

AD-A213 290

DOCUMENTATION PAGE

Form Approved  
OMB No 0704-0188

UNCLASSIFIED		1b RESTRICTIVE MARKINGS	
2a SECURITY CLASSIFICATION AUTHORITY SELECTED		3 DISTRIBUTION AVAILABILITY OF REPORT Approved for public release; distribution unlimited	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE 1989		5 MONITORING ORGANIZATION REPORT NUMBER(S) 1989F	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) 1989F		7a NAME OF MONITORING ORGANIZATION Office of Naval Research	
6a NAME OF PERFORMING ORGANIZATION Donald A. Glaser		6b OFFICE SYMBOL (if applicable)	
6c ADDRESS (City, State, and ZIP Code) Dept. of Molecular and Cell Biology, c/o Stanley/Donner ASU, Univ. of California, Berkeley, CA 94720		7b ADDRESS (City, State, and ZIP Code) 800 North Quincy Street, Arlington, VA 22217-5000	
8a NAME OF FUNDING/SPONSORING ORGANIZATION		8b OFFICE SYMBOL (if applicable)	
8c ADDRESS (City, State, and ZIP Code)		9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-85-K-0692	
		10 SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO 61153N 42	PROJECT NO RR04209
		TASK NO RR0420901	WORK UNIT ACCESSION NO R&T4425089
11 TITLE (Include Security Classification) Computational and Psychophysical Study of Human Vision Using Neural Networks (UNCLASSIFIED)			
12 PERSONAL AUTHOR(S) Glaser, Donald Arthur Kumar, Tribhawan			
13a TYPE OF REPORT Final	13b TIME COVERED FROM 85/8/1 TO 89/1/31	14 DATE OF REPORT (Year, Month, Day) 89/4/28	15 PAGE COUNT 66
16 SUPPLEMENTARY NOTATION			
17 COSATI CODES		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	neural network, computational neuroscience, computer architecture, visual perception, cognitive processes, psychophysics, human visual system	
19 ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>The overall goal of our research program is to construct models of the human visual system that can be implemented on available computers and capture essential abilities of the real thing. These models should be useful in understanding how the human visual system works and for practical applications. In order to incorporate some of the known structural features of the brain in our models, we have chosen a neural net paradigm to mimic some aspects of the real nervous system. These networks contain nodes representing simplified nerve cells and can have an enormous variety of structures, some of which are the subjects of intensive study in many laboratories. Since so many different network structures are possible, it is necessary to use as much information as possible to limit the choice of nets to those most likely to be useful models of the human visual system. Our work in psychophysics is</p>			
20 DISTRIBUTION AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21 ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a NAME OF RESPONSIBLE INDIVIDUAL		22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL

designed to provide limits on the choice of architectures for model nets by requiring them to satisfy certain general conditions indicated by these experiments.

Several experimental projects will be described below concerning perception of relative depth and motion. One generalization that emerges from all of them is that local visual judgments can be grossly influenced by information gleaned from quite distant parts of a scene. To mimic the operation of the human visual system, then, a neural net must collect information from sizeable areas of a scene and use it to influence outputs from local visual processes.

Analysis of these results have led to some unexpected physiological conclusions that will be described below. Our work has also raised doubts about the utility of the idea that the human visual system has an extensive "front end" which functions more or less automatically without being affected by higher level or cognitive processes nor by signals from other "low level" cortical areas. Since the results of many psychophysical experiments depend on verbal instructions given to the observers, on their prior experience, and also on information gathered from parts of the visual scene remote from the region of immediate attention, we have been led to consider neural nets which are consortia of many similar subnets whose outputs are combined in some probabilistic way. This combination process is a way to merge a number of pieces of incomplete or low confidence information to achieve high confidence in the reality of a total percept which depends partially on all of these "weak" clues.

