# NASA Technical Memorandum 104051

## MODES OF VISUAL RECOGNITION AND PERCEPTUALLY RELEVANT SKETCH-BASED CODING OF IMAGES

DANIEL J. JOBSON

FEBRUARY 1991

# NASA

National Aeronautics and
Space Administration

**Langley Research Center**
Hampton, Virginia 23665

# MODES OF VISUAL RECOGNITION AND PERCEPTUALLY RELEVANT SKETCH-BASED CODING OF IMAGES

by

Daniel J. Jobson

## Abstract

A review of visual recognition studies is used to define two levels of information requirements. These two levels are related to two primary subdivisions of the spatial frequency domain of images and reflect two distinctly different physical properties of arbitrary scenes. In particular, pathologies in recognition due to cerebral dysfunction point to a complete split into two major types of processing- high spatial frequency edge-based recognition versus low spatial frequency lightness (and color)-based recognition. The former is more central and general while the latter is more specific and is necessary for certain special tasks. The two modes of recognition can also be distinguished on the basis of physical scene properties - the highly localized edges associated with reflectance and sharp topographic transitions versus smooth topographic undulation. The extreme case of heavily abstracted images is pursued to gain an understanding of the minimal information required to support both modes of recognition. Here the intention is to define the semantic core of visual information and methods for rendering this core after high degrees of data compression and transmission. This central core of processing can then be fleshed out with additional image information and coding and rendering techniques to create more accurate representations as desired or needed.

## Introduction

In previous research[1,2] which defined edge detection methods and sketch abstraction methods, the sketch was considered to be a sufficient basis for visual recognition. During the course of image processing experiments, the sketch for one image was visually incomprehensible. While the sketches of a diverse battery of other test images were sensible as visual phenomena, this particular image cannot be treated as some isolated case. First the image is one of the Mars surface taken by the Viking Lander and therefore is an important case in point for NASA. Secondly, machine vision schemes for image coding which preserve key visual information should be as general as possible. No doubt if this image was not tractable, other "exceptions" would eventually appear to add to this population of one.

A resolution of this issue is found in the literature on visual recognition. Specifically the recognition of human faces seems to explain the problems encountered with the Mars surface image. A more general definition of modes of visual recognition is constructed from a review of research on visual recognition. Further since sketch-alone information is not always adequate for visual interpretation, methods for rendering are needed for recoupling sketch and shading information after compression and transmission. A variety of methods for this rendering process are considered and an attempt is made to define the maximum amount of blur that is permissible while retaining the visual sense of the smooth topography of the scene. An extreme case even in terms of high data compression ratio coding is examined which uses sketch-alone transmission followed by processing edge profiles to create the Craik-Cornsweet illusion. This case should approach the smallest numerical package that is possible while still retaining some sense of lightness values in the original image.

<u>Modes of Visual Recognition, Associated Scene Phenomena, and
Subdivision of the Spatial Frequency Domain of the Image</u>

The treatment of visual recognition can be approached as a beginning-to-end set of processes that proceeds from scene phenomena through image formation and sampling, subsequent processing to extract vital information, and finally the task of recognition itself (fig.1). In previous work this sequence was explored. Here the sequence is reversed by first considering the character of visual deficits which occur with cerebral impairments (such as stroke, trauma, tumors, and multiple sclerosis). Studies of cerebral visual impairment fall into two broad categories- 1) case studies of the perceptual deficits resulting from injury or disease and 2) measurements of changes in the patients' spatial frequency responses associated with cerebral dysfunction. The perceptual deficits are primarily descriptive and provide indirect evidence of underlying mechanisms that point to major subdivision of the spatial frequency domain by the mainstreams of visual processing. On the other hand, spatial frequency measurements supply direct evidence for frequency subdivision without necessarily much descriptive information on the perceptual deficit. While actual measurements of perceptual spatial frequency responses for cerebral dysfunction are sparse, [3-5] these taken with the more descriptive studies[6-8] do produce a reasonably clear subdivision of the spatial frequency domain of images. Finally the study of face recognition by observers with normal vision supports the idea of modes of visual recognition based in a fundamental subdivision of the spatial frequency domain. Although many questions remain to be answered about the nature of the processing by the main pathways of visual system, figure 1b seems consistent with the review of recognition studies which will now be detailed.

Clinical categories[6] of visual dysfunctions of cerebral origin are:

1) visual agnosia- inability to recognize objects
2) visual prosopagnosia- inability to recognize familiar faces
3) color agnosia- inability to recognize colors

2) and 3) are closely related and often occur together. Both 2) and 3) appear to be based on a low spatial frequency (blurred) representation of the image. The exact character of this low frequency channel isn't known. Distinct from this channel is one based on the high spatial frequency content of images. A breakdown somewhere in this pathway appears to be responsible for visual agnosia. This subdivision of spatial frequency adheres to the current idea[9] of a segregation of form and color information and is also supported by an edge detection study.[2]

Face recognition deserves further discussion in order to work backwards from perception to scene properties. The recognition that an object is a face falls under object recognition. Prosopagnosia is a disorder whose principal visual deficit is the inability to recognize familiar faces. Although faces can be matched as identical, they cannot be identified. This implies that the information in the image necessary for face recognition is probably present in other pathways but missing from the one for recognition. This visual deficit could be due either to the loss of the ability to process low spatial frequency image information or loss of the ability to make memory comparisons. In any event the deficit appears to be connected to the absence of smooth topographic information contained in the low spatial frequencies. Line drawings of faces without shading information are not usually recognizable unless they are caricatures which exaggerate prominent unique features. In contrast images of familiar faces such as Lincoln or Mona Lisa are quite recognizable in heavily blurred images.[10] The sketch information on faces may play a role in face recognition but clearly is not necessary in many cases. A more detailed examination of agnosia and prosopagnosia characteristics conveys the sense that the necessary visual information for recognition is present in the brain but simply doesn't make it to the appropriate destination. For example, agnosia patients can read but very slowly on a letter-by-letter basis. These patients can recognize small local features and sometimes identify an object in this way. What appears to be missing is the ability to integrate form information into a global unified construct. Likewise,

prosopagnosia is considered to be a defect in integrative processing. Prosopagnosia patients do not retain complete object recognition skills. In particular, recognition of animals is severely impaired. It isn't clear whether this is due to animals being non-rigid objects or to some other characteristic of this class of objects such as similarity of gross forms and certain details. Additionally, prosopagnosia includes a recognition deficit for objects in partial or cluttered presentation which require "visual closure" by the brain.

Recognition deficits and normal face recognition imply a subdivision of the spatial frequency domain but measurements of perceptual spatial frequency response make this division explicit. Patients with spatial frequency responses only below about 2 cycle/degree have only general sensations of light and dark with no sense of form perception.[3] Patients with spatial frequency responses out to 10 cyc/deg experienced blurred vision and difficulty reading.[4] The most specific evidence for spatial frequency channels comes from the highly preferential losses in spatial frequency response measured in multiple sclerosis patients.[5] Although a few patients were found to have a visual loss extending across the entire frequency domain, most patients experienced a drop in frequency response that was primarily in one of three frequency ranges - less than 8 cyc/deg, between 10-17 cyc/deg, and above 17 cyc/deg. This defines three main subdivisions of the frequency domain-low, middle, and high. Visual performance for patients with a middle loss was the experience of vision as being "washed out" but no deficit existed for high acuity-high contrast tasks (Snellen tests). This supports the idea that a least two high pass channels are necessary for form perception in normal vision. These are 10-17 cyc/deg for low contrast-low acuity image phenomena and about 17-33 cyc/deg for high contrast-high acuity phenomena.

