NASA Technical Memorandum 102739

A DISCREPANCY WITHIN PRIMATE SPATIAL VISION AND ITS BEARING ON THE DEFINITION OF EDGE DETECTION PROCESSES IN MACHINE VISION

DANIEL J. JOBSON

(NASA-TM-102739) A DISCREPANCY WITHIN PRIMATE SPATIAL VISION AND ITS BEARING ON THE DEFINITION OF EDGE DETECTION PROCESSES IN MACHINE VISION (NASA) 31 p CSCL 09B

N91-13941

Unclas

G3/63 0319227

SEPTEMBER 1990

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

A DISCREPANCY WITHIN PRIMATE SPATIAL VISION AND ITS BEARING ON THE DEFINITION OF EDGE DETECTION PROCESSES IN MACHINE VISION

by Daniel J. Jobson

Abstract

The visual perception of form information is considered to be based on the functioning of simple and complex neurons in the primate striate cortex. However, a review of the physiological data on these brain cells cannot be harmonized with either the perceptual spatial frequency performance of primates or the performance which is necessary for form perception in humans. This discrepancy together with recent interest in cortical-like and perceptual-like processing in image coding and machine vision prompted a series of image processing experiments intended to provide some definition of the selection of image operators. The experiments were aimed at determining operators which could be used to detect edges in a computational manner consistent with the visual perception of structure in images. Fundamental issues were the selection of size and circular versus oriented operators (or some combination). In a previous study, circular difference-of-Gaussian (DOG) operators, with peak spatial frequency responses at about 11 and 33 cyc/deg, were found to capture the primary structural information in images. Here larger scale circular DOG operators were explored and led to severe loss of image structure and introduced spatial dislocations in structure which is not consistent with visual perception. Orientation sensitive operators (akin to one class of simple cortical neurons) introduced ambiguities of edge extent regardless of the scale of the operator. For machine vision schemes which are functionally similar to natural vision form perception, two circularly symmetric very high spatial frequency channels appear to be necessary and sufficient for a wide range of natural images. Such a machine vision scheme is most similar to the physiological performance of the primate lateral geniculate nucleus rather than the striate cortex.

Introduction

Vision is central to most human activity and the visual perception of form is the most central of visual skills. There is an interest in the development of a machine vision system which is functionally similar to natural vision. This system would have to include image acquisition, extraction of form information, "visual learning and memory", and recognition and interpretation. This paper is concerned only with the first two elements in this system and the use of natural vision concepts to influence the design of this subsystem.

The perception of form is widely considered to be based in the functioning of the simple and complex neurons of the striate cortex. For the design of a machine vision system, the spatial responses of these cells are of considerable interest, both for image coding and information extraction. Here a review of physiological data on cortical neurons and perceptual data provides a starting point for examining fundamental issues of the type(s) and size(s) of edge detection operators that are necessary and sufficient for extracting form information from arbitrary images. Equivalently for linear systems this amounts to a definition of the spatial frequency channels necessary for edge detection performance that is consistent with the visual perception of primary structure in images. This can be thought of as a question of what spatial frequency channels carry the most important form information. Here the selection of edge detection operators is limited to those occurring in early primate spatial vision and specific configurations which produce zero crossings and, more specifically, only one zero crossing per edge event. Within these limitations, a variety of edge detection operators are applied to determine consistency with visual perception.

Physiological and Perceptual Spatial Frequency Performance of Foveal Primate Vision

A review of primate physiological and perceptual data was conducted in order to provide insight into natural vision mechanisms of form perception. A significant discrepancy emerged when the collective spatial frequency response of foveal striate cortical neurons (Ref. 1) is compared to the perceptual spatial frequency response (Ref. 2) of the same primate (Fig. 1). The perceptual curve is also representative of the human response to luminance sine wave gratings (Ref. 1). With respect to form perception, this discrepancy is particularly disturbing because of the rather coarse spatial resolution of the bulk of the cortical units. The degree of the discrepancy at high spatial frequencies was masked by the original data being plotted on log-log scales (a convention in natural vision science) which heavily compress the high spatial frequency portion of the scale. Further cortical data was assembled (Fig. 2 and Ref. 3-7) in hopes of finding higher spatial frequency units. This did not prove to be the case. In terms of spatial resolution and coverage of the retinal image, high spatial frequency units should not only be present but also occur in populations that should increase as the square of spatial frequency. This persistent discrepancy led to a retreat backward in the visual pathway to examine the physiology of the lateral geniculate nucleus (Fig. 3 and Ref. 8). Here the high spatial frequency units are indeed present. This encourages the belief that any omission of high acuity channels in the striate cortex data is not due to experimental measurement limitations. However, there is still a paucity of numbers of units peaking at a spatial frequency near the limit set by the retinal mosaic of photoreceptors ($\approx$ 33 cyc/deg). Subsequently, I will illustrate that this channel at 33 cyc/deg does seem to be necessary for human form perception. For now it is sufficient to state that the thinness or complete absence of high spatial frequency units raises serious questions about how to use physiology as a source for defining machine vision methods aimed at computational form perception. The edge detection operations necessary for this are now explored via image processing experiments which will shift through low spatial frequencies to high and examine the issue of one circular operator per scale versus multiple orientation sensitive operators per scale.

Basic Spatial Responses of Primate Foveal Vision for Use in Zero Crossing Edge Detection

Since neither the spatial layout of primate receptive fields nor other fundamental aspects of edge detection in primates are known, the liberty must be taken of "playing" with what is known in image processing experiments. The intent here is not to perform an exact simulation of natural vision but rather to gain insight into computational methods which can produce a result consistent with visual perception. In no way can this definition of methods be considered to be a scientific model for natural vision processes because primate vision could use an entirely different means of achieving the same end. We can only attempt to start out in some reasonable "physiological" manner and end up with some reasonable "perceptual" result.

The perception of visual structure is predominantly accurate locational determinations of edge boundaries. Locational accuracy is paramount because such high acuity tasks as reading require locational accuracies approximately equivalent to the original image sampling grid of the photoreceptor mosaic (≈ 0.015 degree sample cells). A direct and computationally concise method of detecting edge locations is a zero crossing determination (Ref. 9). Not all of the spatial responses found in early primate vision are useful for this. Retinal receptive fields do not eliminate the zero spatial frequency response and therefore don't produce exact zero crossings locations that are independent of the actual image intensity values. Complex cortical neuron responses are apparently highly multilobed and introduce a correspondence problem, i.e. multiple zero crossings for one edge event. At the other extreme, one class of simple cortical neurons, one ridge and one valley, produces no zero crossing at all. This process of elimination narrows the types of spatial responses to the two shown (Fig. 4). Details of the specifics of edge detection methods are provided in Ref. 10 and 11. For visual comparisons throughout this paper, image size on the printed page is selected to make the size of each image element equal to the visual acuity limit (≈ 0.015 degrees) for a comfortable reading distance of about 20 inches. This in effect spatially calibrates visual perception to computational processes. Contrast rendition of these images as published is not particularly accurate; however, most of the perceptual content of the images is retained and most of the defects in edge detection are sufficiently glaring that highly accurate contrast rendition is unnecessary.

