AFOSR-TR-95

REPORT DOCUMENTATION PAGE

0545

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED
FINAL 15 Apr 90 To 14 Feb 95

4. TITLE AND SUBTITLE
PSYCHOPHYSICAL ANALYSES OF PERCEPTUAL REPRESENTATIONS

5. FUNDING NUMBERS
AFOSR-90-0274
61103D
3484/HS

6. AUTHOR(S)
Dr Irving Biederman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Minnesota
Dept of Psychology
75 East River Road
Minneapolis, MN 55455

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NL
110 Duncan Ave Suite B115
Bolling AFB DC 20332-0001
Dr John F. Tangney

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)*

This report is divided into two parts. The first part describes studies done at the University of Minnesota. The second part describes studies done at the University of Southern California. In both cases, full lists of citations are given to work supported in full or in part by this grant. Because most of these projects have been described in detail in previous reports, the purpose of this final report is to provide summary of the many studies and a complete list of citations. At both Minnesota and USC, the research focused on linking early sensory representations to higher-level perceptual representations. For this reason, we refer to our Center informally as the "Middle Kingdon." Studies outlined below have examined the sensory/perceptual "middle ground" in object recognition, depth perception, reading, and auditory perception.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

FINAL TECHNICAL REPORT

Submitted to: Air Force Office of Scientific Research
AFOSR/NL
Building 410
Bolling A.F.B., D.C. 20332-6448

Dr. John Tangney, Program Manager

Title: *PSYCHOPHYSICAL ANALYSES OF PERCEPTUAL REPRESENTATIONS*

Grant No. AFOSR-90-0274

Period: 1990-1995

Submitting Organization: University of Minnesota
Office of Res. & Technology Transfer Admin.
Suite 201
1100 Washington Ave. S.
Minneapolis, MN 55415

PI:[1] Professor Gordon E. Legge
(612) 625-0846

Gordon E. Legge
023-42-1595

Date of Report: 6 July 1995

[1]In the 1992-93 grant year, the grant was amended in order to subcontract funds to the University of Southern California for research under the direction of Professor Irving Biederman.

FINAL TECHNICAL REPORT

INTRODUCTION

This report is divided into two parts. The first part describes studies done at the University of Minnesota. The second part describes studies done at the University of Southern California. In both cases, full lists of citations are given to work supported in full or in part by this grant. Because most of these projects have been described in detail in previous reports, the purpose of this final report is to provide summary of the many studies and a complete list of citations.

At both Minnesota and USC, the research focused on linking early sensory representations to higher-level perceptual representations. For this reason, we refer to our Center informally as the "Middle Kingdom." Studies outlined below have examined the sensory/perceptual "middle ground" in object recognition, depth perception, reading, and auditory perception.

Research at the University of Minnesota

Surfaces Segmentation and Representation

The human visual system partitions or "explains" the image in terms of its causes--namely the shapes, materials, lighting, and depth relations. In particular, the visual system links together information from distant parts of the image that are likely to belong to the same surface, even if part of the surface is covered by another. This is the problem of occluding surfaces. The linkage is based on a number of "clues" that include colinearity of bounding edges as well as similar colors or textures of the internal regions. In addition, the visual system makes implicit a priori assumptions about the smoothness of surfaces. We have developed a Bayesian computational scheme called multi-layer segmentation that can solve occlusion (Kersten, D. & Madarasmi, S., 1995; Madarasmi, S., Kersten, D., & Pong, T. C., 1993; Madarasmi S., Kersten, D., & Pong, T.C., 1993).

Object Recognition and Classification for Human and Ideal Observers

We developed a novel paradigm for experimental studies of human object recognition (Liu, Knill and Kersten, 1992). It was based on the use of ideal observer theory to estimate the statistical efficiency with which human subjects use stimulus information for performing a recognition task. We measured the statistical efficiency with which human observers made simple classification judgments of randomly shaped thick wire objects. We were able to show that human performance exceeded that of an ideal 2D template matching strategy, effectively eliminating the class of 2D template matching models as candidates for explaining the data. We also showed viewpoint dependent effects in that subjects' efficiencies were higher for learned views of objects than novel views, but that the effect decreased with increasing structure (e.g. symmetry, planarity) of the objects. Moreover, average efficiencies across all viewpoints increased with increasinmg regularity of objects, indicating that the visual system takes advantage of such regularities in storing object information and comparing it with imagae data for recognition. We also compared human performance with a class of computational models which have been proposed for object recognition known as Hyper Basis Function models. This involved the development of computer implementations of the models and their testing on the same task given to human subjects. We were able to show that even allowing for considerable generalizations of the models, human performance exceeded that of the models, eliminating them as candidate explanations for the results.

The Perception of Spatial Layout from Shadows

This project had two main aspects: psychophysical studies of the role played by cast shadows in the perception of 3D object motion and a theoretical analysis of the geometry of shadow boundaries as they appear in images and are related to the shapes of smooth surfaces. I will briefly describe the latter, as Dan Kersten should have previously described the work on 3D object motion. Drawing on previous work on smooth occluding contours, we analyzed the local geometric structure of shadow contours on smooth and piece-wise smooth surfaces. Particular attention was payed to intrinsic shadows on a surface; that is, shadows created on a surface by the surface's own shape and placement relative to a light source. We derived the invariants relating surface shape to the shapes and singularities of bounding contours of such shadow contours, including the singularities in the evolution of shadows on a surface as it is moved relative to a light source. We showed that the results obtained for point sources of light generalize in a straightforward way to extended light sources, under the assumption that light sources are convex. The results play much the same role for understanding the information provided by shadows as Koenderink's work on occluding contours plays for understanding the information provided by an object's silhouette; namely, they provide the geometric underpinnings on which more applied work can be based.