The subdivision of the spatial frequency domain of images into low and high partitions which was related to distinct modes of visual recognition can also be traced to distinct classes of scene phenomena. High spatial frequencies bear information on sharp reflectance and topographic transitions, while the low range bears information on smooth undulatory surface topography (shading in visual terms).

All of these considerations lead to the definition of a visual processing scheme based on the three spatial frequency channels (fig. 2). Though the discussion here has focussed on the spatial frequency domain, the spatial domain is perhaps more important since the two high frequency channels appear to bear edge location and contrast information. This type of information is decidedly more spatial in character than spatial frequency, since frequency implied a more regional sense for support of sufficient cycles to define a wave phenomena. The two high spatial frequency channels illustrated were defined from edge detection-perception studies[2] but are consistent with both the perceptual deficits of MS patients and the preferential spatial frequency losses measured.

The character of the low spatial frequency channel remains to be defined. The thrust here is to blur and decimate as much as possible for maximum data compression while retaining only as much visual information as is needed to support both modes of visual recognition. Therefore the remainder of this paper will attempt to define the low spatial frequency channel and methods for recoupling this shading information with the edge-based sketch to produce a rendering that is visually comprehensible without being a precise reconstruction in a numerical sense.

## Sketch-Based Coding and Image Rendering

The preceding review of modes of visual recognition and evidence for defining three spatial-spatial frequency channels for recognition processing leads now to applied sketch-based coding and rendering experiments. A major distinction is necessary at the outset. Historically, image coding has been cast as a problem of compressing, transmitting, and reconstructing a reasonably accurate copy of the original image data. This is usually treated as a mathematical problem and the accuracy of results are judged on the basis of statistical measures. More recent efforts[11-12] have introduced the idea of visual relevance into coding methods. Here an even broader view is taken by stating the problem as the acquisition and transmission of visual information in as concise a manner as is possible without introducing serious semantic damage into the rendering. The aim is

3

NOT to reconstruct the image or scene radiance distributions but rather to convey the primary visual information necessary for interpretation.

Here visual significance and relevance are the paramount concern but this approach requires some quantitative definition and demonstration. Numerically severe discrepancies between image and rendering are to be expected and absorbed willingly if visual sense is maintained. This approach is entirely in the spirit of the illustrator seeking to render accurately without attempting the impossible task of duplication.

Some discussion of the relative importance of recognition modes is also of value before proceeding. The sketch-alone is clearly sufficient for the vast majority of natural images and is therefore considered here to be the information of highest significance. Shading information is only required in very specific cases where smooth topography is essential to visual interpretation. These cases may be infrequent but are non-negligible as already noted for face recognition and the visual interpretation of Mars surface images. At this point an illustration of the necessity of shading information is appropriate (fig.3). The sketch alone for these two images does not convey sufficient information for visual recognition. The woman's face is identifiable as such but the specific person is not. The sketch extracted for the Mars surface is not comprehensible as anything ≥in particular. Without any a priori knowledge, this representation is meaningless. So for these two cases, something or everything of visual importance is missing in the sketch by itself.

The first endeavor at combining shading and sketch information was simply to overlay the two (fig.4). The result is visually unconvincing and gives a feeling of insufficient coupling. The images have a flat distorted quality that is visually wrong.

In order to numerically recouple sketch and shading some crude information is needed beyond just edge location - namely, edge orientation and polarity of contrast across the edge. Very crude contrast magnitude information is already available by virtue of tagging which of the two high channels detected the edge. This will be made use of in rendering. Previous edge detection methods and sketch extraction methods[1] were modified to classify edges into four classes of orientation and polarity (fig.5a). Diagonal edges are coded as horizontal to avoid an additional class and to examine the visual quality resulting from this simplification. This crude edge information can now be recoupled with shading in a variety of ways. First however the extreme case of recreating a crude sense of lightness values from sketch alone is examined.

## Sketch Rendering Using the Craik-Cornsweet Illusion

This case is of interest when visual communication is restricted to sketch alone such as video transmission over telephone lines for video conferencing[13] or visual sign language communication for the deaf.[14] Here the crude edge information is coupled into a uniform middle gray background using intensity profiles adjacent to the edge locations which invoke the Craik-Cornsweet optical illusion. By carrying edge polarity information the illusion is invoked in a manner that fills in lightness values via brain computation. Results are shown in Fig. 5 and using the CRAIK edge profile is shown in Fig. 6. The images are successful illusions because the large regional zones that appear to have different lightness values are numerically identical. It is interesting that the illusion works on reasonably complex natural images since its usual presentation is as a simple figure against a blank background. The images are rather convincing even though they have a ghostly insubstantial feeling about them. This presentation may be more useful than sketch-alone information presented as a line drawing.

## Sketch Rendering Using Low Spatial Frequency Information

Visual shading information can be transmitted separately and recoupled with sketch information in at least two ways a Gaussian blurred low pass image or a low frequency bandpass. The former allows for edge profiles to be carved directly into the low pass image while the latter proceeds as an

4

extension of the Craik-Cornsweet illusion. The edge profiles are carved out of a uniform middle gray and then the low bandpass image values are simply added to this. Note the bandpass image contains about as many negative as positive values so that addition here produces as many subtractions on average as additions. The latter approach was dropped after experimentation indicated poorer visual quality unaccompanied by any significant advantage. The types of edge profiles employed (fig. 6b) range from a spatially tight incised profile as was used for the Craik-Cornsweet illusion through looser wider spreading profiles to finally a pure step function that is not even tapered in intensity into low spatial frequency values. Computation proceeds from the edge location outward along either vertical or horizontal directions and ceases if an adjacent edge location is encountered. The simplest profile is called STEP and is a straightforward insertion of a step function into the low channel signal values according to:

$$I_{\pm i} = I_E(x,y) \pm 1/2 \overline{CS}$$

where $i = 1, 2,.....6$ and $\overline{CS} = 192$ for the highest frequency channel and $\overline{CS} = 83$ for the middle channel. These are the respective mean values of contrast for edges detected in these channels. The next profile type is called TAPER because it does blend outlying values of the edge profile into the low channel values according to:

$$I_{\pm i} = I_E(x,y) \pm 1/2 \overline{CS}$$

$$I_{\pm 2} = I_{\pm 1}$$

$$I_{\pm 3} = I_{\pm 1}$$

$$I_{\pm 4} = (I_{\pm 4} + I_{\pm 1})/2$$

$$I_{\pm 5} = (2I_{\pm 5} + I_{\pm 1})/3$$

$$I_{\pm 6} = (3I_{\pm 6} + I_{\pm 1})/4$$

where all I's are low pass intensity values adjacent to edge locations. Note STEP and TAPER differ only at several image elements away from the edge location. A spatially tight profile is called CRAIK and is computed as:

$$I_{\pm 1} = I_E(x,y) \pm \frac{1}{2}\overline{CS}$$

$$I_{\pm 2} = (I_{\pm 2} + I_{\pm 1})/2$$

$$I_{\pm 3} = (2I_{\pm 3} + I_{\pm 1})/3$$

$$I_{\pm 4} = (3I_{\pm 4} + I_{\pm 1})/4$$

The tightest profile called NARROW CRAIK is simply

$$I_{\pm 1} = I_E(x,y) \pm \frac{1}{2}\overline{CS}$$

CRAIK and NARROW CRAIK allow both a narrower spread and the effect on an illusory profile to be visually assessed. Now results are given for variable amounts of blur for the low channel.