Edge Detection at 3 Cycles/Degree

The spatial frequency range centered on 3 cycles/degree is the collective operating range of the bulk of measured cortical neurons considered to be responsible for form perception and is therefore of initial interest in comparing edge detection experiments with visual perception.

A. Circular DOG Operator

With the calibration of computation to perception of one image element corresponding to 0.015 degrees (≈ 33 cycles/deg), a DOG operator at 3 cycles per degree has a center diameter of 11 image elements. The visual effect of this amount of blur is illustrated by convolving an original image with a Gaussian blur function whose circular full width at half of maximum value is 11 image elements (Fig. 5). Examples of edge representations for the 3 cycle/degree DOG operator (Fig. 6) consistently show little or no form information left after this amount of blurring. Cases where an impression of some form information is given, suffer from serious spatial dislocation from true edge locations in original image space. These results do not convey any convincing feeling that edge detection at 3 cyc/deg is the basis for form perception.

B. Oriented Operators at 3 Cycles/Degree

Before shifting from 3 cyc/deg to higher spatial frequency operators, we can examine operators with an orientation axis to see if this very cortical character for an operator provides some better representation than the circular case. One vertical and one horizontal ridge-with-valleys operator are used. The merged results (Fig. 7) illustrate a problem which may be fundamental to oriented operators, i.e., an ambiguity in edge extent and location for edges with geometrical shape at the scale of the operator. This problem appears to arise from the locational uncertainty introduced along the orientation axis simply by making the operator have an oriented response. This ambiguity will reappear at higher spatial frequencies when oriented operators are used. In any event, there seems to be no pathway to the capture of form information using either type of operator at 3 cycles/degree.

Edge Detection - Circular Versus Oriented Operators at 33 Cycles/Degree in a High Acuity Visual Task

The reading of printed text is a high acuity visual task since the narrowest linewidths of lettering for most print occupies a visual angle of ≈ .015 degrees. Edge detection of printed text is then a reasonable test for high spatial resolution comparison with visual perception. The well defined finely detailed character of letters exposes flaws in edge representations readily. For this experiment the operators used are shown in Table 1. The operator for the circular DOG has been shown to produce the equivalent to a convolution of the original scene intensity distribution with a circular DOG by making use of the Gaussian blur of the image forming optics (Ref. 10). The oriented operator is intended to be a crude approximation to the smallest possible ridge with two valleys. The actual continuous character of the operator in the scene domain has not been investigated. A vertical and horizontal operator are again used for initial visual assessment. Various contrast thresholds for edge detection are explored (Fig. 8). No particular threshold can be found which eliminates the "cross stitch" artifacts while retaining all edges whose orientation matches that of the operator. Perhaps a fuller set of orientation would improve performance but fundamentally the oriented operators confound edge extent with edge contrast in a manner which is not clearly resolvable. This difficulty is not encountered with the circular DOG where locational uncertainties are equal and can be minimized in all directions at once by choice of diameter. The case of obtaining an accurate edge representation with the circular DOG in the printed text case (Fig. 9) is obtainable over a wide range of edge detection contrast thresholds while the best result with the oriented operator does not approach the accuracy of the visual perception of printed text.

Recapitulation of Previous Results - Edge Detection and Contour Information Extraction at 11 and 33 Cycles/Degree Using Circular DOG Operation

Experimental results of edge detection and contour information extraction (Ref. 11) are revisited to provide a more comprehensive perspective on the visual information content of different spatial frequency channels within the overall human-primate spatial frequency response. In the previous work primary contour bearing information channels (Fig. 10) were found to be the highest spatial frequency channels, 11 and 33 cycles/deg. Further, the 11 cyc/deg channel is a higher contrast subset of all image phenomena visible at 11 cyc/deg. Contour extraction methods were based on the combined visual significance of relatively high contrast, minimum degree of connectivity, and limitations on maximum spatial density of edge events. The two channels are merged by giving the 33 cyc/deg channel priority in any local rivalry. The resulting contour images (Fig. 11) do differ in some particulars from those of Ref. 11. This reflects some changes in contour extraction methods in the interim.

Perceptual Evidence for a 33 Cycle/Degrees Channel

The idea, that two high spatial frequency channels with complementary spatial resolution and contrast sensitivity are necessary and sufficient basis for human-like form perception in some machine implementation, is considered further. Clearly, the higher contrast sensitivity - coarse spatial resolution channel is essential to capture lower contrast phenomena that make up much of perceived structure in images. But the need for the 33 cyc/deg channel is not so obvious. The geometrical fidelity of detected edges at 11 cyc/deg and 33 cyc/deg for a image of printed text is now considered (Fig. 12-14) in relation to visually perceived outlines of letters. The original image is shown side by side with the original image blurred by an amount approximately equal to the 11 cyc/deg channel. Below the images are the detected edge representations for the 33 and 11 cyc/deg channels respectively. The 11 cyc/deg representation is rich in geometrical defects not found in the visual perception of the original image. These defects are especially obvious in the details (Fig. 13-14) of the edge representations. Therefore, it is necessary to require a 33 cyc/deg channel in order to duplicate the geometrical fidelity present in the visual perception of high contrast finely detailed image phenomena.

Discussion

Taken as a whole, the results herein prompt speculation about the role of oriented response simple neurons in primate vision. Considering both the coarse spatial resolution and the extent ambiguities encountered in edge detection experiments, these neurons seem more amenable to encoding shading information or perhaps connecting shading information to contour information derived from other classes of neurons (perhaps the parvocellular subsystem of the LGN). The apparent absence of high visual acuity responses in the striate cortex is acutely disturbing since units of this type should be present in very large numbers.

For machine vision, these results argue strongly against the use of oriented responses especially when coupled with the unavoidable computational complexity of such a family of operators. Further, the circular operator seems capable of producing all information that oriented responses are capable of as well as accurate locational definition of corners, tight loops, and other high acuity image phenomena which are not captured easily or at all by oriented responses. Naturally the orientation of extended straight or moderately curved edges can be computed if needed from the detected edge loci of circular DOG operators.

References

1. DeValois, R. L.; Morgan, H.; and Snodderly, D. M.: Psychophysical Studies of Monkey Vision - III. Spatial Luminance Contrast Sensitivity Tests of Macaque and Human Observers. Vision Res., Vol 14, 1974, pp. 75-81.

2. Thorell, L. G.; DeValois, R. L.; and Albrecht, D. G.: Spatial Mapping of Monkey VI Cells with Pure Color and Luminance Stimuli. Vision Res., Vol 24, 1984, pp. 751-769.

3. Dow, B. M.; Snyder, A. Z.; Vautin, R. G.; and Bauer, R.: Magnification Factor and Receptive Field Size in Foveal Striate Cortex of the Monkey. Exp. Brain Res., Vol. 44, 1981, pp. 213-228.

4. Schiller, P. H.; Finlay, B. L.; and Volman, S. F.: Quantitative Studies of Single-Cell Properties in Monkey Striate Cortex III - Spatial Frequency. J. Neurophysio., Vol. 39, No. 6, Nov. 1976, pp. 1334-1351.

5. Silverman, M. S.; Grosof, D. H.; DeValois, R. L.; and Elfar, S. D.: Spatial-Frequency Organization in Primate Striate Cortex. Proc. Natl. Acad. Sci., Vol. 86, Jan. 1989, pp. 711-715.