Workshop on Visual Perception: Computation and Psychophysics

Over the course of the Summer and Fall of 1992, David Knill planned and organized a workshop on computational and psychophysical approaches to visual perception which was held in Chatham Bar, Massachusetts Jan. 14-17. The workshop brought together researchers in computational vision and psychophysics to discuss ways of conceptualizing and modeling problems in visual perception. The workshop was a tremendous success and has resulted in the publication of a book with contributions from the participants by Cambridge University Press ("Perception as Bayesian Inference, to appear in 1995). The costs of the workshop itself were paid by a seperate grant from the AFOSR (AF/F49620-93-1-0124).

The Role of Color in Object Recognition

Does color improve object recognition? If so, is the improvement greater for blurred images where there is less shape information? Do people with low visual acuity benefit more from color than people with normal acuity? Wurm, Legge, Isenberg & LaMay (1992) addressed these questions in three experiments by comparing naming reaction times (RTs) for food objects displayed in four ways: achromatic or color, and blurred or unblurred. Normally sighted subjects had faster reaction-times with color that did not change significantly with blur. Low-vision subjects were also faster with color and the difference did not depend significantly on acuity. In two additional experiments, we asked if the faster RTs for color stimuli were related to objects' prototypicality or color diagnosticity. We conclude that color does improve object recognition and the mechanism is probably sensory rather than cognitive in origin.

Psychophysics of Complex Auditory Signals

The focus of the work during the grant period was on temporal aspects of auditory perception with an emphasis on the detection and discrimination of modulation. Amplitude and frequency modulation exist in almost all complex auditory signals. In Edwards and Viemeister (1994b) we showed that under certain conditions FM is encoded as AM; we also delineated the conditions under which a second coding mechanism for FM is used. Edwards and Viemeister (1994a) examined when FM, a relatively complicated signal, can be approximated by quasi FM, a simple 3-component signal that offers considerable advantages over FM in psychoacoustic studies. Viemeister et al. (1992) pursued a "multiple look" model for temporal processing and showed that at the level of the auditory nerve it can provide an account of temporal integration equivalent to simple spike summation. Chang and Viemeister (1991) extended the often-used probe technique to the temporal domain and showed that the temporal window measured this way was surprising broad.

Statistical Efficiency for Categorization of Curvature

To assess the manner in which curvature is internally represented by the visual system, we measured the statistical efficiency with which observers were able to classify circular arcs into pretrained curvature categories (Mansfield, Biederman, Legge, and Knill, 1991). We found that highest efficiencies were obtained when one of the curvature categories included zero curvature (i.e., a straight line). These efficiencies are not accounted for by differences in the discriminability of curvature. This outcome is consistent with the notion that viewpoint invariance differences are exploited in object recognition.

The Role of Font Information in Reading

What is the role of font in reading? Reading is usually effortless despite the wide range of different fonts found in modern-day printed text. Is font "transparent" to the reading process or does it play an explicit role? Occasionally words written in a b old or italic font "pop-out" from a page of text, which suggests that font may be involved in a global analysis of the page of text. We measured the perceptual similarity of pairs of fonts using a reaction-time font-discrimination task (Klitz, Mansfield, and Legge, 1992). Subjects indicated whether a page of text was rendered in only one font, or whether there was a region of text rendered in a different font than the background. For some paris of fonts the detection reaction time was very fast irrespective of the size of the target region. This result is consistent with the use of font information in rapid global analysis of the page of text.

In a second study we measured reading speed for text passages printed in either a single font, or in mixtures or two or more fonts (Klitz, Mansfield, and Legge, 1995). We found that reading speeds with font mixtures were slower than the average of the reading speeds for the component fonts alone, indicating that there was a reading-speed "cost" involved in reading texts with multiple fonts. This result argues against a font-generic reading mechanism extracting letter information without explicit access to the details of the font (e.g., using font-invariant features). Instead, letters may be recognized via a font-specific mechanism that tunes the reading process appropriately for the font being read.

Mr. Chips: An Ideal Observer Model of Reading

Existing models of reading do not explicitly specify how visual data are combined with other sources of information, nor do they explain how visual disorders affect reading. Ideal-observer models have been useful in vision because they are explicit in identifying sources of information and task constraints. The perceptual component of reading can be formalized as the interpretation of a string of stimulus symbols (text), sampled through a window whose position is determined by a sequence of saccades. An ideal reader can be defined that accurately interprets the text in the minimum number of saccades. Its computation uses three sources of information: 1) visual data, normally a few recognized letters in central vision and the locations of spaces in the periphery; 2) lexical data, including allowable words and their probabilities; and 3) eye-movement data, including distribution of saccade lengths.

Results from a computer simulation of the ideal reader may be informative about human readers. For example, the ideal reader exhibits regressive saccades (which also occur in human reading but are usually regarded as "errors") because ideal saccades of greatest expected length occasionally result in ambiguous interpretation of text. The ideal reader with scotomas has more regressions than normal and erratic eye movements (much larger standard deviation of saccade lengths), a pattern like that reported for some patients with central-field loss. The ideal reader is an explicit model for the combination of visual and other sources of information in reading. Its performance with abnormal retinal data may help us to understand the adverse effects of visual-field loss on human reading.

Results have been presented at two conferences (Legge, 1992; Klitzt & Legge, 1994) and a major paper is currently in preparation.