## ONR Technical Reports Distribution List Perceptual Sciences Program

Dr. Earl A. Alluisi  
OUSDR (A)/R&A(E&LS)  
Pentagon 3D129  
Washington, DC 20301-3080

Dr. Gary Aston-Jones  
New York University  
Department of Biology  
Washington Sq Ctr for NS  
New York, NY 10003

Dr. Christopher G. Atkeson  
M. I. T.  
Artificial Intelligence Lab  
545 Technology Square  
Cambridge, MA 02139

Dr. Alan D. Baddeley  
Director, Appl. Psychol. Unit  
Medical Research Council  
15 Chaucer Road  
Cambridge CB2 2EF  
ENGLAND

Dr. Bruce Barnes  
Program Director, Div. Info  
Robotics, & Intell. Systems  
National Science Foundation  
1800 G Street, N.W.  
Washington, DC 20550

Mr. Charles Bates  
Director, HE Division  
USAF AAMRL/HE  
WPAFB, OH 45433

Mr. Norman Beck  
Naval Underwater Systems  
Center  
Combat Control Systems  
Department Code 221  
Newport, RI 02840

Dr. Emilio Bizzi  
M. I. T.  
Whitaker College  
Cambridge, MA 02139

Dr. Kenneth R. Boff  
AAMRL/HE  
WPAFB, OH 45433

Mr Luiz Cabral  
Naval Underwater Sys. Ctr.  
Code 2212  
Bldg. 1171/1  
Newport, RI 02841

Cdr. Robert C. Carter USN  
Naval Research Laboratory  
Code 5532  
Washington, DC 20375-5000

Dr. Marvin S. Cohen  
Dec. Sci. Consortium, Inc.  
1895 Preston White Drive  
Suite 300  
Reston, VA 22091

Dr. Stanley C. Collyer  
Office of Naval Technology  
Code 222  
800 N. Quincy Street  
Arlington, VA 22217-5000

Defense Tech. Info. Center  
Cameron Station, Bldg 5  
Alexandria, VA 22314  
(2 copies)

Prof. Jay Elkerton  
University of Michigan  
Dept. of I & O Engineering  
Center for Ergonomics  
1205 Beal Avenue  
Ann Arbor, MI 48109-2117

Dr. Robert A. Fleming  
Naval Ocean Systems Center  
Human Factors Support Grp.  
1411 South Fern Street  
Arlington, VA 22202-2896

Dr. Alfred R. Fregly  
AF Office of Sci. Res.  
Life Sciences Directorate  
Bldg. 410  
Bolling AFB, DC 20332-6448

Dr. Jaocuin Fuster  
Univ of Calif - Los Angeles  
School of Medicine  
760 Westwood Plaza  
Los Angeles, CA 90024

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution /	
Availability Codes	
Dist	Availability for Special
A-1	

## ONR Technical Reports Distribution List Perceptual Science Program

Dr. Sherman Gee  
Office of Naval Technology  
Command & Control Technol.  
Code 221  
800 N. Quincy Street  
Arlington, VA 22217-5000

Dr. Apostolos P. Georgopoulos  
Johns Hopkins University  
School of Medicine  
Department of Neuroscience  
Baltimore, MD 21205

Dr. Donald A. Glaser  
Univ of California  
Dept of Molecular Biology  
Berkeley, CA 94720

Mr. Dan Greenwood  
Netrologic, Inc.  
5080 Shoreham Place  
Suite 201  
San Diego, CA 92122

Jeffrey Grossman  
Naval Ocean Systems Center  
Code 4403  
Bldg. 334  
San Diego, CA 92152-6800

Dr. Thomas E. Hanna  
Naval Sub Med Res Lab  
Behavioral Sciences Dept  
Naval Submarine Base  
Groton, CT 06349-5900

Mr. Paul Heckman  
Naval Ocean Systems Center  
San Diego, CA 92152

Dr. Steven Hillyard  
Univ of Calif - San Diego  
Department of Neuroscience  
La Jolla, CA 92093

Dr. Susan Hockfield  
Yale University  
School of Medicine  
333 Cedar Street  
New Haven, CT 06510

Prof. Donald D. Hoffman  
U. of California - Irvine  
School of Social Sciences  
Irvine, CA 92717

Dr. John M. Hollerbach  
M. I. T.  
Artificial Intelligence Lab  
545 Technology Square  
Cambridge, MA 02139

Prof. James H. Howard, Jr.  
Department of Psychology  
Human Performance Lab  
Catholic University  
Washington, DC 20064

Prof. Douglas E. Hunter  
Defense Intelligence College  
Washington, DC 20374

Dr. Edgar M. Johnson  
Technical Director  
US Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333-5600

Dr. William B. Johnson  
Search Technology, Inc.  
4725 Peachtree Corners Circle  
Suite 200  
Norcross, GA 30092

