the mean luminance of the image.

Procedure. As in Experiment 1, an image was rapidly alternated 12 times with a mean luminance blank field on each trial. Image duration was fixed at 96 msec. Blank field duration range was 16-320 msec, with a 16-msec step size. Subjects made a forced choice response on each trial, whether the image was continuous in time, or interrupted by a press of the appropriate key on a computer terminal keyboard. The computer recorded the responses and controlled stimulus presentation. Before an experimental block was started, an obvious example of continuity (ISI = 16 msec) and interruption (ISI = 320 msec) was demonstrated twice, using the full-spectrum vase. A block of trials consisted of each NP and SP image being randomly presented once. Persistence was measured, using a PEST procedure (Taylor & Creelman, 1967). Data were averaged over two blocks.

Results

The mean ISI for each image is presented in Figure 3. The horizontal lines indicate the persistence measurements for unfiltered NP and SP images. As in Experiment 1, increased persistence is seen only with the high-pass SP image. Analysis of variance revealed significant main effects for filters [$F(4,12) = 19.85$, $p < .0001$] and image type [$F(1,3) = 51.49$, $p < .0056$], and a significant interaction between these two factors [$F(4,12) = 14.09$, $p < .0002$]. Subsequent paired comparisons indicated that the persistence was significantly increased only for the high-pass SP image ($p < .0001$).

Discussion

The results of Experiment 2 agree quite well with those of Experiment 1 and indicate that the persistence for non-object-like SP images increases with spatial frequency content, whereas persistence for object-like NP stimuli does not. Prior to attributing this difference to an object-superiority effect, we conducted a third experiment to determine whether or not this effect could be accounted for by differences in the contrast of a few local areas in the NP and SP images. Only the high-passed images were employed, since the major effect occurred with these images.

EXPERIMENT 3

Method

Subjects. Five observers served as subjects in this experiment. Three had participated in Experiment 2. Three were naive as to the hypotheses under investigation, and all had corrected 20/20 visual acuity. They ranged in age from 19 to 48 years.

Stimuli. The pixel gray-level range was evaluated for the high-pass (32.1-64.0 cpi) NP and SP vase. The pixel gray-level range was found to be greater for the NP (R = 61-196) vase compared to the SP (R = 102-161) vase, although the means (M) and standard deviations (SD) were identical (M = 132.75, SD = 7.08). A new high-pass NP vase was created, using the pixel gray-level range from the original SP vase, and a new SP vase was created, using the pixel gray-level range from the original NP vase. The two new stimuli were quite similar (NP, M = 131.38, SD = 16.24; SP, M = 133.33, SD = 3.10). Thus, the four stimuli used in this experiment were: an SP vase (the original with low pixel gray-level range), a new SP vase (with high pixel gray-level range), an NP vase (the original with high pixel gray-level range), and a new NP vase (with low pixel gray-level range). The two new versions were still identical in spatial frequency content, but the SP version had greater average contrast than the NP version did. A clear outline of the vase was still easily discerned in the new NP version. Except for the use of only four stimuli, all other viewing conditions and procedures were the same as in Experiment 2.

Results

The mean ISI as a function of pixel gray-level range is presented in Figure 4 for each high-pass image. The most persistence is obtained for images of highest gray-level range. Analysis of variance revealed no significant main effects, but the interaction between gray-level range and image type was significant [$F(1,4) = 11.36$, $p < .028$]. Subsequent tests (Newman-Keuls) revealed no significant differences between image types, but the

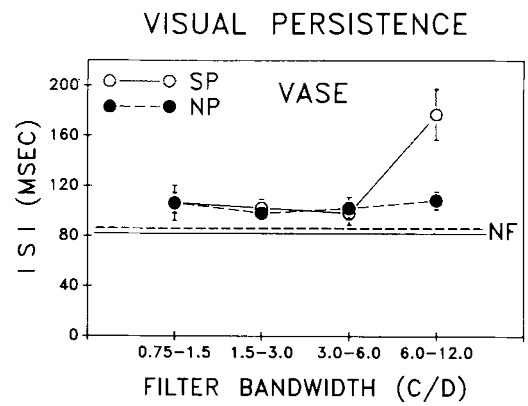

Figure 3. Mean persistence as a function of filter bandwidth and image type. The horizontal lines indicate the mean persistence for the nonfiltered (NF) images.