5

Some initial selection of the amount of blur is necessary. As a starting point, the low pass image (Fig. 7) is created by a two-dimensional (2-D) convolution of the image with a 2-D circular Gaussian where full width at half maximum (FWHM) is 6 image pixels. It is recalled at this point that the two high pass channels have spatial extents of about 3 pixels (for 1X) and 9 pixels (for 3X). Therefore, the coupling (overlap in spatial frequency domain) between high pass and low pass is reasonably strong for FWHM = 6. Results then progress through larger amounts of blur up to FWHM = 20 image pixels. For this last case, the visual decoupling from the edge structure is more noticeable but even for this extreme case the impression is not particularly bad. Further, the face is probably still recognizable, unless a comparison were drawn with a very similar face.

The degree of spatial spread for computing from both 1X and 3X edges is the same. This is done as an attempt to stretch both 1X and 3X edges back into the heavily blurred lightness values and to assess visual effectiveness of this mathematical "shortcut". Perhaps this is not a shortcut and is justifiable in terms of the absence of adjacent edges being interpreted as visually insignificant signal changes in this area of the image.

A comparison (Fig. 7-10) of the four types of edge profiles and various amounts of blur produces no definitive sense of either a best profile or a maximum amount of tolerable blur. The broader spatial spread of STEP and TAPER does produce a better rendition at very high levels of blur while CRAIK is better at lower levels of blur. NARROW CRAIK does not purchase a successful effect except at FWHM = 6. Perhaps the lack of sharp differences in visual quality is due to the CRAIK between STEP, TAPER, AND CRAIK involving the optical illusion in a manner which somewhat offsets the narrowness of its spatial spread.

Results are now presented for a variety of images (Fig. 11), most of which have some smooth topographic information for the case of blur FWHM = 20 pixels. While all of the renderings are visually convincing and informative, this severe amount of blur is noticeable in image zones with very high contrast extended edges lacking geometrical detail (shape or close spacing of adjacent edges) with a relatively large area of black forming one side of the edge group (i.e., some areas near back of the shuttle wing and the page boundary in the text image).

This issue of maximum tolerable blur likewise has no clear-cut definition. Increasing blur allows greater spatial decimation of the image during compression but a point is reached where compression ratio approaches the minimum needed for the high pass edge data points. Edge locations for images of moderate complexity usually run about 10,000 in number and this coding scheme requires 3 bits per edge or 30,000 bits devoted to edge structure. Considering this allocation of bits, pushing blur to extremes (fig. 12) or greatly increasing spatial spread of lightness computations are not likely to be worthwhile. Modest increases in spatial spread are worth exploring especially if the visual quality improves markedly. This is almost certainly because the recognition mode based on low spatial frequency information is one of subtlety and specificity. Careful psychological studies might reveal some quantitative relationship between amount of blur tolerable and degree of similarity between two faces. This would still not be a generalization though. One very important point though is clear - very large amounts of blur relative to the spatial resolution of images are tolerable and still produce an interpretable and even pleasing image. These large amounts of blur are much greater than would be allowed for a mathematically complete sampling of spatial frequency domain. In other words, rather drastic mathematical short cuts are possible while preserving all necessary visual information.

## Conclusions

Two fundamental modes of visual recognition are assumed that are consistent with the major classes of visual cerebral dysfunction, normal recognition processes and the subdivision of the spatial frequency domain of images into high and low frequency information. The high portion is the basis for most general object recognition while the low portion contains shading information which is important for specific tasks such as face recognition. This basic spatial frequency

subdivision reflects a distinct division of scene properties into smooth topography for the low channel and sharp topography and reflectance boundaries for the high portion.

Since shading information in images is essential for one mode of visual recognition, methods of recoupling edge based sketch information with low spatial frequency shading information were investigated. The hope that some clear definition of the low spatial frequency channel and some optimal computation for recoupling edge and shading would emerge was not realized. Tolerances for these characteristics appear to be considerable and associated visual effects are often subtle. A painstaking psychophysical study would likely be necessary to use the subtle visual changes to determine some optimum point of visual quality. While this lack of precise definition is unsatisfactory from a purely scientific viewpoint, it allows for considerable practical latitude in designing systems to transmit and render minimal visual information.

## References

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

2. Jobson, D. J.: A Discrepancy Within Primate Spatial Vision and Its Bearing on the Definition of Edge Detection Processes in Machine Vision. NASA TM 102739, September 1990.

3. Bodis-Wollner, I.: Vulnerability of Spatial Frequency Channels in Cerebral Lesions. Nature, Vol. 261, 1976, pp. 309-311.

4. Bodis-Wollner, I.: Visual Acuity and Contrast Sensitivity in Patients with Cerebral Lesions. Science, Vol. 178, 1972, pp. 769-771.

5. Reagan, D., et al.: Visual Acuity and Contrast Sensitivity in Multiple Sclerosis - Hidden Visual Loss. Brain, Vol. 100, 1977, pp. 563-579.

6. Baron, R. J.: The Cerebral Computer, Lawrence Erlbaum Associates, Hillsdale, NJ, 1987.

7. Levine, David N. and Calvano, R.: Prosopagnosia: A Defect in Visual Configural Processing. Brain and Cognition, Vol. 10, 1989, pp. 149-170.

8. Mendez, M. F.: Visuoperceptual Function in Visual Agnosia. Neurology, Vol. 38, Nov. 1988, pp. 1754-1759.

9. Livingstone, M. and Hubel, D.: Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception. Science, Vol. 240, May 1988, pp. 740-749.

10. Harmon, Leon D.: The Recognition of Faces. Scientific American, Nov. 1973, pp. 70-82.

11. Carlson, Stefan: Sketch-Based Coding of Grey Level Images. Signal Processing, Vol. 15, 1988, pp. 57-83.

12. Kunt, Murat, et al.: Second Generation Image-Coding Techniques. Proc. of IEEE, Vol. 73, No. 4, April 1985, pp. 549-574.

13. Pearson, Don E. and Robinson, John A.: Visual Communication at Very Low Data Rates. Proc. of IEEE, Vol. 73, No. 4, April 1985, pp. 795-812.

