# REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0571



| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | 1 September 2000 | Final - 15 November 1997 - 14 November 2000 |

**4. TITLE AND SUBTITLE**
Early Scene Analysis: Rapid Procession of Contours, Surfaces and Objects in Human Vision

**5. FUNDING NUMBERS**
F49620-98-1-0122

**6. AUTHOR(S)**
Dr. Patrick Cavanagh
Psychology Department

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Harvard University
33 Kirkland Street
Cambridge, MA 02138

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
AFOSR/NL
801 North Randolph Street, Room 732
Arlington, VA 22203-1977

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION AVAILABILITY STATEMENT**
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

Peter Murphy and I have created an image analysis program which labels image discontinuities as either occlusions, shadow, or pigment boundaries. The algorithm relies on regularities in the brightness flow surrounding each different contour type--each of contour has a characteristic form of isophots in its neighborhood. For example, the isophotes of a surface lie parallel to an occluding boundary on the side of the nearer surface and have no particular relation to the boundary on the occluded surface. The isophotes of a surface are collinear across any pigment change, independently of the orientation. For some ambient lights though not all, the same relation holds for the isophotes across a cast shadow boundary. Highlights and attached shadows have additional characteristics forms. Due to Peter's untimely loss to a lucrative computer science job, this project is on hold.

20001106 053

**14. SUBJECT TERMS**
HUMAN VISION

**15. NUMBER OF PAGES**
7

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclas | Unclas | Unclas | |

Report AFOSR-F49620-98-1-0122

# EARLY SCENE ANALYSIS: RAPID PROCESSING OF CONTOURS, SURFACES, AND OBJECTS IN HUMAN VISION

Patrick Cavanagh
Psychology Department
Harvard University
33 Kirkland Street
Cambridge, MA 02138

1 September 2000

Final Technical Report for Period 1 September, 1997 - 15 November, 2000

Prepared for

LIFE SCIENCES DIRECTORATE
Bldg 410
Bolling AFB, DC 20332-6448

# EARLY SCENE ANALYSIS: RAPID PROCESSING OF CONTOURS, SURFACES, AND OBJECTS IN HUMAN VISION

## Summary

During the past grant period, we have reached several goals. We have developed a novel theory of object representation in the visual system. We have examined how objects are seen to move and change shape as new parts are added. We have shown how shadows are interpreted from sparse image data and developed algorithms for identifying object and shadow borders in camera images. We have analyzed artists' techniques for clues to efficient but non-veridical representations of 3D shape in human vision. Overall, these initiatives have filled in many missing details of early scene analysis. The work has been published in 21 articles and 12 conference papers. Two projects remain to be completed.

## Objectives

How does the human brain recognize objects so rapidly even when camouflaged in partial shadow? Recognition is most often described as a sequence which starts by identifying the parts, their relations, and then finally the object. In contrast, we have demonstrated a striking short-cut: a direct contact with memory, using 2D views and occurring prior to any part-based analysis. The advantage is not only the speed, but also the ability to overcome the often intractable problems caused by shadows. Seemingly unimportant to us as observers, shadow contours are the bane of any contour recovery scheme — it is very difficult to know which contours are meaningless shadow contours and which are critical object contours. Much of our research has concentrated on images where the shadow problem is intractable and yet human vision is undeterred (2-tone images). Our direct identification process resolves the problem and experiments show that this approach only works, in human observers, for familiar objects. We have also addressed the identification of unfamiliar objects camouflaged by shadows by identifying characteristic properties of shadows in natural images. We have catalogued those properties (as well as those of attached shadows and highlights, self-occluding and pigment edges) and shown how they can be recovered from an analysis of the "isophote" field of the image (the lines joining points of identical brightness). Once we have identified and discounted the shadow contours, the remaining contours permit a reconstruction of object shape.

## Status of Effort

With two exceptions, all the proposed projects of the grant are completed. During the grant period, based on work funded by this grant we have published 21 articles and presented 12 conference papers. Details are given in the next section.