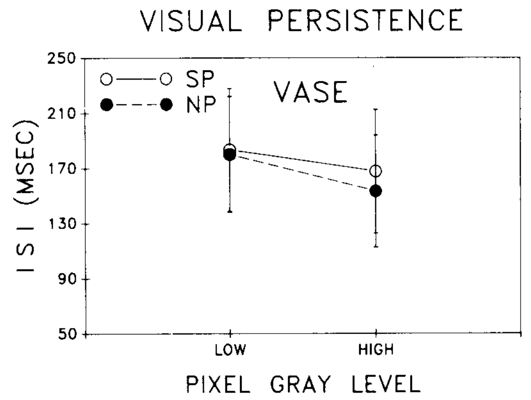

Figure 4. Mean persistence as a function of pixel gray-level range and image type.

two highest means were significantly different from the lowest with Duncan's multiple range test (*ps* < .03 and .04).

Discussion

Since the original high-pass NP vase image had a greater gray-level range due to the high contrast of a few local features in the image, adjusting the gray-level range of the new SP image was a conservative arrangement that resulted in greater net contrast across the entire image relative to the original NP vase image. Conversely, the new NP vase image contained less net contrast than did the original SP image, but the structure of the vase was still quite evident. The fact that this manipulation reversed the trends in the persistence of these two types of images supports the notion that the results of Experiment 2, and probably also Experiment 1, can be understood in terms of the concentration of higher contrast in the area of a few local features of the high-pass NP images.

GENERAL DISCUSSION

The results of Experiments 1 and 2 agree and indicate that the persistence of the three lowest passed images is the same for objects and non-object-like SP control stimuli. Experiment 2 showed that persistence of these images does not differ from the persistence of unfiltered, full-spectrum images. This suggests that the low spatial frequency content of unfiltered images mediates the persistence to them, due in part to the fact that these images contain more power than the high-passed images do. More interestingly, in both Experiments 1 and 2, the persistence of the high-passed, non-object-like SP control images, which had the same contrast and spatial frequency content as did the object-like NP high-passed images, was significantly greater than persistence of that image. The results of sine-wave grating studies would predict that persistence would increase as spatial frequency increases and contrast decreases. This was the case with the SP image, but not the object-like image. Thus, it appears that a two-dimensional Fourier description of the magnitude of images predicts persistence measures for low spatially filtered images, but fails for high-passed images.

Is this a form of object superiority or is there a more parsimonious explanation? One version of the object-superiority point of view might suggest that images of objects usually contain a cluster of local features, and that such features are defined mainly by high spatial detail in the image. Furthermore, it is the unique arrangement of these features that conveys the information necessary for one to recognize the object (Fiorentini, Maffei, & Sandini, 1983) and to infer its three-dimensional form. Since the SP image is devoid of such structure, it is processed on the basis of spatial frequency and contrast alone. Another view, however, is that the local features in the high-passed NP image contain more local contrast than does any comparable local area in the SP image. If persistence is mediated by contrast of local areas, as opposed to the average contrast in the whole image, then persistence of the NP image would be expected to be less than that of 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 reverse the differences in persistence. In Experiment 3, we constructed two more images of the high-passed stimuli and scaled them to reverse the differences in peak contrast. The results of Experiment 3 clearly indicate that the persistence differences observed in Experiments 1 and 2 derive from differences in local contrast between the normal and SP high-pass images.

The present findings suggest that one should be cautious in using earlier results with sinusoidal gratings to predict how complex images might be processed temporally. In those studies, the spatial frequency and space-averaged contrast were found to be prime determinants of visual persistence. With complex scenes, however, local features, and not just the average scene content, must be considered separately. It is, perhaps, the spatial frequency and contrast of these local areas which underlie object superiority effects. Although it is not clear how three-dimensional objects might differ systematically in phase spectra or local feature composition, such comparisons seem to be indicated.