6. Lennie, P.: Krauskopf, J.; and Selar, G.: Chromatic Mechanisms in Striate Cortex of Macaque. J. Neurosci., Vol. 10, No. 2, Feb. 1990, pp. 649-669.

7. Poggio, G. F.; Doty, R. W., Jr.; and Talbot, W. H.: Foveal Striate Cortex of Behaving Monkey: Single Neuron Responses to Square-Wave Gratings During Fixation of Gaze. J. Neurophysio, Vol. 40, No. 6, Nov. 1977, pp. 1369-1391.

8. Derrington, A. M.; and Lennie, P.: Spatial and Temporal Contrast Sensitivities of Neurones in Lateral Geniculate Nucleus of Macaque. J. Physio, Vol. 357, 1984, pp. 219-240.

9. Marr, D.: Vision. W. H. Freeman and Co., 1982.

10. Jobson, D. J.: Spatial Vision Processes: From the Optical Image to the Symbolic Structures of Contour Information. NASA TP 2838, Nov. 1988.

11. Jobson, D. J.: Isolating Contour Information from Arbitrary Images. Visual Information Processing for Television and Telerobotics. NASA CP 3053, May 1989, pp. 170-190.

SMALLEST CIRCULAR DIFFERENCE-OF-GAUSSIAN (DOG)

-.0675	-.182	-.0675
-.182	1.0	-.182
-.0675	-.182	-.0675

SMALLEST DIRECTIONAL RIDGE-VALLEYS

-.002	-.125	-.499	-.125	-.002
.004	.249	1.0	.249	.004
-.002	-.125	-.499	-.125	-.002

HORIZONTAL (AND VERTICAL BY ROTATION)

Table 1. Operators for Edge Detection Experiments at 33 Cycles/Degrees

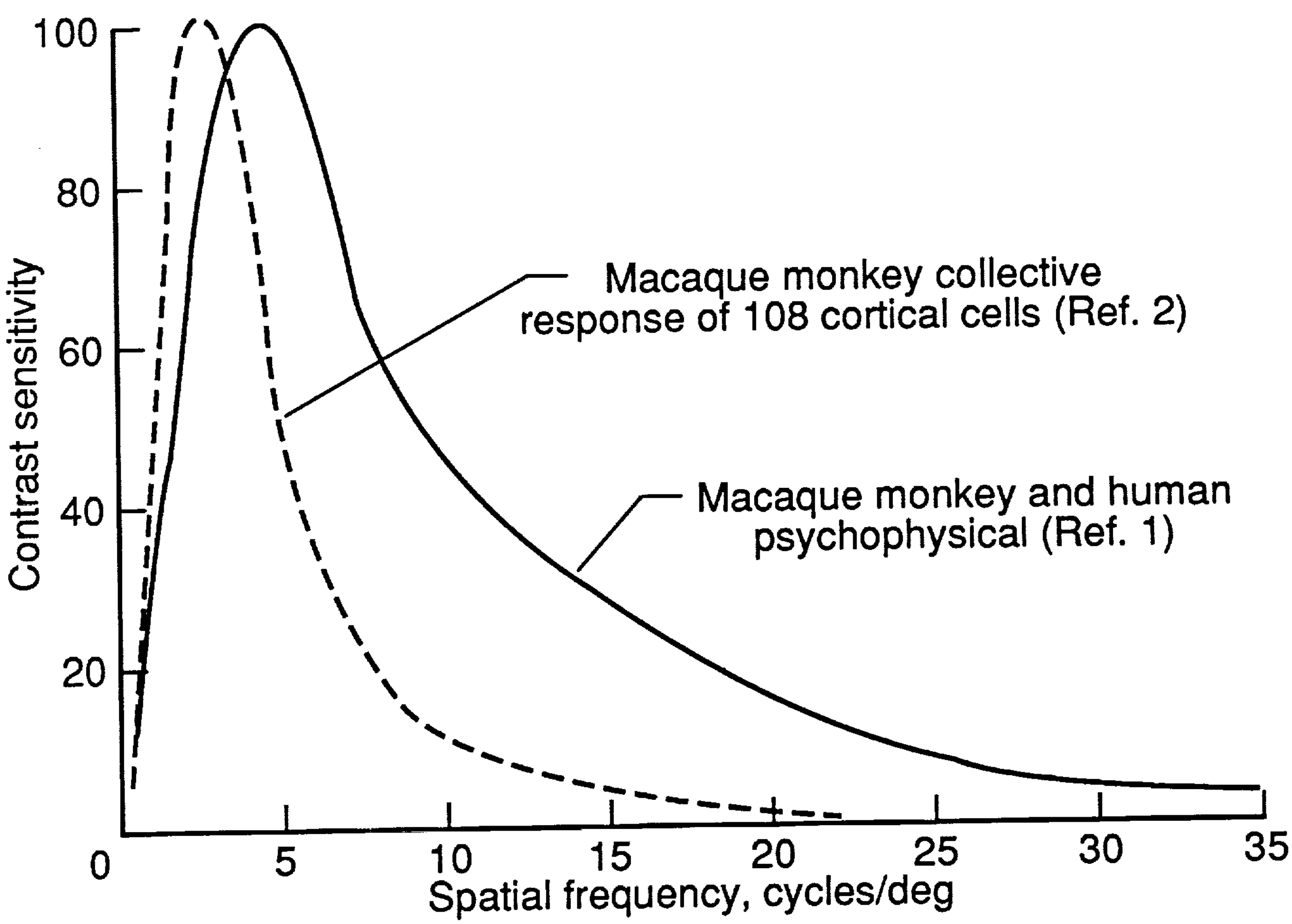


Fig. 1 - Comparison of Collective Physiological with Perceptual Spatial Frequency Responses for Foveal Primate Vision