Binocular Visual Direction

How is the binocular visual direction of a feature in depth determined from the views seen by the left and right eyes? According to the geometry of binocular vision, binocular visual direction ought to be the average of the directions to the feature from the left and right eye. However, we have shown that if the views seen by the two eyes have different contrasts then the perceived direction of the feature is closer to the direction from the eye with higher contrast (Mansfield, Akutsu, and Legge, 1992; Mansfield and Legge, 1995b). We have proposed a new model in which binocular visual direction is the "most-likely" visual direction given the monocular direction signals and their associated directional uncertainties (Mansfield and Legge, 1993a;

Mansfield and Legge, 1995b). Our new model also makes predictions for the accuracy of judgements of binocular direction (binocular vernier acuity). However, the model fails to predict binocular vernier acuity when the the elements in the vernier target are very close to each other (Mansfield and Legge, 1993b).

The shift in binocular visual direction observed when the monocular images have different contrasts is consistent with the 'Cyclopean eye' being located closer to the eye seeing the higher contrast image (Mansfield and Legge, 1995 a,b). This observation challenges the currently accepted premise that, in binocular vision, the world is perceived as if from a single viewpoint located midway between the left and right eyes (the so-called 'Cyclopen eye'). We devised a new stimulus in which two features in depth have different interocular contrast ratios. By simultaneously measuring the visual direction of these features we have shown that the scene is perceived as if viewed from two spatially-separated Cyclopean eyes (Mansfield and Legge, 1995a). This observation refutes the notion of a single Cyclopean eye in binocular vision, and suggests an alternative scheme by which binocular vision is used to determine spatial layout. Instead of assessing visual direction globally from a single point, we have proposed that a global percept of scene layout is built-up from multiple local estimates of relative visual direction between pairs or clusters or neighboring features. In this way, the perception of visual direction is similar to the local and global stages proposed for the perception of depth from binocular stereopsis.

Citations: University of Minnesota

Publications

Edwards, B. W. and Viemeister, N.F. (1994a). Psychoacoustic equivalence of frequency modulation and quasi-frequency modulation. *J. Acoust. Soc. Am., 95*, 1510-1513.

Edwards, B. W. and Viemeister, N.F. (1994b). Modulation detection and discrimination with three-component signals. *J. Acoust. Soc. Am., 95,* 2202-2212.

Kersten D. & Madarasmi S. (1995) The visual perception of surfaces, their properties, and relationships. In *Partitioning Data Sets: With applications to psychology, vision and target tracking*, Edited by I. J. Cox, P. Hansen and B. Julesz, AMS.

Knill, D. C., Mamassian, P. and Kersten, D. (1993) The geometry of shadows, University of Minnesota Computer Science Technical Report, TR 93-47.

Liu Z. Knill D. C. & Kersten D. (1995). Object classification for human and ideal observers. *Vision Research, 35*, 549-568.

Madarasmi S. Kersten D. & Pong T.C. (1993) Multi-layer surface segmentation using energy minimization. *IEEE Computer Vision & Pattern Recognition.* 774-775.

Madarasmi S. Kersten D. & Pong T. C. (1993). The computation of stereo disparity for transparent and for opaque surfaces. In C. L. Giles, S. J. Hanson, & J. D. Cowan (Ed.), *Advances in Neural Information Processing Systems 5.* San Mateo, CA: Morgan Kaufmann Publishers. pp. 385-392.

Mansfield, J. S. and Legge, G. E. (1995b). The binocular computation of visual direction. *Vision Research,* in press.

Viemeister, N.F., Shivapuja, B.G., and Recio, A (1992). Physiological correlates of temporal integration. In Cazals, Y., Demany, L., and Horner, K. (Eds.) *Auditory Physiology and Perception.* Pergamon, Oxford, pp. 323-329.

Viemeister, N.F. and Plack, C.J. Time analysis. In Yost, W., Popper, A., and Fay, R. (Eds.) *Human Psychophysics* (Springer Series in Auditory Research, Vol. 3. Springer-Verlag, New York. (In press). Review chapter partially supported by AFOSR.

Wurm L.H., G.E. Legge, L.M. Isenberg & A. Luebker. Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 899-911, 1993.

Conference Presentations and Abstracts

Chang, P. and Viemeister, N.F. (1991). Temporal windows for signals presented at uncertain times. *J. Acoust. Soc. Am. 90*, 2248. (Abstract)

Klitz T.S. & Legge G.E. Modeling the visual span in reading. ARVO, Sarasota, May 1-6, 1994. *Supp. IOVS.,* 35, 1949, 1994.

Klitz, T. S., Mansfield, J. S., and Legge, G. E. (1992). Font "pop out" in text images. In *OSA Annual Meeting Technical Digest,* Optical Society of America, Washington, DC., 1992. 23, 170.

Klitz, T. S., Mansfield, J. S., and Legge, G. E. (1995). Reading speed is affected by font transitions. *Suppl. to Investigative Ophthalmology and Visual Science, 36*, S670.

Legge G.E. An ideal observer model of reading. ARVO, Sarasota, May 3-8, 1992. *Supp. IOVS., 33,* 1414, 1992.

Liu, Z., Kersten, D. & Knill, D. C. (1992). Object discrimination for human and ideal observers. *ARVO*. Sarasota, Florida. *Supplement to Investigative Ophthalmology and Visual Science, 33*, 825.

Madarasmi, S., Kersten, D. and Pong, T.C.. (1992). The computation of stereo disparity for opaque and transparent surfaces. Neural Information Processing Systems: Natural and Synthetic, Denver.

Madarasmi, S., Kersten, D., and Pong, T.C. (1992). A multi-layer approach to segmentation and interpolation with application to stereo vision. ARVO.