14. Letellier, Phillipe, et al.: The Telesign Project. Proc. of IEEE, Vol. 13, No. 4, April 1985, pp. 813-827.
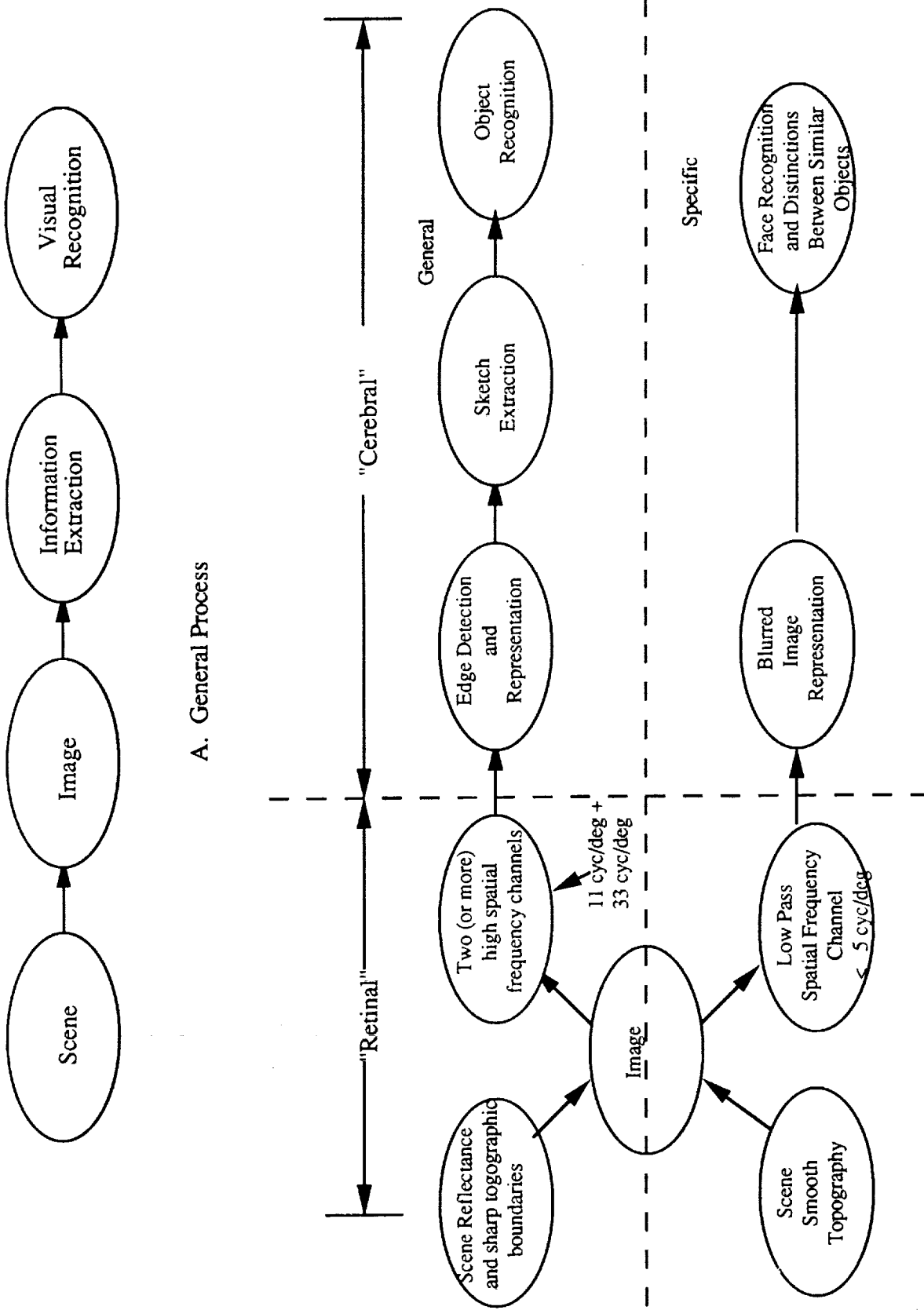
A. General Process

B. Two Modes of Visual Recognition Consistent with Cerebral Impairments

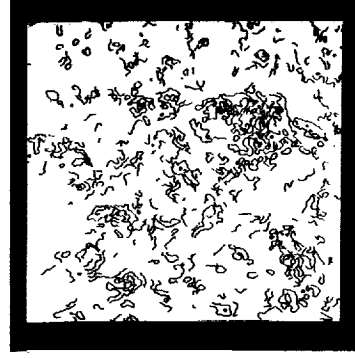Figure 1.  Framework for Visual Recognition Processes
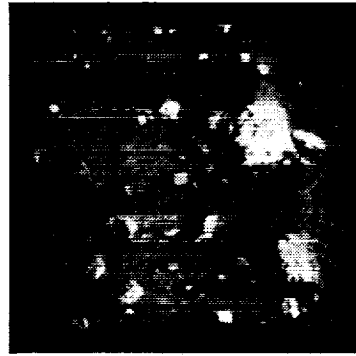
8
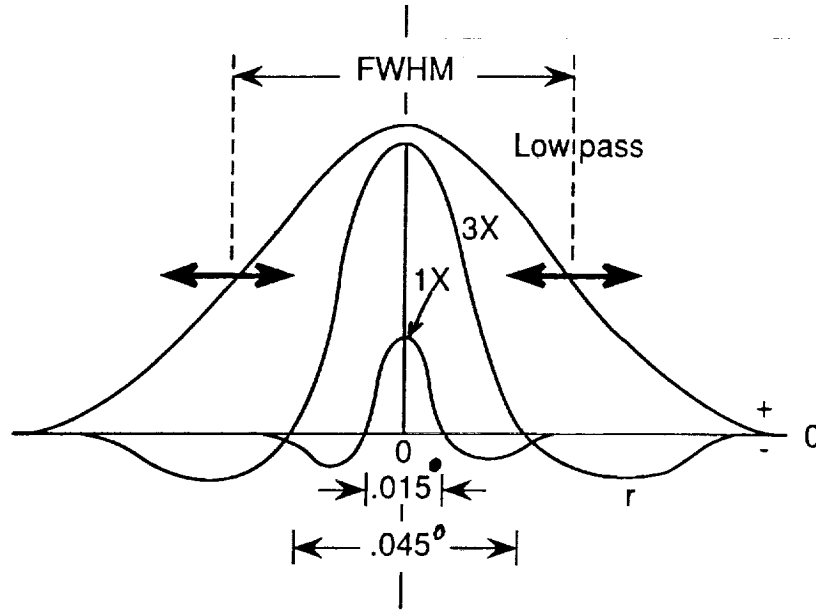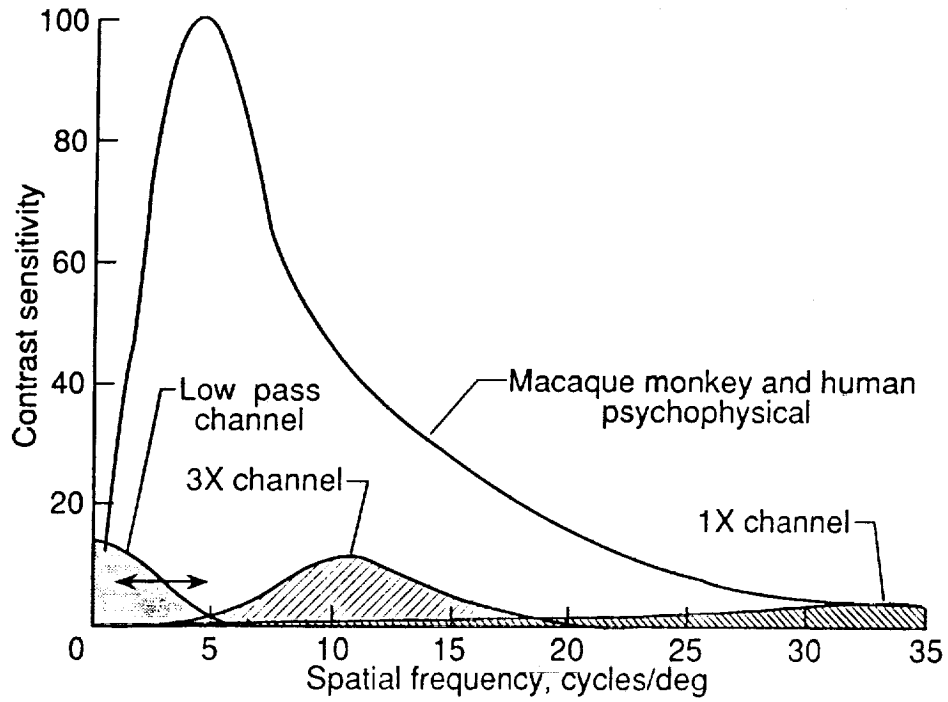
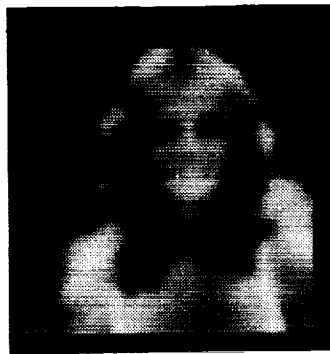a) Spatial Domain