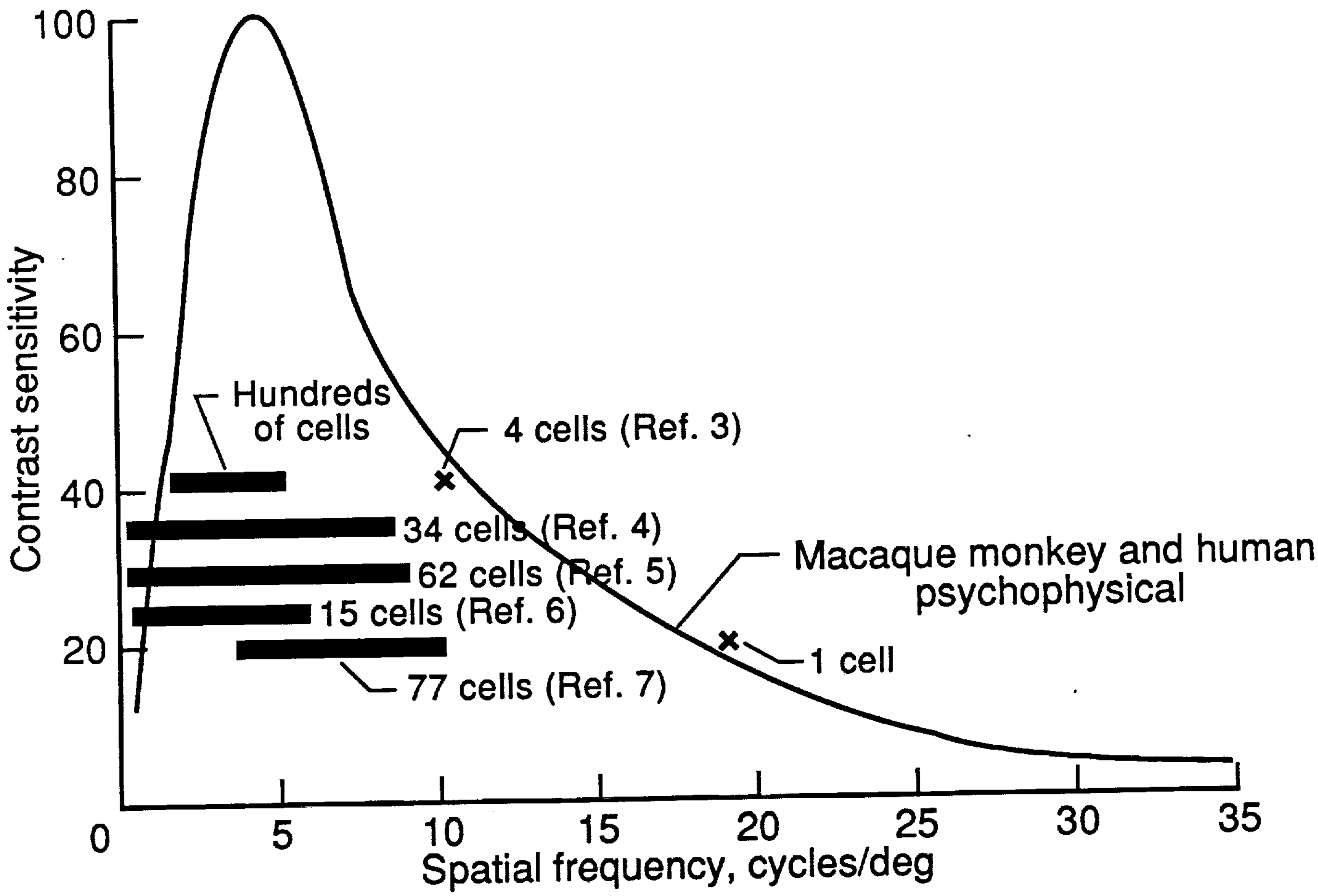


Fig. 2 - Further Physiological Data on Foveal Cortical Neurons - Range of Peak Spatial Frequency Responses.

G/PRMT CRTCL DATA|90-1268

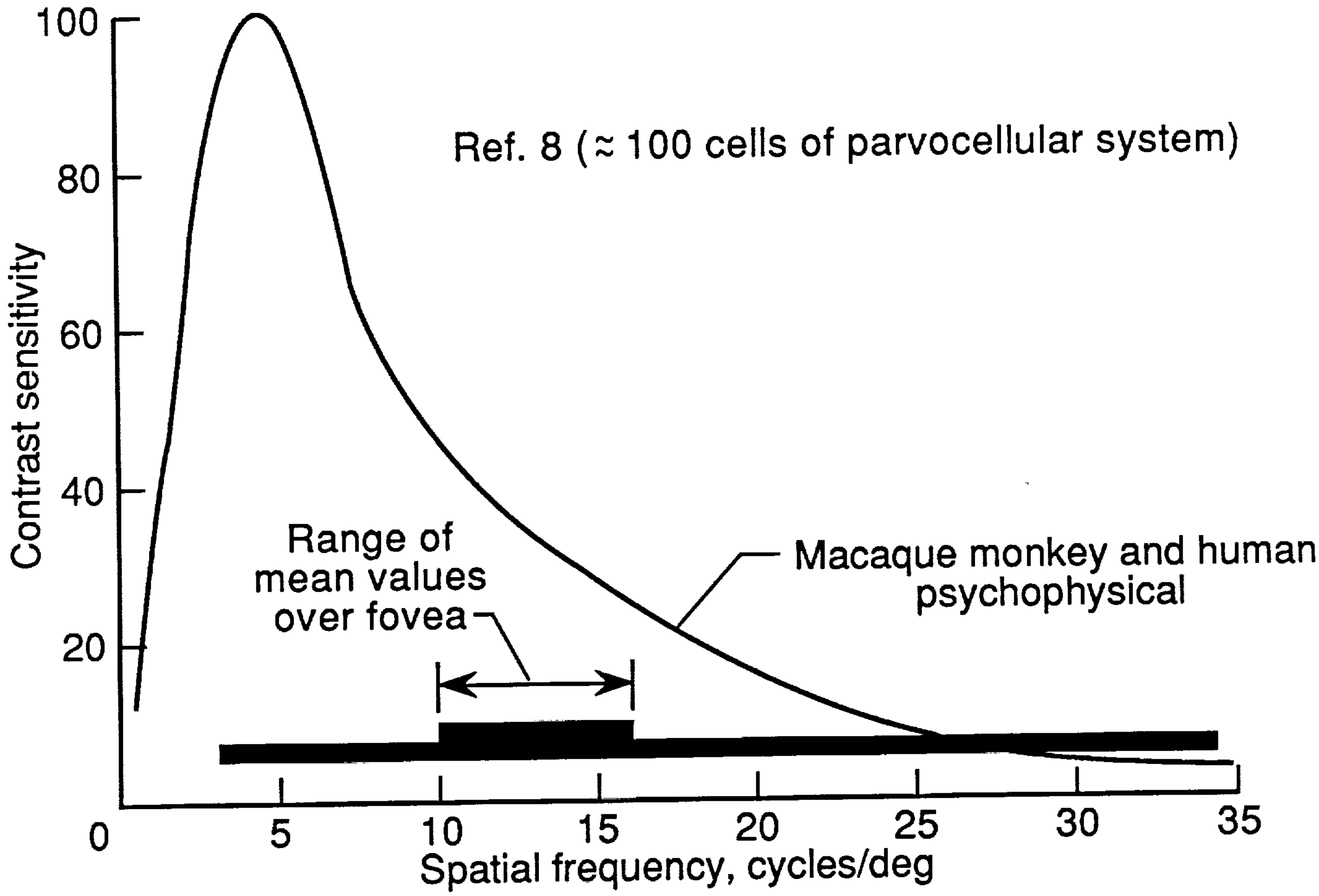
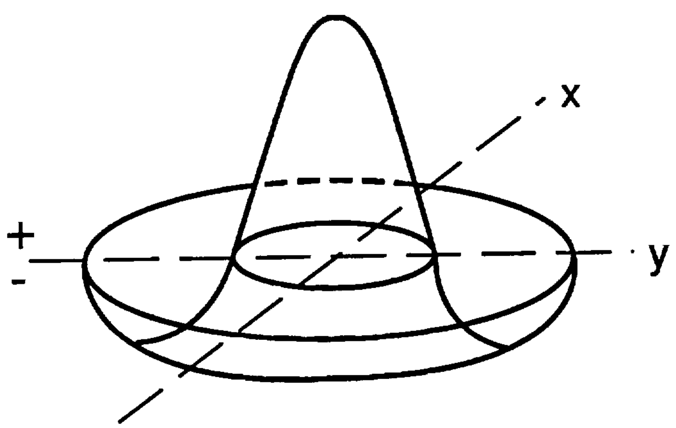
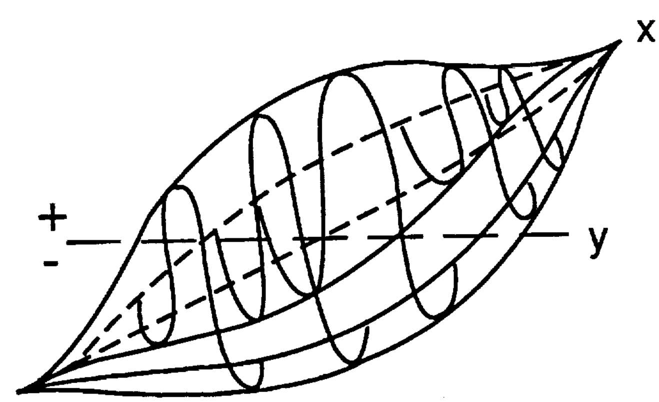