## Accomplishments / New Findings

**Image analysis: line labeling.** Peter Murphy and I have created an image analysis program which labels image discontinuities as either occlusions, shadow, or pigment boundaries. The algorithm relies on regularities in the brightness flow surrounding each different contour type -- each type of contour has a characteristic form of isophots in its neighborhood. For example, the isophotes of a surface lie parallel to an

occluding boundary on the side of the nearer surface and hae no particular relation to the boundary on the occluded surface. The isophotes of a surface are collinear across any pigment change, independently of the orientation. For some ambient lights though not all, the same relation holds for the isophotes across a cast shadow boundary. Highlights and attached shadows have additional characteristic forms. A manuscript in preparation describes this work. Due to Peter's untimely loss to a lucrative computer science job, this project is on hold.

**What art tells us about the brain.** Artists have been the pioneers of visual science for 40,000 years, discovering techniques of representation that provide compelling impressions of surfaces, light, and shadow. Many of these techniques work because they reproduce the structure of light from the original scene. However, artists can also convey three-dimensional structure with representations that never occur in real world scenes, These work because they tap the internal codes of the vision, exploiting the shortcuts and backdoors of its architecture. Examination of these techniques shows that vision is much simpler than previously thought and, moreover, that artists have been responsible for identifying all of these simple properties of vision because they permit effective art with a minimum of effort. Although this work was not originally outlined in the grant proposal, it has fit in very well with our analysis of shadows and shading. This analysis of vision and art has been published in a chapter (Cavanagh, 1999) and was part of an exchange of comments in Science (Cavanagh & Kennedy, 2000).

**Motion extrapolation, position distortion.** When a target is briefly flashed beside a moving object, the flash appears to trail behind the object. Recent articles have suggested that the perceived location of a moving item is assigned ahead of its sensed location to compensate for the continued motion of the object during the inevitable delays of processing prior to perceiving the object. David Whitney showed that the effect is based on latency differences. He published two notes, one in *Nature Neuroscience* (Whitney & Murakami, 1998) and one in Science (Whitney & Cavanagh, 2000), and two articles in *Vision Research* (Whitney, Murakami, & Cavanagh, 2000a, 2000b). He followed this up with a discovery of a novel distorting effect of motion on the apparent position of distant, stationary targets. This was just published in *Nature Neuroscience* (Whitney & Cavanagh, 2000). This work was also the subject of five conference presentations.

**Object recognition: positive priming.** In our model, recognition starts with an initial, crude 2-D match that selects a "best" prototype to explain the image data. This is followed by more sophisticated 3-D analyses to complete the recognition process. Our first experiment showed a priming effect of contours in recognition even though the contours alone were uniformative for the task. David Whitney, supported by AASERT, has extended this to priming of gender recognition in images. The results again support our early match model. An undergraduate, Susan Murunga, completed a classical conditioning study of 2 -tones images and their outlines. She examined whether there is any preconscious identification of the outlines which can trigger a skin conductance response even though the subject is unaware that the outline is related to a face. She submitted this work as her senior thesis. David Whitney's gender priming experiment and a related size change priming experiment need to be finished and analyzed and the manuscript published before this project is completed.

**Object recovery from 2 tone images.** Dr. Moore was an AFOSR supported postdoc until September 1997. Her work on simple and complex objects depicted in 2 tone images was completed before she left and revisions on the manuscript continued into the current grant period. It appeared in *Cognition* in 1998. She imaged single, simple shapes with direct lighting that produces sharp cast shadows. Her first observation is that

no single, simple shape is recognizable in a two-tone format (high-contrast, black and white). Complex objects like faces are fully interpretable in 2 tone versions but the very same parts rearranged into an unfamiliar shape loses its three-dimensionality when presented as a 2 tone image. Evidently, familiarity is a requirement for interpretation, arguing strongly against any part-based approach. This project is completed.