REFERENCES

Badcock, D., & Lovegrove, W. (1981). The effects of contrast, stimulus duration, and spatial frequency on visible persistence in normal and specifically disabled readers. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 495-505.

Bowen, R. W., Pola, J., & Martin, L. (1974). Visual persistence: Effects of flash luminance, duration, and energy. *Vision Research*, **14**, 295-303.

Bowling, A., & Lovegrove, W. (1980). The effect of stimulus duration on the persistence of gratings. *Perception & Psychophysics*, **27**, 574-578.

Bowling, A., & Lovegrove, W. (1981). Two components to visible persistence: Effects of orientation and contrast. *Vision Research*, **21**, 1241-1251.

Bowling, A., Lovegrove, W., & Mapperson, B. (1979). The effects of spatial frequency and contrast on visual persistence. *Perception*, **8**, 529-539.

Carr, T. H., Lehmkuhle, S. W., Kottas, B., Astor-Stetson, E. C., & Arnold, D. (1976). Target position and practice in the identification of letters in varying contexts: A word superiority effect. *Perception & Psychophysics*, **19**, 412-416.

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

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory & Cognition*, **4**, 176-185.

Johnson, J. C. (1978). A test of the sophisticated casing theory of word perception. *Cognitive Psychology*, **10**, 123-153.

Krueger, L. E., & Shapiro, R. G. (1979). Letter detection with rapid serial visual presentation: Evidence against word superiority at feature extraction. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 657-673.

Marx, M. S., & May, J. G. (1983). The relationship between temporal integration and persistence. *Vision Research*, **23**, 1101-1106.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 80-91.

Meyer, G. E., & Maguire, W. M. (1977). Spatial frequency and the mediation of short-term visual storage. *Science*, **198**, 524-525.

Meyer, G. E., & Maguire, W. M. (1979). The effects of bar width and spatial frequency-specific adaptation on visual persistence. *Bulletin of the Psychonomic Society*, **14**, 64-66.

Meyer, G. E., & Maguire, W. M. (1981). Effects of spatial frequency specific adaptation and target duration on visual persistence. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 151-156.

Purcell, D. G., & Stewart, A. L. (1986). The face-detection effect. *Bulletin of the Psychonomic Society*, **24**, 118-120.

Purcell, D. G., & Stewart, A. L. (1988). The face-detection effect: Configuration enhances detection. *Perception & Psychophysics*, **43**, 355-366.

Schwartz, B., Winstead, D. K., & May, J. G. (1982). Meridional differences in temporal resolution. *Perception*, **11**, 25-35.

Solman, R. T., May, J. G., & Schwartz, B. D. (1981). The word superiority effect: A study using parts of letters. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 552-559.

Spector, A., & Purcell, D. G. (1977). The word superiority effect: A comparison between restricted and unrestricted alternative set. *Perception & Psychophysics*, **21**, 323-328.

Staller, J. D., & Lappin, J. S. (1979). Word and nonword superiority effects in a letter detection task. *Perception & Psychophysics*, **25**, 47-54.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, **41**, 782-787.

van Santen, J. P. H., & Jonides, J. (1978). A replication of the face-superiority effect. *Bulletin of the Psychonomic Society*, **12**, 378-380.

Weisstein, N., Williams, M. C., & Harris, C. S. (1979). Line segments are harder to see in flatter patterns: The role of three-dimensionality in "object-line" and "object-superiority" effects. *Bulletin of the Psychonomic Society*, **14**, 229-230. (Abstract)

Williams, A., & Weisstein, N. (1978). Line segments are perceived better in a coherent context than alone: An object-line effect in visual perception. *Memory & Cognition*, **6**, 85-90.

Williams, M. C., & Weisstein, N. (1981). Spatial frequency response and perceived depth in the time-course of object-superiority. *Vision Research*, **24**, 1279-1288.

(Manuscript received August 17, 1989;
accepted for publication December 28, 1989.)