Fig. 3 - Physiological Data on Foveal Cells in Primate Lateral Geniculate Nucleus - Peak Spatial Frequency Responses



Mexican hat
(Difference of Gaussian (DOG))
or
Circular Gaussian Cosine



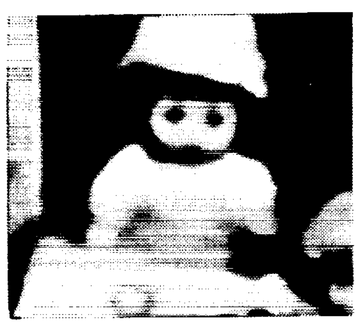
Ridge and valleys
(Gaussian + DOG)

Fig. 4 - Spatial Responses in Primate Vision Used for Zero-Crossing Edge Detection Experiments

I/MJR SPTL RSPNSS|90-1268

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

a) original



b) blurred

Figure 5. Visual Effect of Blurring That is Approximately Equivalent to a 3 Cyc/Deg Operator

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

ORIGINAL PAGE IS
OF POOR QUALITY

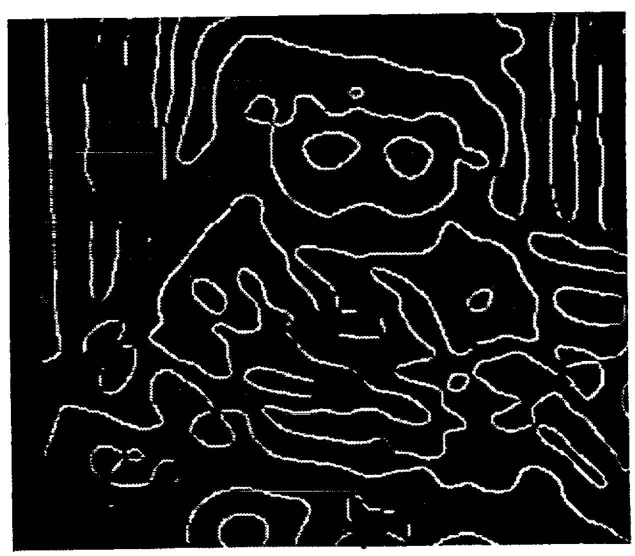
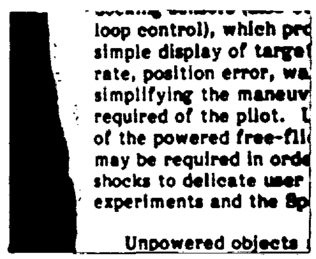


Figure 6. Results - Edge Detection at 3 Cyc/Deg Using Circular DOG Operator



loop control), which pr
simple display of targe
rate, position error, wa
simplifying the maneuv
required of the pilot.
of the powered free-fli
may be required in ord
shocks to delicate user
experiments and the Sp

Unpowered objects

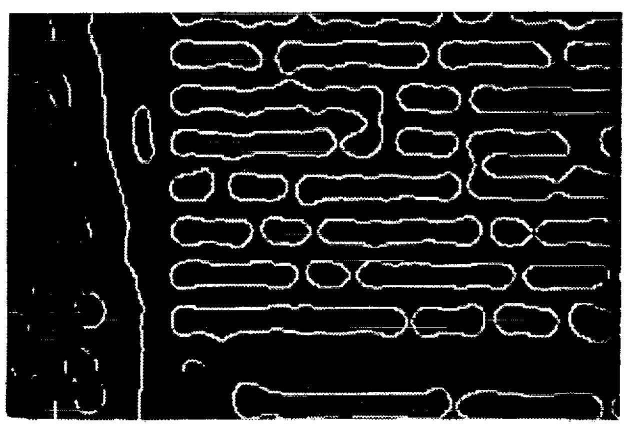
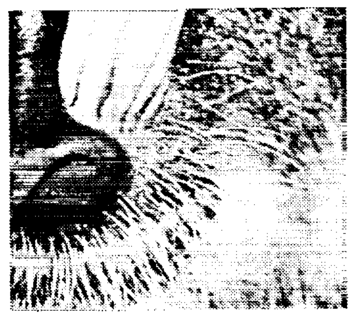


Figure 6. Continued



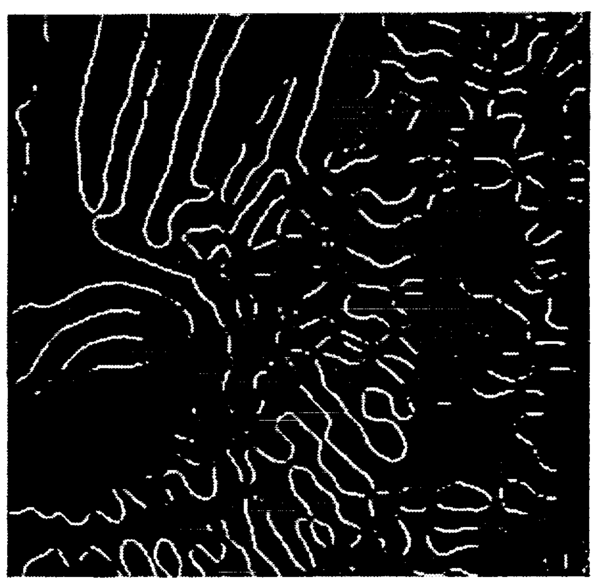
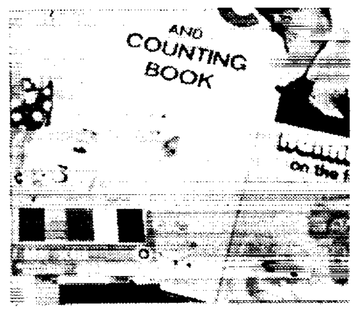