b) Spatial Frequency Domain

Figure 2. Sketch-Based Coding - Image Operators
(Circularly Symmetric Functions in Cross Section)

**Sketch**

**Image**

**Low pass image**

Figure 3. Illustration of Two Cases Where Sketch-By Itself is Insufficient for Visual Interpretation.

Figure 4. Illustration of Inadequacy of Simply Overlaying Sketch and Low Pass Image Information.
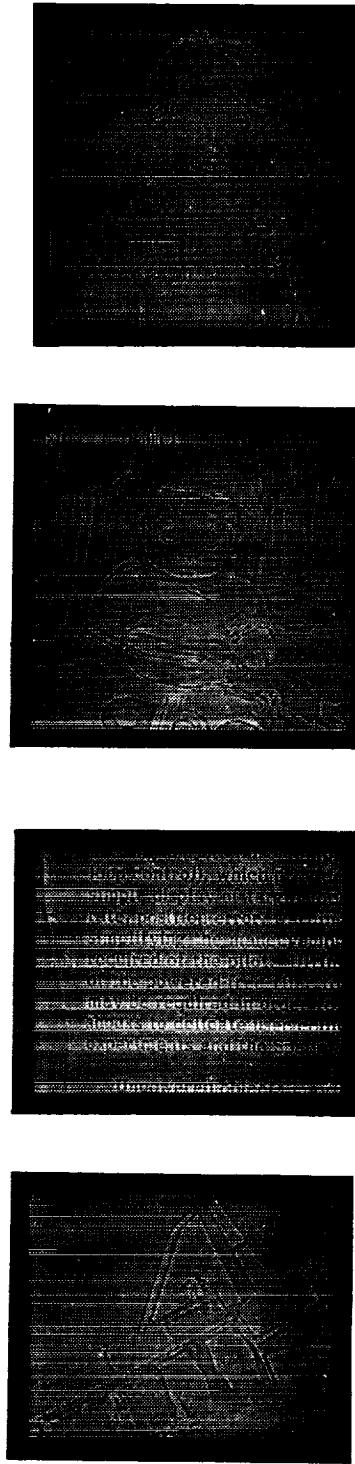
Figure 5. The Visual Effect of the Craik-Cornsweet Illusion as a Method for Coupling Sketch Information with a Uniform Gray Background.

Figure 5. (Concluded)

Vertical

$+$ | $-$

$-$ | $+$

Left plus

Right plus

Horizontal

$+$ / $-$

$-$ / $+$

Top plus

Bottom plus
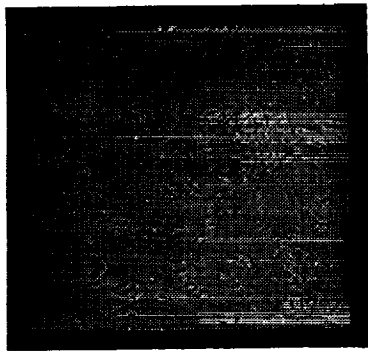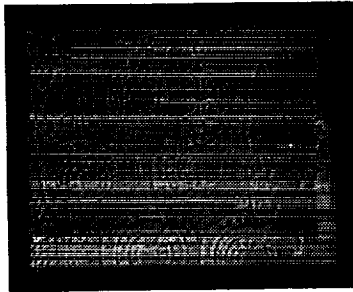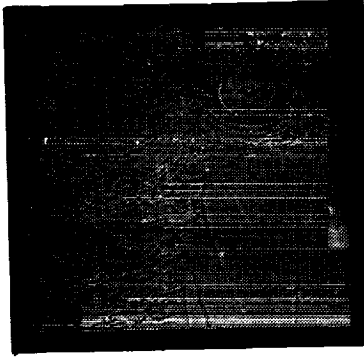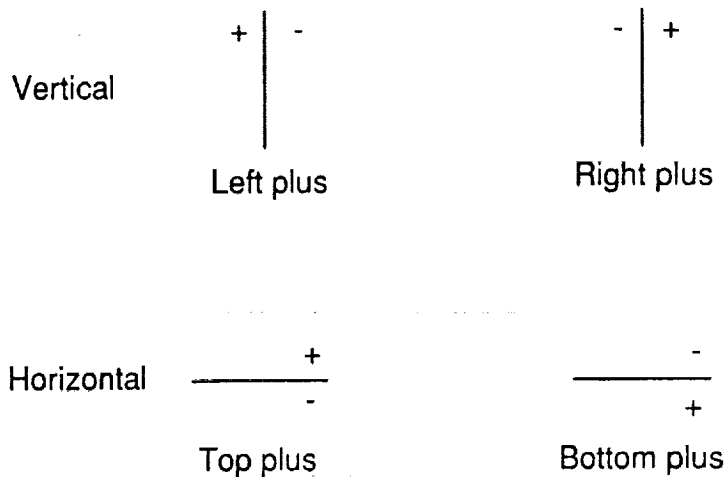
a) Classes of Edge Direction and Polarity Coding of IX and 3X Channels.