**Priming 2 tone images with gray-scale images.** Dr. Moore also began a study of the effects of a preview of a gray-scale version on the interpretation of a 2-tone image. Without the preview, the 2-tone images were seldom seen as the original object. The preview prime was the same object in the same or different view, or the same or different lighting. The goal was to determine whether the internal representation which was being primed was object centered or viewer centered and whether it was illumination dependent. Several classes of objects were used: familiar objects, unfamiliar objects, simple and mulitpart objects. This work was completed before she left and analyzed after taking up her new position at UCLA. The work was presented at the 1998 meeting of ARVO and a manuscript is in preparation. Once the manuscript is published, the project will be completed.

**Object models and motion perception.** Peter Tse (AASERT supported) developed a new theory for apparent motion that relies on parsing each scene into objects before matching takes place. The novel aspect of the work is that the shapes in the first frame of the motion sequence overlap spatially with those in the following frame. This enables Peter to test for principles of shape parsing (continuity, surface similarity, contiguity) that do not come into play in standard apparent motion where the shapes do not overlap. This give us a new tool for understanding image segmentation. This work is published in a recent book (Tse, Cavanagh, & Nakayama, 1998) and a related study was published in *Cognition* this year (Tse & Cavanagh, 2000). This project is completed.

**Theory of volume.** Peter Tse and Marc Albert (grant supported in 1999) developed a new theory for the level at which objects are represented in understanding visual scenes. This work is exceptionally novel and important. Rather than depending on relations between image contours or inferring surfaces, Peter shows that the underlying mode of representation is one of volumes or occupied space. Several critical demonstrations show that his formulation accounts for the broad range of image interpretations whereas representations of objects by their contours or surfaces fail. During the period of this grant, he published five papers on this topic (Albert & Tse, 1999; Tse, 1999a, 1999b; Tse & Albert, 1998; Tse, 1998) and present five talks (Tse, 1998a; Tse 1998b; Tse & Albert, 1998; Tse, 1997a; Tse, 1997b). This project is completed.

**Shape distortions.** The shape of a briefly presented test can be influenced by the shape of a preceding cue. Satoru Suzuki (AASERT supported in previous grant) and I used this to develop a completely new paradigm which can catalog the dimensions of shape. A briefly presented line target will make a subsequent circle appear elliptical with the major axis orthogonal to the orientation of the line. This distortion appears to be global — its effect is largely indepenent of the offset between the line and the circle. Equivalent interactions are found between curved lines and straight lines and between trapezoids and squares. Our current hypothesis is that the distortion arises in shape-specific units (perhaps in inferotemporal cortex) which mutually inhibit each other. This work was published in*JEP:HPP*. This project is completed.

**Personnel supported**

Personnel on the grant from 1997 to 2000 were myself, Peter Murphy (consulting Research Associate, 1997-1999), Marc Albert (Postdoc 1999, 25% salary), Raynald Comtois, our Senior Systems Analyst (25% salary), and Seth Hamlin, our research assistant (25% salary). AASERT supported students are listed in the accompanying AASERT report.

**Publications during grant period supported by grant (* new, copies appended)**

*Cavanagh, P, & Kennedy, J. M. (2000). Close Encounters: Details Veto Depth from Shadows. *Science (Letters)*, **287**, 2421.

*Tse, P. U., & Cavanagh, P. (2000). Chinese and Westerners see opposite apparent motions in a kanji stimulus. *Cognition*, **74**, B27-B32.

*Whitney, D., Cavanagh, P. (2000). The position of moving objects. *Science (Technical Comments)*, **289**, 1107.

*Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research*, **40**, 137-149.

*Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote stationary objects. *Nature Neuroscience*, **3**, 954-959.

*Whitney, D., Murakami, I., & Cavanagh, P. (2000). Temporal facilitation for moving stimuli is independent of changes in direction. *Vision Research*, in press.

Albert, M., & Tse, P. U. (1999). The role of surface attachment in perceived volumetric shape. *Perception*. **29**, 303-312.

Cavanagh, P. (1999). The Cognitive Impenetrability of Cognition: Commentary on Pylyshyn. *Behavioral and Brian Sciences* (1999) **22**, 370-371. (NR)

Cavanagh, P. (1999). Pictorial art and vision. In Robert A. Wilson and Frank C. Keil (Eds.), *MIT Encyclopedia of Cognitive Science*, (pp. 648-651) Cambridge, MA: MIT Press.