Figure 6. Continued

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH



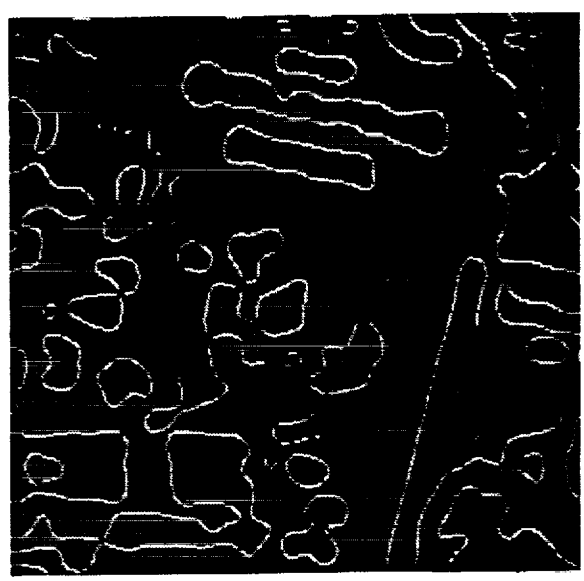


Figure 6. Continued

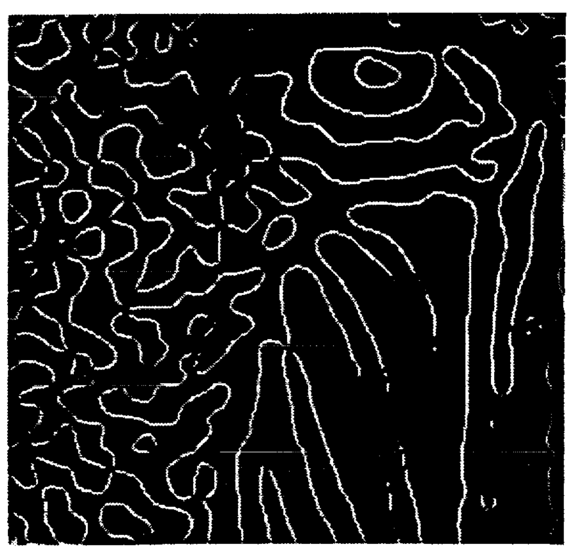
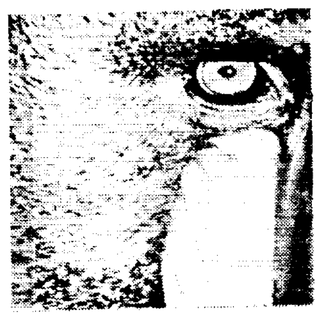
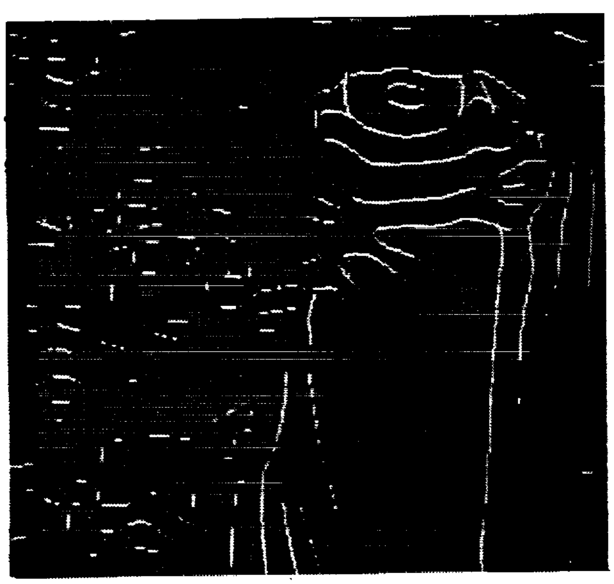


Figure 6. Continued





ETH = 13

ETH = 6

Two different contrast thresholds for edge detection.

Figure 7. Results - Edge Detection at 3 Cyc/Deg Using Vertical and Horizontal Operators

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

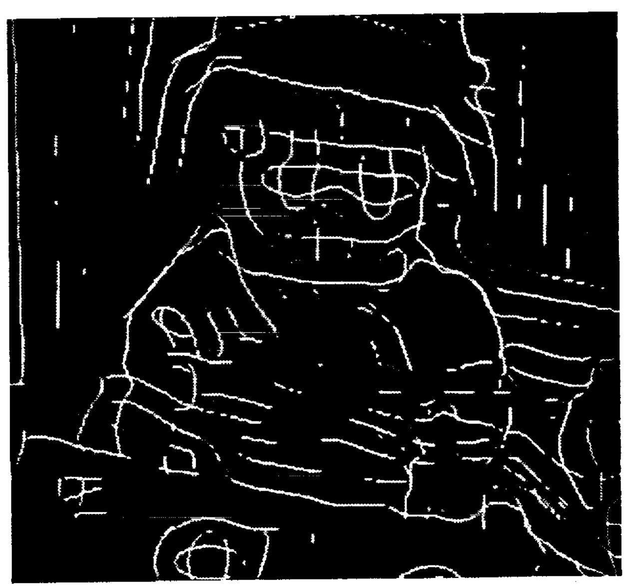
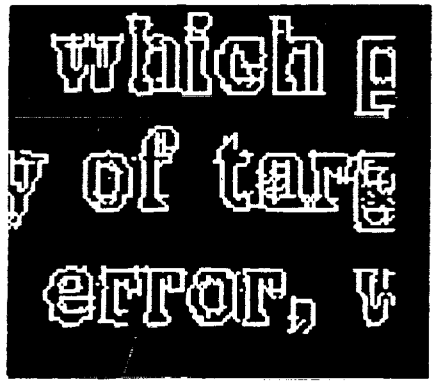
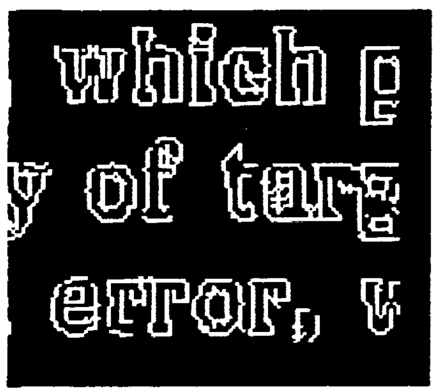


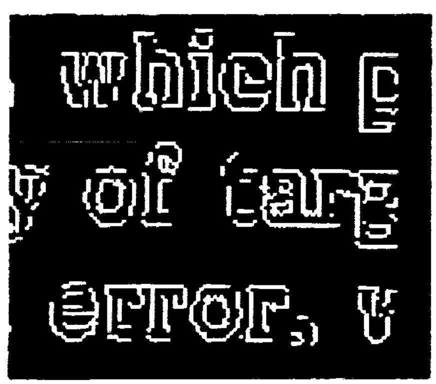
Figure 7. Continued



High

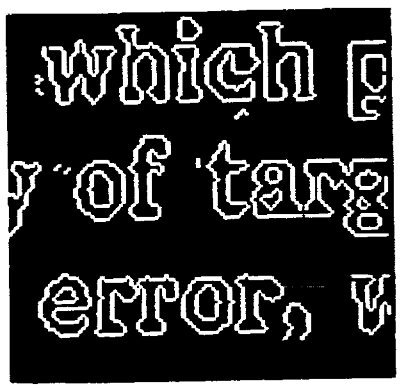


Medium

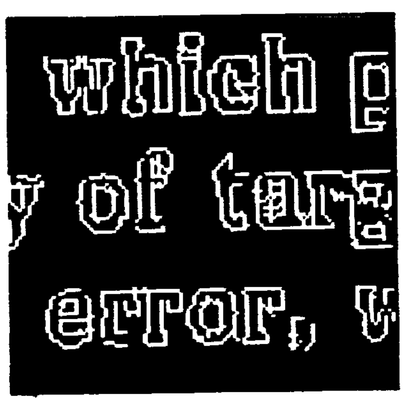


Low

Figure 8. Results - Edge Detection at 33 Cyc/Deg in a High Acuity Task Using Vertical and Horizontal Operators for a Range of Contrast Thresholds