| — Edge location

$+$ | $-$ Contrast polarity

Step

Taper

Craik

Narrow Craik

Intensity

$I_E (x,y)$

1/2 $\overline{cs}$

Low pass image
$I (x,y)$

x or y
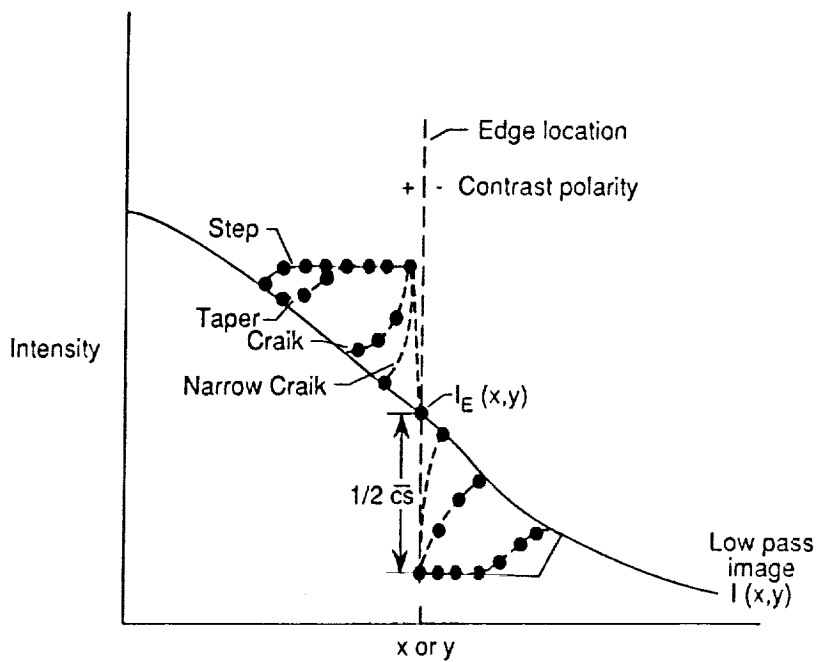
b) Various Methods for Rendering Spatial Intensity Profile of an Edge.

Figure 6. Schematic of Method for Recoupling Low Pass Image with Edge Structure.

FWHM = 6.0

FWHM = 11.0
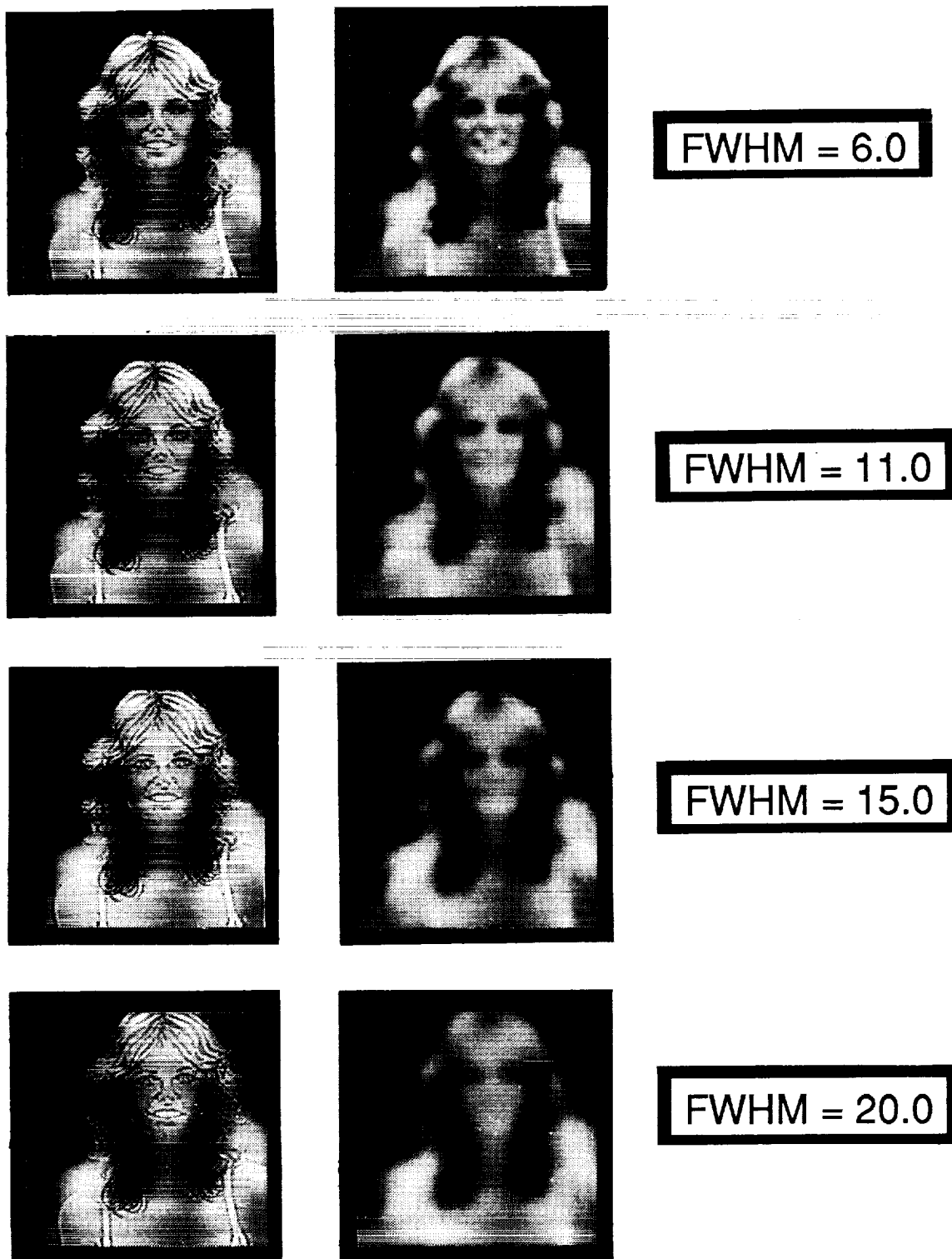
FWHM = 15.0

FWHM = 20.0

a) Face

Figure 7. Rendering Results for TAPER Method as a Function of Amount of Blur in Low Pass Image.