Madarasmi, S., Kersten, D., & Pong, T. C. (1993). Depth from occlusion using multiple surface representations. ARVO. Sarasota, Florida. *Investigative Ophthalmology and Visual Science, 34*, 1130.

Mansfield, J. S., Biederman, I., Legge, G. E., and Knill, D. C. (1991). Greater statistical efficiency for viewpoint invariant differences in the categorization of curves. In *OSA Annual Meeting Technical Digest*, (Optical Society of America, Washington, DC ., 1991.) 17, 191-192.

Mansfield, J. S., Akutsu, H. A., and Legge, G. E. (1992). Interocular contrast differences produce lateral shifts in the perceived location of binocular depth targets. *Suppl. to Investigative Ophthalmology and Visual Science, 33*, 530.

Mansfield, J. S. and Legge G. E. (1993a). Binocular computation of direction and depth. *Suppl. to Investigative Ophthalmology and Visual Science, 34*, 1053.

Mansfield, J. S. and Legge G. E. (1993b). Vernier acuity for targets having different depths. Poster presented at the NATO Workshop on Binocular Stereopsis and Optic Flow, York University, Toronto, Canada.

Mansfield, J. S. and Legge G. E. (1995a). Is there more than one cyclopean eye for binocular visual direction? *Suppl. to Investigative·Ophthalmology and Visual Science, 36,* S813.

Research: University of Southern California

The research is described in 15 papers published or in press, and four submitted. Here are the citations and an outline of findings.

1. Biederman, I., Hummel, J. E., Gerhardstein, P. C., & Cooper, E. E. (1992). From Image Edges to Geons to Viewpoint Invariant Object Models: A Neural Net Implementation. Applications of Artificial Intelligence X: Machine Vision and Robotics, 1708, 570-578, SPIE Proceedings Series.

2. Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and further evidence. Canadian Journal of Psychology, 46, 191-214.

3. Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. Psychological Review, 99, 480-517.

4. Biederman, I. (1992). Human Image Understanding. In P. Johansen & S. Olsen (Eds.) Theory and Applications of Image Analyses. (Pp. 3-14). Singapore: World Scientific.

5. Biederman, I., Hummel, J. E., Cooper, E. E., & Gerhardstein, P. C. (1993). Shape recognition in mind, brain, and machine. In P. Rudomen, M. A. Arbib, F. Cervantes-Perez, & R. Romo (Eds.) Neuroscience: Neural Networks to Artificial Intelligence (Pp. 282-293.). Berlin: Springer-Verlag.

6. Biederman, I. (1993). Geon theory as an account of shape recogntiion in mind and brain. The Irish Journal of Psychology, 14, 314-327.

7. Biederman, I., Cooper, E. E., Hummel, J. E., & Fiser, J. (1993). Geon theory as an account of shape recognition in mind, brain, and machine. In J. Illingworth (Ed.) Proceedings of the 4th British Machine Vision Conference, 1, 175-186. Surrey, Guildford, U.K.: BMVA Press

8. Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for 3D viewpoint invariance. Journal of Experimental Psychology: Human Perception and Performance, 19, 1162-1182.

Five experiments on the effects of changes of depth orientation on a) priming the naming of briefly flashed familiar objects, b) detecting individual simple volumes (geons), and c) the classification of unfamiliar objects (that could readily be decomposed into an arrangement of distinctive geons), all revealed immediate (i.e., not requiring practice) depth invariance. The results can be understood in terms of three conditions derived from a model of object recognition (Biederman, 1987; Hummel & Biederman, 1992) that have to be satisfied for immediate depth invariance: a) that the stimuli be capable of activating viewpoint invariant (e.g., geon) structural descriptions (GSDs), b) that the GSDs be distinctive (different) for each stimulus, and c) that the same GSD be activated in original and tested views. The stimuli employed in several recent experiments documenting extraordinary viewpoint dependence violated these conditions.

9. Biederman, I. (1995). Some Problems of Visual Shape Recognition to Which the Application of Clustering Mathematics Might Yield Some Potential Benefits. In I. J. Cox, P. Hansen, B. Julesz (Eds.) Partitioning Data Sets, Pp. 313-329. Providence, R. I.: American Mathematical Society.

10. Biederman, I. (1995). Geon theory as an account of shape recognition in mind, brain, and network. Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 2, 46-59.

In a fraction of a second--from a single visual fixation--humans are able to comprehend novel images of objects and scenes, often under highly degraded and novel viewing conditions. Recent research on how the brain achieves this remarkable feat suggests that objects are represented as an arrangement of simple viewpoint-invariant shape primitives, termed "geons," that serve to distinguish visual classes, so that a given image can be determined, for example, to be that of a chair, fork, or penguin. As long as two or three geons in their specified relations can be extracted from the image, entry-level classification will almost always be successful despite drastic variations in the object's silhouette and its local context. Progress on neural and neural network modeling of these capacities and their relation to face recognition are discussed.

11. Biederman, I. (1995). Visual Object Recognition. In S. Kosslyn (Ed.). Invitation to Cognitive Science, 2nd edition. MIT Press, In press.

12. Biederman, I., & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: A critical analysis. Journal of Experimental Psychology: Human Perception and Performance, in press.

Biederman and Gerhardstein (1993) proposed that a representation specifying a distinctive arrangement of viewpoint-invariant parts (geons) affords great reduction in the costs of rotation in depth relative to viewpoint dependent information and appears to be sufficient for characterizing easy shape classifications. Tarr and Bulthoff (1995) attempt to make a case for viewpoint dependent mechanisms, such as mental rotation, to explain the effects of depth rotation in shape-based object recognition. Their arguments against geon theory's account of entry level classification rest on the mistaken and unwarranted attribution that all entry level classes have to be equally distinguishable. Instead, geon theory offers an explanation of those cases where entry level classification is relatively difficult and subordinate level classification is relatively easy.