Circular DOG



Directional

Figure 9. Performance Comparison of Circular and Oriented Operators at 33 Cyc/Deg

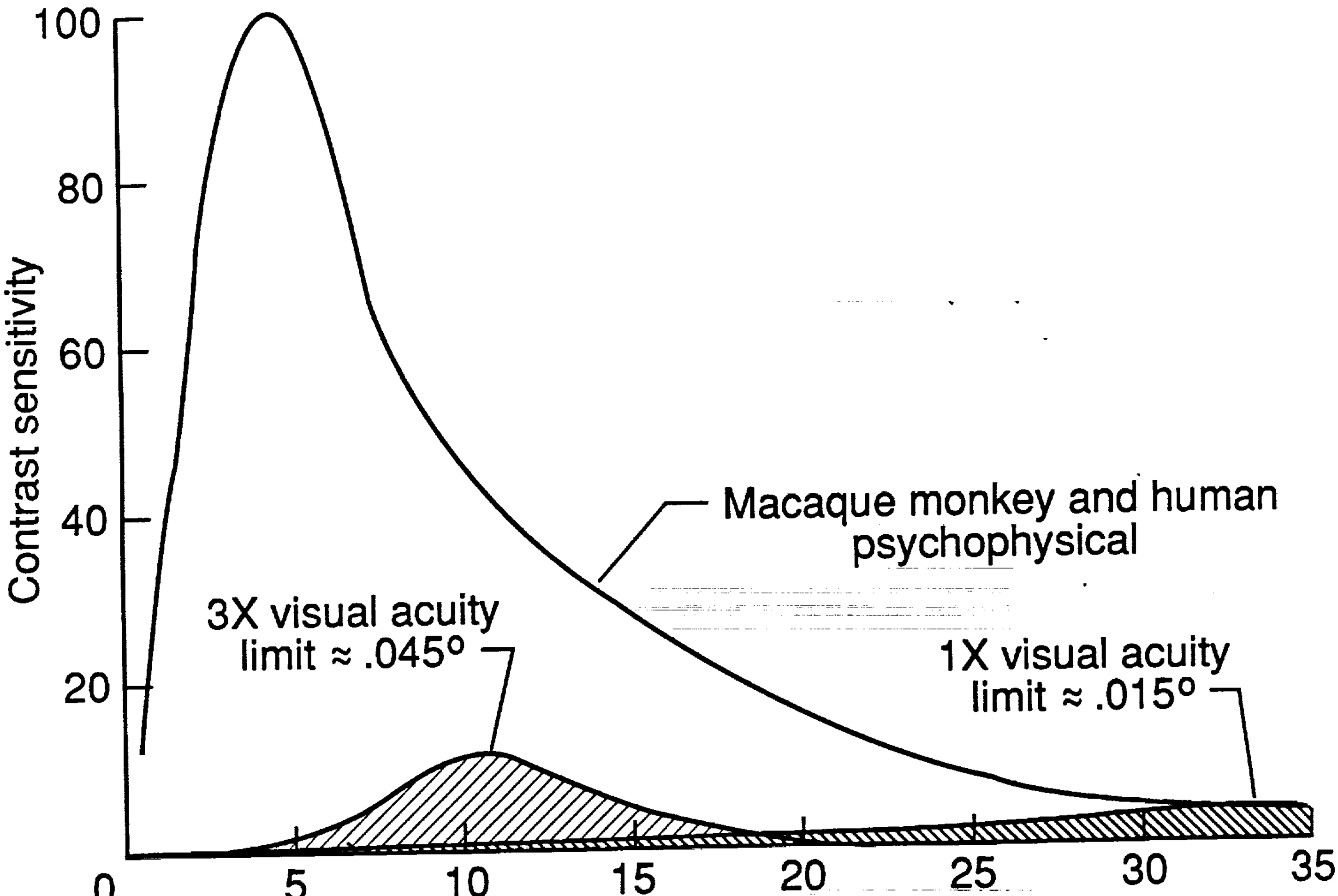
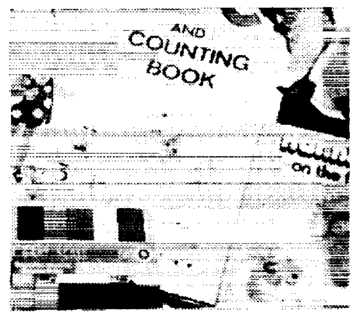


Fig. 10 - Contour Information Channels



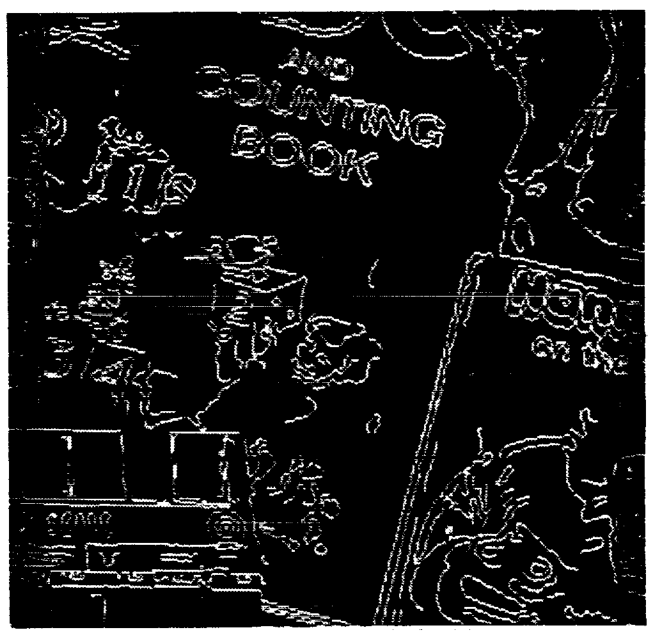


Figure 11. Recapitulation of Merged Results of Edge Detection at 11 Cyc/Deg and 33 Cyc/Deg with Additional Processing for Contour Extraction

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

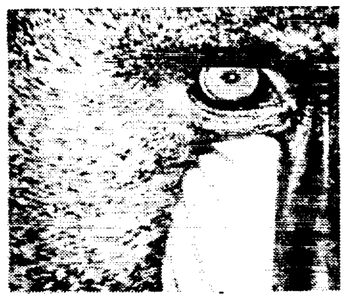


Figure 11. Continued

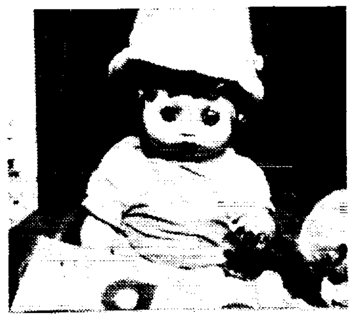
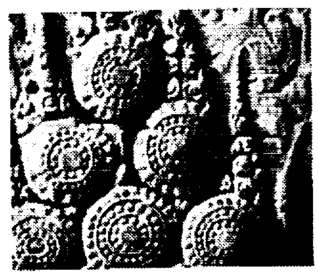