Rendering    Low pass component

FWHM = 6.0

FWHM = 11.0
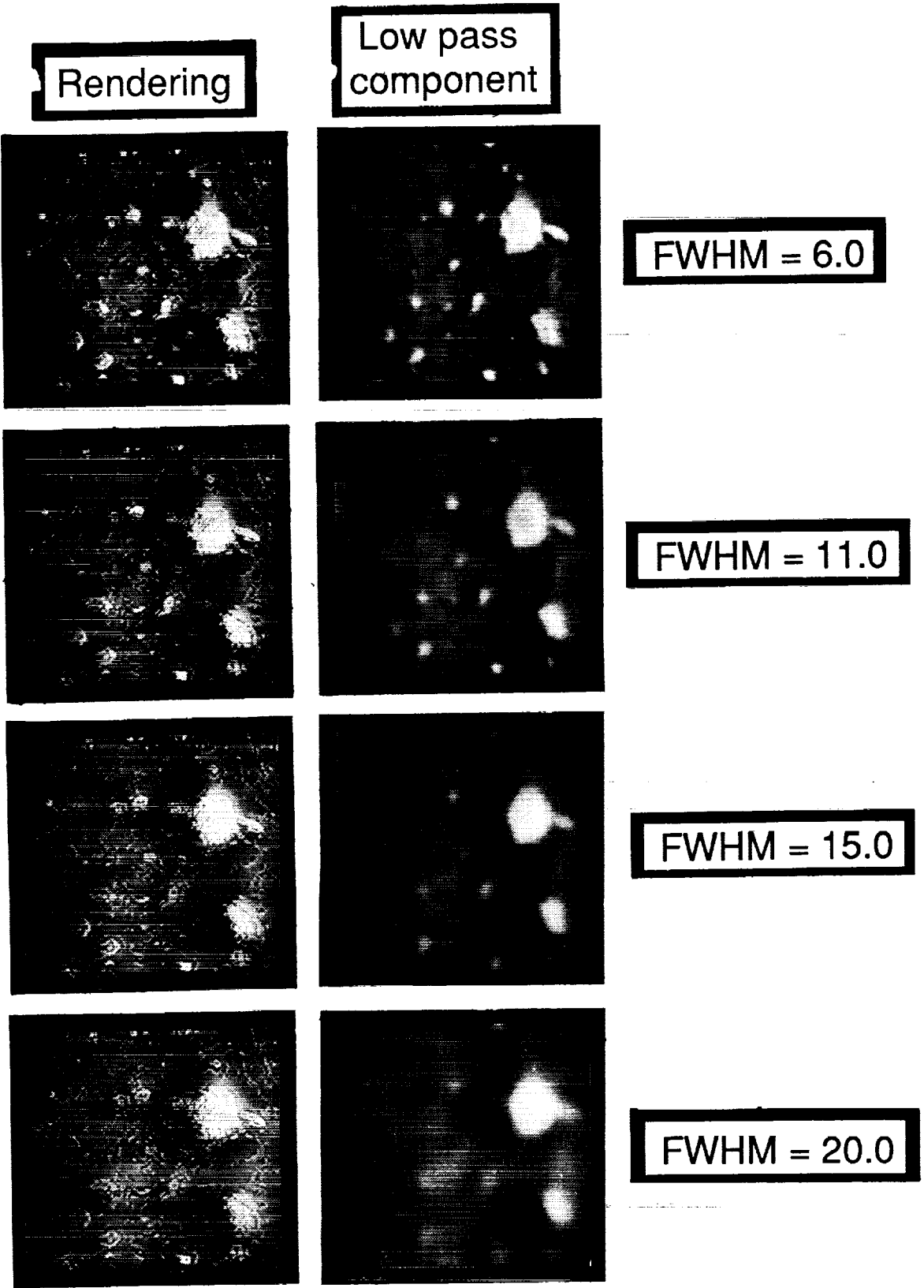
FWHM = 15.0

FWHM = 20.0

b) Mars Surface

Figure 7. (Concluded)
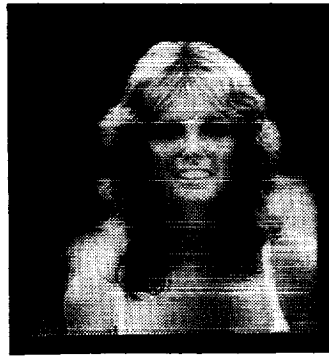
FWHM = 6.0

FWHM = 11.0
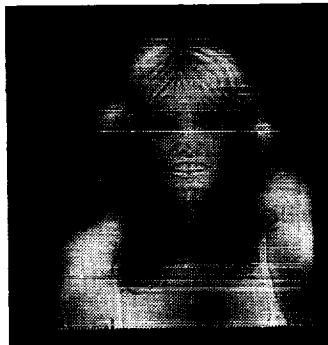
FWHM = 15.0
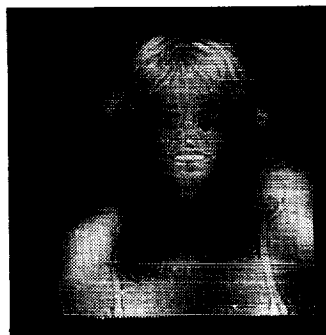
FWHM = 20.0

Figure 8. Rendering Results Using CRAIK Method
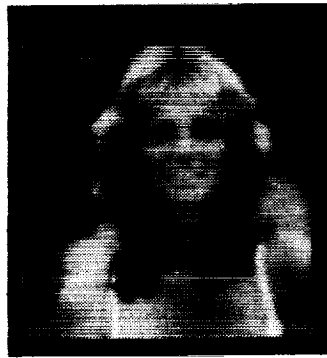
**FWHM = 6.0**

**FWHM = 11.0**
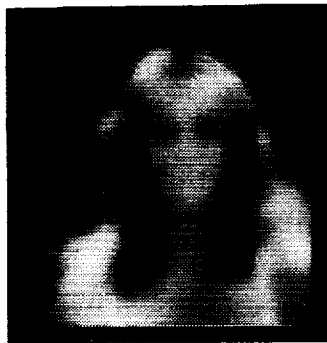
**FWHM = 15.0**

**FWHM = 20.0**

Figure 9. Rendering Results Using NARROW CRAIK.

FWHM = 6.0

FWHM = 11.0

FWHM = 15.0

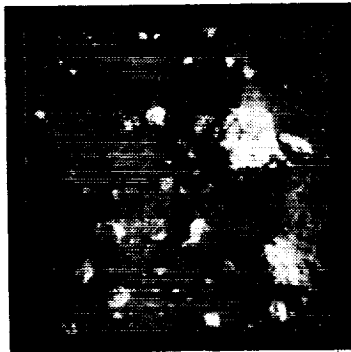FWHM = 20.0

Figure 10. Rendering Results Using STEP.
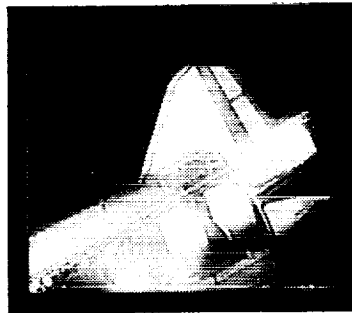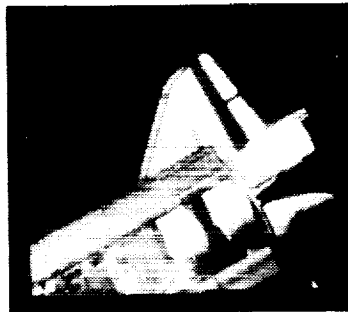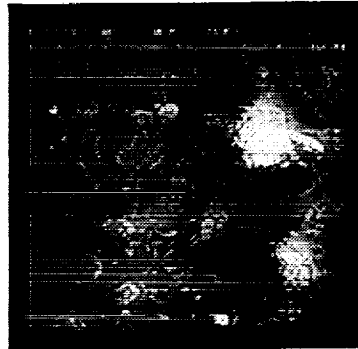
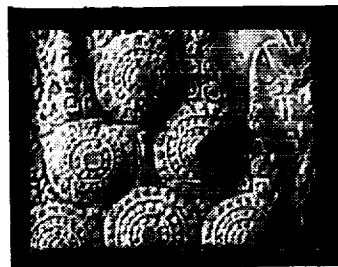Figure 11. Rendering Diverse Images Using TAPER and FWHM ≈ 20.