13. Fiser, J., & Biederman, I. (1995). Size invariance in visual object priming of gray scale images. Perception, in press.

14. Fiser, J., Biederman, I., & Cooper, E. E. (1995). Test of a two-layer network as a model of human entry-level object recognition. J. M. Bower (Ed.) The Neurobiology of Computation:

Procedings of the Third Annual Computational Neuroscience Meeting. Amsterdam: Kluever.

15. Biederman, I., Cooper, E. E., & Hummel, J. E. (1995). Recognition-by-Geons: 1995's Current Progress and Current Challenges. Image and Vision Computing, in press.

16. Biederman, I., & Cooper, E. E. (1995). Viewpoint invariant differences during object recognition are more salient than metric differences, Submitted for publication.

Abstract. Object images in which a single part differed in either: (a) a viewpoint invariant property (VIP), such as whether an edge is straight or curved, or (b) a viewpoint dependent Metric property, such as aspect ratio, were scaled according to a model of simple cell similarity and a physical identity judgment task. The two scaling operations allowed, for the first time, a principled test of the fundamental assumption of several recent theories of shape recognition that properties of images that are not likely to change with small variations in viewpoint in depth (viz., VIPs), receive greater weight in object recognition than properties that do change with viewpoint. Even though the scaling procedures, presumed to reflect early cortical representations, indicated that the Metric differences of these stimuli were greater than the VIPs, the reverse was true on an object classification task, presumed to reflect later cortical representations. Models that do not posit a role for VIPs thus neglect a fundamental aspect of human object representation.

17. Fiser, J., Biederman, I., & Cooper, E. E. (1995). To what extent can matching algorithms based on direct outputs of spatial filters account for human shape recognition? Submitted for publication.

18. Subramaniam, S., Biederman, I., & Cooper, E. E. (1995). Perceiving irregular objects. Submitted for publication.

Abstract. Subjects judged whether two object images, S1 and S2, presented briefly and sequentially, were or were not members of the same basic level class, e.g., both lamps. On same trials, the images could be identical or differ in a large part that was either regular or irregular. A regular change would always entail a difference in a viewpoint invariant property. for example, if the base of a lamp in S1 was a cylinder it could be a brick in S2. Irregular parts resembled "free form" sculptures, with the magnitude of the change for the irregular parts equated to the magnitude of the regular part changes according to a model of similarity based on V1-type spatial filters. A change in a regular part was much more disruptive on the speed and accuracy in judging that the two images were members of the same basic level class compared to a change in an irregular part. This result is expected from theories that posit a special status to viewpoint invariant regularities as a basis for determining 3D objects from 2D images. It was not that the irregular parts were not represented, in that a change from an irregular part to a regular part or vice versa was as disruptive to object classification as a change in a regular part. Thus the irregularities appear to be represented as irregularities, but without specification of their detailed configuration.

19. Subramaniam, S., Biederman, I., & Cowie, R. I. D. (1995). Priming the naming of impossible familiar objects. Submitted for publication.

Abstract. Subjects named possible or impossible line drawings of common objects, presented for 100 msec, on two blocks of trials. Error rates were somewhat higher for the impossible objects, but the large and reliable reduction in RTs and error rates from the first to the second blocks were equivalent for the two types of objects. Moreover, there was no effect of whether the image for an object on the first block was of the same or different type (possible or impossible) as on the second block. Rendering object images impossible has no discernible effect on name priming. Although this result confirms an empirical prediction of Schacter et al. (1991) we question whether it supports their assumption of a structural description specifying global shape and orientation.

Papers at Scientific Meetings

1. Biederman, I., Hummel, J. E., & Cooper, E. E. (1990) Human Object Recognition. Invited address presented at a Conference on Visual Information Assimilation in Man and Machine, Ann Arbor, Michigan, June.

2. Hummel, J. E., & Biederman, I. (1990). Dynamic Binding: A Basis for the Representation of Shape by Neural Networks. Paper presented at the 12th Annual Meeting of the Cognitive Science Society, Cambridge, MA. July.

3. Biederman, I, & Cooper, E. E. (1990). Intermediate, invariant representations mediate visual object recognition. Invited presentation at a Workshop on Object and Scene Perception. University of Leuven, Leuven, Belgium. September.

4. Hummel, J. E., & Biederman, I. (1990). Binding invariant shape descriptors: A neural net architecture for structural description and object recognition. Invited presentation at a Workshop on Object and Scene Perception. University of Leuven, Leuven, Belgium. September.

5. Biederman, I. (1990) Visual Image Understanding. The Fourth Annual Fern Forman Fisher Lecture, University of Kansas, November.

6. Hummel, J. E., & Biederman, I. (1990). Binding invariant Shape Descriptors for Object Recognition: A Neural Net Implementation. Paper presented at the 31th Annual Meeting of the Psychonomic Society, New Orleans, LA, November 17-19.

7. Biederman, I. (1991). How an account of Shape Recognition can be Achieved by a Neural Network that Solves the Binding Problem through Phase Locking. Invited paper presented at the Workshop on Rhythmic Oscillations in Cortex: Their Form and Function, Tucson, April.