Figure 11. Continued



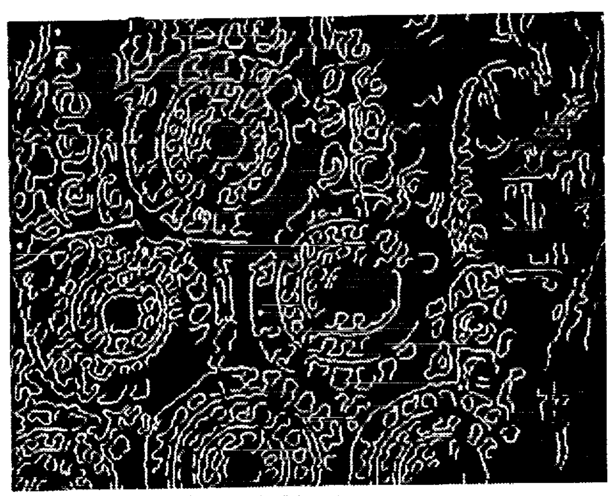
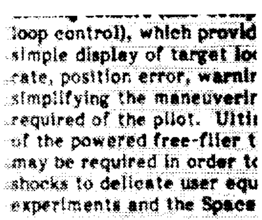
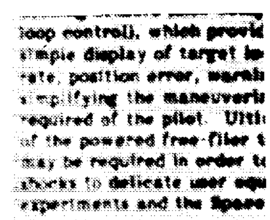


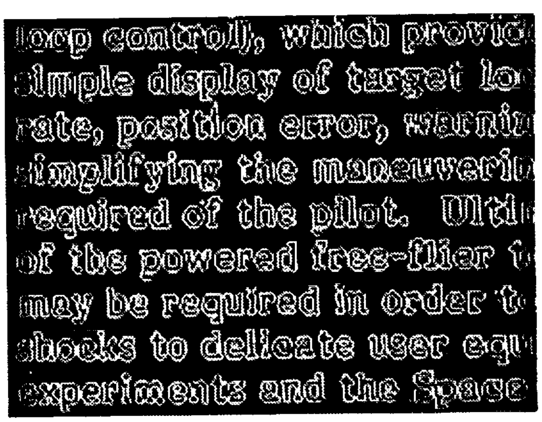
Figure 11. Continued



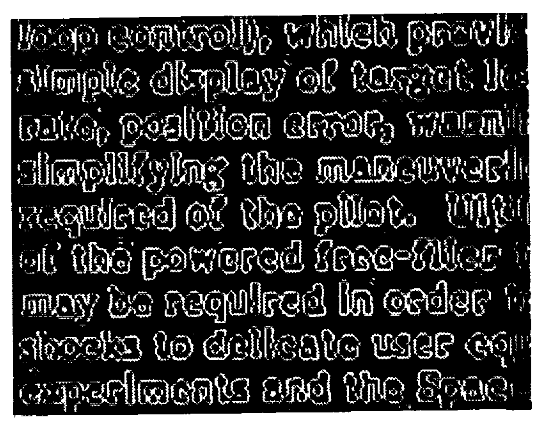
original



blur equivalent to 11 cyc/deg channel



a) Edge representation - 33 cyc/deg DOG



b) Edge representation - 11 cyc/deg DOG

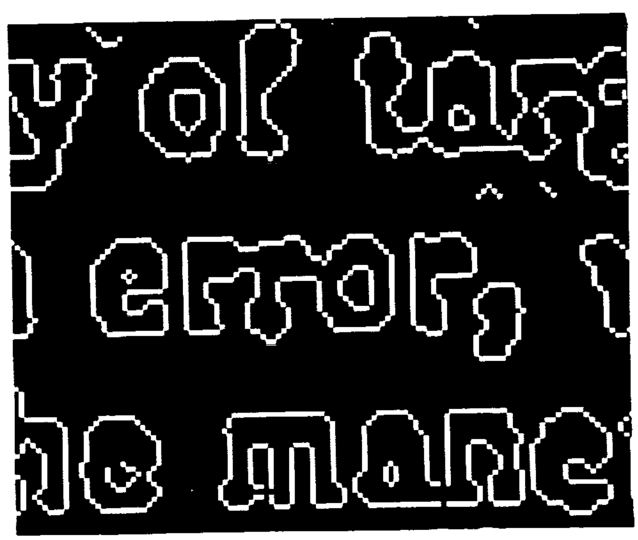
Figure 12. The Perception of Printed Text as Evidence of the Need for a 33 Cyc/Deg Channel

33 cyc/deg

11 cyc/deg

Figure 13. Details of Printed Text Edge Detection Results

33 cyc/deg



11 cyc/deg

Figure 14. Additional Details of Figure 12 a) & b)

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-102739		

4. Title and Subtitle	5. Report Date
A Discrepancy Within Primate Spatial Vision and Its Bearing on the Definition of Edge Detection Processes in Machine Vision	September 1990
	6. Performing Organization Code
7. Author(s) Daniel J. Jobson	8. Performing Organization Report No.
	10. Work Unit No. 307-51-10
9. Performing Organization Name and Address NASA Langley Research Center Hampton, Virginia 23665-5225	11. Contract or Grant No.
	13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract
The visual perception of form information is considered to be based on the functioning of simple and complex neurons in the primate striate cortex. However, a review of the physiological data on these brain cells cannot be harmonized with either the perceptual spatial frequency performance of primates or the performance which is necessary for form perception in humans. This discrepancy together with recent interest in cortical-like and perceptual-like processing in image coding and machine vision prompted a series of image processing experiments intended to provide some definition of the selection of image operators. The experiments were aimed at determining operators which could be used to detect edges in a computational manner consistent with the visual perception of structure in images. Fundamental issues were the selection of size (peak spatial frequency) and circular versus oriented operators (or some combination). In a previous study, circular difference-of-Gaussian (DOG) operators, with peak spatial frequency responses at about 11 and 33 cyc/deg, were found to capture the primary structural information in images. Here larger scale circular DOG operators were explored and led to severe loss of image structure and introduced spatial dislocations (due to blur) in structure which is not consistent with visual perception. Orientation sensitive operators (akin to one class of simple cortical neurons) introduced ambiguities of edge extent regardless of the scale of the operator. For machine vision schemes which are functionally similar to natural vision form perception, two circularly symmetric very high spatial frequency channels appear to be necessary and sufficient for a wide range of natural images. Such a machine vision scheme is most similar to the physiological performance of the primate lateral geniculate nucleus rather than the striate cortex.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Edge Detection, Spatial Frequency Channels, Primate Physiology, Form Perception Cybernetic Visual Methods	Unclassified - Unlimited Subject Category 63

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	30	A03

NASA FORM 1626 OCT 86