Figure 11. (Concluded)
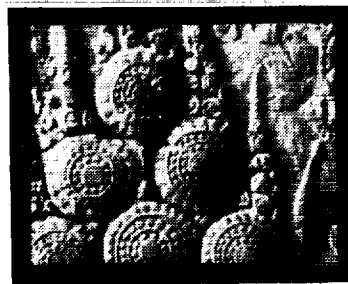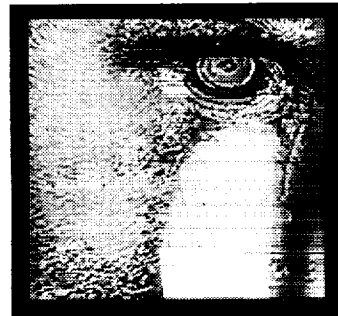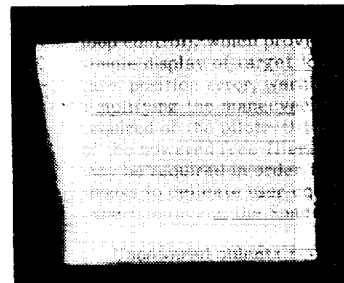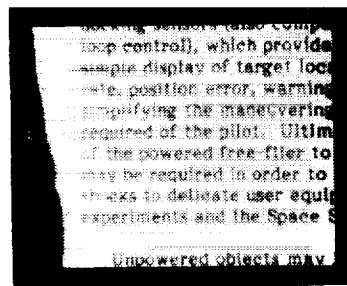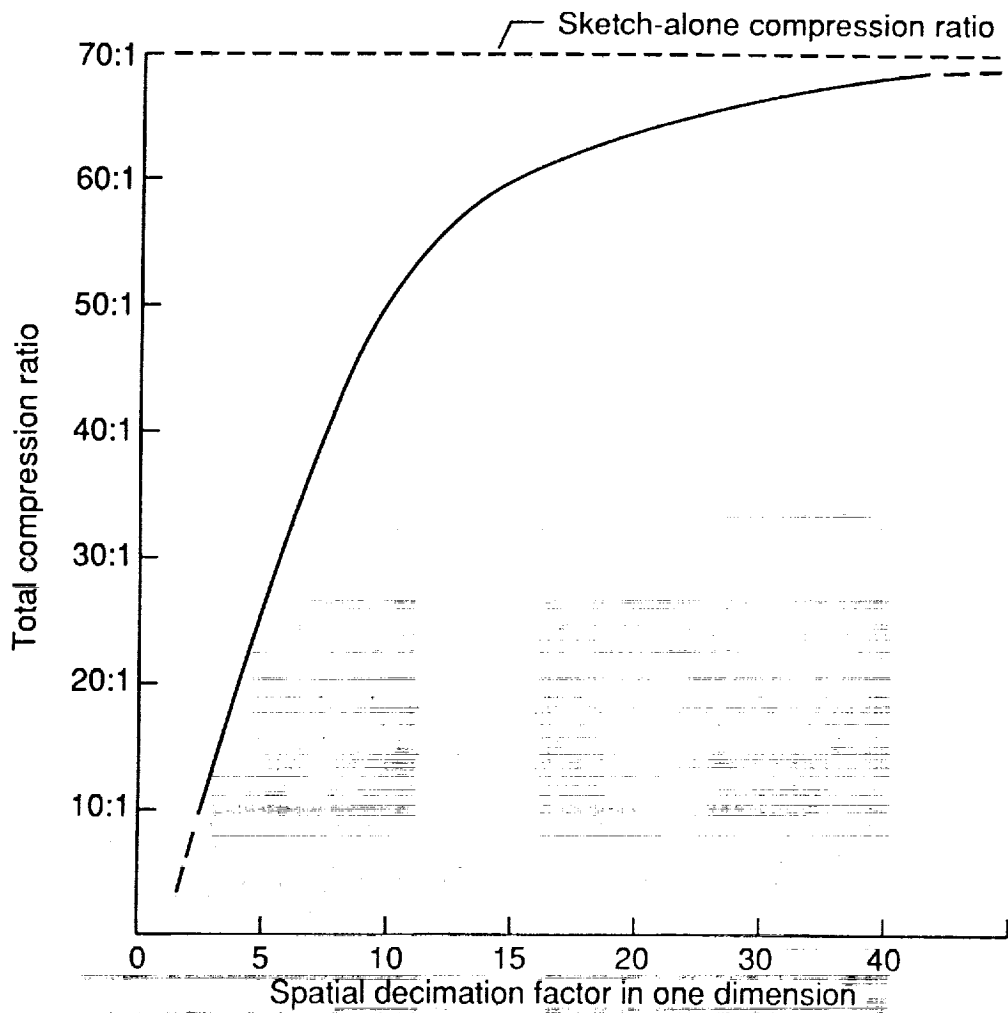
Figure 12.  Compression Ratio as a Function of Spatial Decimation of Low Pass Image
(for a 10,000 Edge Element Typical Image)

# NASA

National Aeronautics and
Space Administration

# Report Documentation Page

| 1. Report No.<br>NASA TM-104051 | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 4. Title and Subtitle<br><br>MODES OF VISUAL RECOGNITION AND PERCEPTUALLY RELEVANT SKETCH-BASED CODING FOR IMAGES | | 5. Report Date<br>February 1991 |
| | | 6. Performing Organization Code |
| 7. Author(s)<br><br>Daniel J. Jobson | | 8. Performing Organization Report No. |
| | | 10. Work Unit No.<br>307-51-10 |
| 9. Performing Organization Name and Address<br>NASA Langley Research Center<br>Hampton, VA 23665-5225 | | 11. Contract or Grant No. |
| | | 13. Type of Report and Period Covered<br>Technical Memorandum |
| 12. Sponsoring Agency Name and Address<br>National Aeronautics and Space Administration<br>Washington, DC 20546 | | 14. Sponsoring Agency Code |

15. Supplementary Notes

16. Abstract

    A review of visual recognition studies is used to define two levels of information requirements. These two levels are related to two primary subdivisions of the spatial frequency domain of images and reflect two distinctly different physical properties of arbitrary scenes. In particular, pathologies in recognition due to cerebral dysfunction point to a complete split into two major types of processing- high spatial frequency edge-based recognition versus low spatial frequency lightness- (and color-) based recognition. The former is more central and general while the latter is more specific and is necessary for certain special tasks. The two modes of recognition can also be distinguished on the basis of physical scene properties - the highly localized edges associated with reflectance and sharp topographic transitions versus smooth topographic undulation. The extreme case of heavily abstracted images is pursued to gain an understanding of the minimal information required to support both modes of recognition. Here the intention is to define the semantic core of visual information and methods for rendering this core after high degrees of data compression and transmission. This central core of processing can then be fleshed out with additional image information and coding and rendering techniques to create more accurate representations as desired or needed.

| 17. Key Words (Suggested by Author(s))<br><br>Visual Recognition, Image Coding, Spatial Frequency Bands | 18. Distribution Statement<br><br>Unclassified - Unlimited<br><br>Subject Category 63 |
|---|---|

| 19. Security Classif. (of this report)<br>Unclassified | 20. Security Classif. (of this page)<br>Unclassified | 21. No. of pages<br>23 | 22. Price<br>A03 |
|---|---|---|---|

NASA FORM 1626 OCT 86