Cavanagh, P. (1999). Top-down processing in vision. In Robert A. Wilson and Frank C. Keil (Eds.), *MIT Encyclopedia of Cognitive Science*, (pp. 844-845) Cambridge, MA: MIT Press.

Holcombe, A. O., Macknik, S., Intriligator, J., Seiffert, A. E., Tse, P. U. (1999). Wakes and spokes: new motion-induced brightness illusions. *Perception*, **28**, 1231-1242.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, **39**, 37-68.

Tse, P. U. (1999). Complete mergeability and amodal completion. *Acta Psychologica*, **102**, 165-201.

Whitney, D., Murakami, I., & Cavanagh, P. (1999). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research*, **40**, 137-149.

Moore, C., & Cavanagh, P. (1998). Recovery of 3D volume from 2-tone images of novel objects. *Cognition*, **67**, 45-71.

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1315-1341.

Tse, P. U. & Albert, M. (1998). Amodal completion in absence of image tangent discontinuities. *Perception*, **27**, 455-464.

Tse, P. U. (1998). Illusory volumes from conformation. *Perception*, **27**, 977-92.

Tse, P. U., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. In Takeo Watanabe (ed.), *High level motion processing*. (pp. 249-266), Cambridge, MA: MIT Press.

Whitney, D., & Murakami, I.(1998). Latency difference, not spatial extrapolation. *Nature Neuroscience (Scientific Correspondence)*, **1**, 656-657.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance*, **23**, 443-463.

**Interactions, conference papers during grant period supported by grant**

Whitney, D. V., & Cavanagh, P. (2000). Motion adaptation shifts the apparent positions of remote objects. *Perception*, **29 Suppl.**, 80. (A)

Whitney, D., Cavanagh, P. (2000). Motion distorts visual space: Shifting the perceived position of remote stationary objects. *Investigative Ophthalmology & Visual Science*, **41**, S741.

Whitney, D., Murakami, I., & Cavanagh, P. (1999). Moving stimuli have latency advantage over flashed stimuli: evidence for a temporal facilitation mechanism. *Investigative Ophthalmology & Visual Science*, **40**, S744. (A)

Whitney, D., Murakami, I., & Cavanagh, P. (1999). Persistence does not influence the perceived location of a flash relative to a moving stimulus. *Perception*, **28 Suppl.**, 81. (A)

Holcombe, A. O., Tse, P., Macknik, S., Seiffert, A. E., & Intriligator, J. (1998). Wakes: a new motion-induced brightness illusion. *Investigative Ophthalmology & Visual Science*, **39**, S671

Moore, C., Engel, S. A., & Cavanagh, P. (1998). Seeing what you expect to see: perceptual organization of 2-tone images of novel objects. *Investigative Ophthalmology & Visual Science*, **39**, S1113.

Tse, P. U. (1998a). Volume relatability, volume completion, and illusory volumes. *Investigative Ophthalmology & Visual Science*, **39**, S1132.

Tse, P. U, & Albert, M. (1998). Amodal completion in the absence of image tangent discontinuities. ECVP, Oxford University.

Tse, P. U. (1998b). Volume Completion. Eastern Psychological Association, February, Boston.

Whitney, D., Murakami, I., & Cavanagh, P. (1998). Motion extrapolation cannot account for apparent position offset of a flashed disk relative to unpredictable motion. *Investigative Ophthalmology & Visual Science*, **39**, S1075

Tse, P. U. (1997a). Plasmas: A new class of motion-induced brightness illusions. ECVP, Helsinki.

Tse, P. U. (1997b). Volume Completion. OPAM, Psychonomics Meeting, November, Philadelphia.

**New discoveries, inventions, or patent disclosures**

Discoveries are reported in the progress section above. There were no inventions or patent disclosures during the grant period.

**Honors / Awards**

None.