8. Hummel, J. E., & Biederman, I. (1991). Binding by phase locked neural activity:

Implications for a theory of visual attention. Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

9. Cooper, E. E., & Biederman, I. (1991). Evidence for size invariant representations in visual object recognition. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

10. Gerhardstein, P. C., & Biederman, I. (1991). 3D Orientation invariance in visual object recognition. Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

11. Biederman, I. (1991). The neuroscience of object recognition. Invited featured speaker at The First Annual Meeting of the Canadian Society for Brain, Behavior, and Cognitive Science, Calgary, June.

12. Biederman, I. (1991). Shape recognition in eye and brain. Invited presentation at the Stockholm Workshop on Computational Vision, Rosenen, Sweden, August.

13. Biederman, I. (1991). Human Image Understanding. Invited address presented at the 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark. August.

14. Mansfield, J. S., Biederman, I., Legge, G. E., & Knill, D. C. (1991). Greater statistical efficiency for viewpoint-invariance differences in the categorization of curves. Paper presented at the Meetings of the Optical Society, San Jose: CA. November.

15. Biederman, I., Cooper, E. E., & Gerhardstein, P. C. (1991). Picture naming reveals the major invariances expected of a shape recognition system. Poster presented at the Meetings of the Psychonomic Society, San Francisco, CA. November.

16. Biederman, I. (1991). Shape recognition in mind, brain, and machine. Invited paper presented at the NSF-CONACYT sponsored Symposium on Natural and Artificial Intelligence: A Meeting Between Neuroscioence and Artificial Intelligence, Jalapa, Mexico, December.

17. Biederman, I. (1992). The neural basis of shape recognition. Invited paper presented at the Seminar on Cognitive Neuroscience at the Meetings of the AAAS, Chicago, February, 1992.

18. Biederman, I. (1992). Shape recognition in Mind and Brain. Invited paper presented to the Helmholtz Society, Irvine, California, February, 1992.

19. Biederman, I. (1992). Reverse Engineering the Psychology of Shape Recogntiiton. Invited paper presented at the Office of Naval Research Workshop on Intermediate and Higher Level VIsion. Laguna Beach, CA. March.

20. Biederman, I., & Hummel, J. E. (1992). From Image Edges to Geons to Viewpoint

Invariant Object Models: A Neural Net Implementation. Invited paper to be presented at the International Society for Optical Engineering Conference on Intelligent Information Systems, Orlando, FL. April.

21. Biederman, I., Gerhardstein, P. C., Cooper, E. E., & Nelson, C. A. (1992). High level object recognition without a temporal lobe. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

22. Biederman, I. (1992). Human Image Understanding. Invited paper presented at the Meetings of the International Congress of Psychology, Brussels, Belgium, August.

23. Biederman, I. (1992). Challenges to Machine Vision. Invited paper presented to a Workshop on Active Vision, Ruzinagaard, Denmark, August.

24. Biederman, I. (1992). Shape Recognition in Mind and Brain. Invited paper presented at a Workshop on Pattern Organization and Object Recognition, Brussels, Belgium, September,

25. Biederman, I. (1992). Shape Recognition in Mind and Brain. Invited paper presented at the National Research Council's Committee on Vision Symposium on Vision and Cognitive and Behavioral Psychology, Washington, D. C., October.

26. Biederman, I. (1992). Shape Recognition in Mind, Brain, and Machine. Invited address presented at Dedication of Computer Science Building, Heriot-Watt University, Scotland, October.

27. Biederman, I., Gerhardstein, P. C., Cooper, E. E., & Nelson, C. A. (1992). High level shape recognition without an inferior temporal lobe. Paper presented at the Annual Meeting of The Psychonomic Society, St. Louis, Mo. November.

28. Biederman, I (1993). Shape Recognition. Invited paper at a DIMACS sponsored Meeting on Mathematical Clustering and Vision. Rutgers University, April.

29. Cooper, E. E., & Biederman, I. (1993). Metric versus viewpoint-invariant shape differences in visual object recognition. Poster presented at the Annual Meeting of The Association for Researchin Vision and Ophthalmology, Sarasota, Fl. May.

30. Gerhardstein, P. C., & Biederman, I. (1993). Viewpoint invariance in recognizing unfamiliar depth-rotated objects. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.

31. Biederman, I. (1993). Shape recognition in mind and brain. The William F. Prokasy Lecture at The University of Utah. May.

32. Biederman, I. (1993). Visual object recognition in mind and brain. Invited paper

presented at a Conference on Object Representation in Visual and Haptic System, Madrid, Spain. May.

33. Biederman, I. (1993). The neural basis of shape recognition. Invited paper presented at a Symposium of Parallel Processing in the Nervous System, Toronto, Canada, July.

34. Biederman, I., Gerhardstein, P. C., Cooper, E. E., Fiser, J., & Hummel, J. E. (1993) Recognition-by-Geons: Current Progress and Current Challenges. Invited paper presented at the International Joint Conference on Artificial Intelligence, Savoie, France, September.

35. Biederman, I. (1993). Geon theory as an account of shape recognition in mind, brain, and machine. Invited address presented to the British Machine Vision Association, Surrey University, Guildford, U.K. September.

36. Biederman, I. (1993). Object recognition in mind, brain, and machine. Invited address to the Sixth Annual Meeting of the Irish Association of Aritificial Intelligence and Cognitive Science, Belfast, Northern Ireland, September.

37. Biederman, I. (1993). Can a successful face recognizer serve as a model of entry level object recognition? Paper presented at the International Conference on Face Processing, Cardiff, Wales, September.

38. Biederman, I. (1993). Grounding Mental Symbols in Object Images. Invited paper presented at the Meetings of the Psychonomics Society, Washington, D. C., November.

39. Cooper, E. E., & Biederman, I. (1993). Geon Differences During Recognition are more Salient than Metric Differences. Poster presented at the Meetings of the Psychonomics Society, Washington, D. C., November.

40. Biederman, I., Fiser, J., Cooper, E. E., & Gerhardstein, P. C. (1993). Intermediate representations and visual shape recognition. Paper presented at the Meetings of the Psychonomics Society, Washington, D. C., November.

41. Biederman, I. (1994). Shape Recognition in Mind and Brain. Invited address to a special meeting of the Japanese Cognitive Science Society, Toyko, March.

42. Biederman, I. (1994). The Neural Basis of Object Recognition. Invited presentation at a Syposium on Cerebral Cortex and Object Perception. Jerusalem, Isreal. March.

43. Biederman, I. (1994). Shape recognition in Mind and Brain. Invited paper at a Symposium on Cortex and Object Recognition, Syracuse University, April.

44. Fiser, J., Biederman, I., & Cooper, E. E. (1994). Are the direct outputs of Gabor filters sufficient for human object recognition or are they only the prior stage for intermediate

representations? Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Sarasota, Fl., May.

Recent two-stage models of human object recognition (e.g., Edelman & Weinshall, 1991; Lades, Vorbruggen, Buhmann, Lange, von der Malsburg, Wurtz, et al., 1993) map the output of a lattice of local filters directly onto an object layer thus dispensing with the intermediate representations (e.g., lines, surfaces, aspects, or simple volumes posited by multilayer models (e.g., Dickinson, Pentland, & Rosenfeld, 1992; Hummel & Biederman, 1992; Poggio & Edelman, 1990). Are these intermediate representations necessary? Methods. A model of the two-stage type (Buhmann, Lange, & von der Malsburg, 1989), a highly successful face recognition system) was evaluated by comparing its performance to that of humans in several real-time, object recognition experiments. The model's complete representation and matching strategy allow it to approach the performance limits of two-stage models. The test stimuli were the same object pictures used in the human experiments, which included images with complementary contour deletions, geon-recoverable and nonrecoverable images with the same amount of contour deletion, and mirror reflected versions. Results. Although the system was able to recognize images with relatively high accuracy, its performance did not qualitatively match that of humans. Whereas people are much better at recognizing recoverable compared to nonrecoverable images, and show no effect of mirror reflecting or complementizing images, the system revealed none of these fundamental effects. Conclusion. The results suggest that although the Gabor-filter stage may be appropriate for initial representation of visual information, modeling human object recognition requires specification of intermediate representations. Two-stage filter models likely derive their recognition power from precise specification of metric spatial relations for gray scale variation--information that people may directly employ for face recognition but not for real-time, entry-level object recognition.

45. Kalocsai, P., Biederman, I., & Cooper, E. E. (1994). To what extent can the recognition of unfamiliar faces be accounted for by a representation of the direct output of simple cells. Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Sarasota, Fl., May.

46. Fiser, J., Biederman, I., & Cooper, E. E. (1994). Test of a two-layer network as a model of human entry-level object recognition. Poster presented at the Third Annual Computational Neuroscience Meeting, Montery, CA, July.

47. Biederman, I., & Hummel, J. E. (1994). Real-time shape recognition: Implications for temporal asynchrony as an account of the binding problem. Invited paper presented at a symposium on Temporal Coding in the Brain at the 17th Annual Meeting of the European Neuroscience Association, Vienna, Austria, Sept.

48. Subramanian, S., Biederman, I., & Cowie, R. I. D. (1994). Priming the naming of impossible familiar objects. Paper presented at the Second Annual Workshop on Object Perception and Memory. St. Louis, Nov.

49. Kalocsai, P., & Biederman, I. When controlled for physical similarity, differences in emotional expression and gender produce much larger effects on face recognition than rotation in depth. Paper presented at the Second Annual Workshop on Object Perception and Memory. St. Louis, Nov.

Abstract:

Purpose. What computations are performed by the specialized areas tuned to the identification of faces? One possibility is that these areas transform the outputs of a lattice of hypercolumns of Gabor-like simple cells that characterize the early stages in the ventral pathway so that an intermediate representation which alters the simple cell similarity space is created. This intermediate representation might then be employed for the activation of identity. Another possibility is that these areas map the outputs of these simple cells without much change onto recognition units, similar to a two-layer network. The degree to which a two-layer network could account for the effects of rotational, emotional expression and gender changes on the speed and accuracy of the recognition of unfamiliar faces was assessed. Method. Subjects judged whether a pair of brief (100 msec), masked, sequential presentations of face images were of the same or different individuals. The images could differ in orientation in depth, emotional expression (neutral, smiling, surprised) and gender (for "different" trials). The similarity of each pair of faces was assessed by the Buhmann, Lades, & von der Malsburg's (1990) two-stage face recognition system. The system develops links between adjacent columns in the lattice so that relations among the filter activation values are coded. Results. Orientation, expression, and gender differences produced highly reliable effects on RTs and error rates. The effects of rotation of "same" RTs and errorrates were highly correlated with the similarity values calculated by the model, r =3D -.90. The correlation for the different expressions and gender were lower, r =3D -.59 (for same trials) and r =3D .41 (for different trials), respectively. However, the ranges of the similarity values for expression and gender were much smaller than that for rotation. When corrected for range attenuation (so that the ranges were equivalent to that for rotation), the correlations between similarity variations dues to differences in expression was -.97 and that for gender differences (on different RTs) were .98. Conclusion. When compared to the effects of rotation in depth, extremely small image variation produced by differences in expression or gender resulted in disproportionately large effects on reaction time and error rates.

50. Biederman, I., Subramaniam, S., & Madigan, S. F. (1994). Chance forced choice recognition memory for identifiable RSVP object pictures. Paper presented at the meetings of the Psychonomics Society, St. Louis, Nov.

51. Biederman, I. (1994). Recognition of faces and objects: implications for a general theory of shape recognition. Invited presentation at the ATR Symposium on Face and Object Recognition '95, Kansai, Japan, January.

52. Biederman, I. (1994). Invited panelist. Discussion of 3D object representation in the brain. ATR Symposium on Face and Object Recognition '95, Kansai, Japan, January.

53. Biederman, I. (1995). From image edges to geons to viewpoint-invariant object representations. Invited address (featured speaker) presented to the Vision Society of Japan, Tokyo, January.

54. Biederman, I. & Bar, M. (1995). An inadvertant experiment fails to confirm the the employment of viewpoint dependent mechanisms in human object recognition. Paper presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Ft. Lauderdale, Fl., May.

55. Kalocsai, P., & Biederman, I. (1995). Selective attention among presumed classifiers in the huamn face recognition system. Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Ft. Lauderdale, Fl., May.

56. Cooper, E. E., Subramaniam, S., & Biederman, I. (1995). Recognizing objects with an irregular part. Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Ft. Lauderdale, Fl., May.

57. Fiser, J., & Biederman, I. (1995). Priming with complementary gray-scale images in the spatial-frequency and orientation domains. Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Ft. Lauderdale, Fl., May.

Biederman and Cooper (1991) showed that line drawings with complementary contour deletions, that had no contours in common, but allowed activation of the same parts (geons), prime each other as well as they prime themselves in object recognition tasks. We assessed such priming in the spatial frequency (S.F.) domain with complementary pairs defined by either different SFs or different orientations. There was complete visual priming across scales but none across orientations. The results provide a spatial filter analog of the contour deletion results. In both cases visual priming occurred if the complementary images preserved the same geon structure (as with the S.F. complements) but no priming was evident if different parts were apparent in the complements (as with the orientation complements).

58. Subramaniam, S., Biederman, I., Kalocsai, P., & Madigan, S. R. (1995). Accurate identification, but chance forced-choice recognition for RSVP pictures. Poster presented at the Annual Meeting of the Association for Research in Vision and Opthalmology. Ft. Lauderdale, Fl., May.

After viewing several thousand pictures, human forced-choice recognition memory is extraordinarily high, with accuracy rates exceeding 90%. However, this level of performance is obtained when the pictures can be studied for several seconds. Methods and Results. At brief exposure durations (72 or 126 msec/image) with RSVP presentations, identification of an arbitrarily designated target was highly accurate (90-100% for most images) but forced choice recognition memory was at chance. RSVP presentations at these rates did not result in visual priming. Conclusions. Why is it that the same image that can be readily identified cannot leave a representation that can be recognized or affect subsequent perception? Tuned responding of

IT cells can occur in less than 100 msec and this activity may be sufficient for identification. However, several hundred msec of additional neural activity may be required for temporal binding so that a reduced code can be conveyed to the hippocampus or to other cortical structures involved in memory or priming. (By one account, the additional time allows the segregation of the activity from different object parts into different phase sets so that a viewpoint invariant structural description specifying the parts and their relations can be specified.) This hypothesis may explain the paradox that most of the information of IT cells is coded in the first 50 msec of their firing, yet the cells continue to fire for several hundred msec. The conditions of brief RSVP presentations may be well designed to allow the activation of the initial perceptual component of IT activity and to interfere with the sustained activity required for binding.

Invited Colloquiua: 1990-95:

1. University of Southern California (Department of Psychology)
2. University of Southern California (Department of Computer Science)
3. University of Paris
4. University of Kansas
5. Rice University
6. University of California, Berkeley
7. Stanford University
8. University of California, Santa Cruz
9. Navel Ocean Systems Center, Kailua Bay, Hawaii
10. University of Hawaii at Manoa
11. MIT Cognitive Neuroscience Program
12. Indiana University
13. University of Pennsylvania (Computer Science Department)
14. Pennsylvania State University
15. California Institute of Technology
16. UCLA
17. University of Toronto (Erindale, St. George)
18. University of California, Irvine
19. Bochum University (Germany)
20. Birmingham University (England)
21. Herriot-Watt University (Scotland)
22. Sterling University (Scotland)
23. Oxford University (England)
24. UC/Irvine
25. University of Southern California
26. California Institute of Technology
27. University of California
28. San Diego
29. Night Vision Laboratory (Fort Belvoir)
30. National Institutes of Mental Health (Neuropsychology Division)
31. Brooklyn College

32. University of Utah
33. INRIA (Sophia, France)
34. Trinity University (Dublin, Ireland)
35. Emory University
36. Yale University
37. Massachusetts Institue of Technology
38. AIR Kansai, Japan
39. RIKEN Tokyo
40. NTT Toyko
41. University of Haifa
42. Weizmann Institute, Rehovat, Israel
43. Hebrew University of Jerusalem
44. Utrecht University, Utrecht, Netherlands
45. Cornell University
46. University of Budapest, Hungary
47. Max Planck Institut for Brain Research, Frankfurt, Germany
48. Max Planck Institut for Biological Cybernetics, Tuebingen, Germany
49. NEC Research Institute
50. University of Tokyo, National Institue of Bioscience and Human-Technology, Tsukuba, Japan, Kyushu University
51. Chinese University of Hong Kong
52. University of Science and Technology, Beijing, China.