Capt T. E. Jones, MSC, USN  
Aviation Medicine & Human  
Performance (Code 404)  
Naval Medical R & D Cmd  
National Capital Region  
Bethesda, MD 21814-5044

Dr. Michael Kaplan,  
Director, Office Basic Res  
US Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333-5600

Dr. Milton S. Katz  
US Army Research Institute  
ARIESCO  
Box 65  
FPO, NY 09510-1500

## ONR Technical Reports Distribution List Perceptual Science Program

Dr. Steven W. Keele  
University of Oregon  
Department of Psychology  
Eugene, OR 97403-1219

Dr. J. A. Scott Kelso  
Florida Atlantic University  
Center for Complex Systems  
P. O. Box 3091  
Boca Raton, FL 33421-0991

Prof. Roberta Klatzky  
Department of Psychology  
U. of Calif.-Santa Barbara  
Santa Barbara, CA 93106

Prof. David L. Kleinman  
Electrical Engineering &  
Computer Science Dept.  
University of Connecticut  
Storrs, CT 06268

Dr. Christof Koch  
Calif Inst of Technology  
Division of Biology  
Pasadena, CA 91125

Dr. Arthur F. Kramer  
Department of Psychology  
University of Illinois  
Champaign, IL 61820

Mr. Keith Kramer  
Naval Research Laboratory  
Code 5532  
Washington, DC 20375-5000

Dr. Robert H. LaMotte  
Yale University  
School of Medicine  
333 Cedar Street  
New Haven, CT 06510

Dr. David Le Berge  
U. of California - Irvine  
Cognitive Sciences Dept.  
Irvine, CA 92717

Mr. I. Lemaire  
Naval Ocean Systems Center  
Code 09  
Engrg. & Computer Science  
San Diego, CA 92152

Prof. Gregory R. Lockhead  
Department of Psychology  
Duke University  
Durham, NC 27706

Dr. Richard Loutitt  
National Science Foundation  
Division of Behavioral &  
Neural Sciences  
1800 G Street, N. W.  
Washington, DC 20550

Dr. Robert A. Lufti  
Waisman Center  
University of Wisconsin  
Madison, WI 53706

Dr. John H. R. Maunsell  
University of Rochester  
Strong Sch of Med & Dent  
Box 642  
Rochester, NY 14642

Dr. John C. Middlebrooks  
University of Florida  
Dept. of Neuroscience  
Box J-J44, JHMHC  
Gainesville, FL 82610

Mr. Harold G. Miller  
Technical Director, War  
Gaming Department  
Naval War College  
Newport, RI 02841

Capt. William Moroney, USN  
Naval Air Development Ctr.  
Code 602  
Warminster, PA 18974

Dean of the Academic Dept.  
U. S. Naval Academy  
Annapolis, MD 21402-5018

Naval Aerospace Medical  
Research Laboratory  
Sensory Division Code 23  
Pensacola, FL 32508

## ONR Technical Reports Distribution List Perceptual Science Program

Commanding Officer  
Naval Air Systems Command  
Crew Station Design  
NAVAIR 5313  
Washington, DC 20361

Commanding Officer  
Naval Biodynamics Lab  
Michoud Station  
Box 29407  
New Orleans, LA 70189

Commanding Officer  
Navy Health Research Ctr.  
P.O. Box 85122  
San Diego, CA 92138

Commanding Officer  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Director  
Technical Information Div.  
Code 2627  
Naval Research Laboratory  
Washington, DC 20375-5000

US Naval Test Center  
Aircrew Systems Branch  
Systems Engineering Test  
Directorate  
Patuxent River, MD 20670

Commanding Officer  
Naval Weapons Center  
Human Factors Branch  
Code 3152  
China Lake, CA 93555

Dr. Allen Newell  
Dept. of Computer Science  
Carnegie-Mellon University  
Schenley Park  
Pittsburgh, PA 15213

Dr. Michael Letsky  
Office of the Chief of  
Naval Operations  
OP-01B7  
Washington, DC 20350

Office of the Chief of  
Naval Operations  
OP-933D3  
Washington, DC 20350-2000

Office of Naval Research  
Special Asst. for Marine  
Corps Matters  
Code 00MC  
800 N. Quincy Street  
Arlington, VA 22217-5000

Office of Naval Research  
Perceptual Science Program  
Code 1142PS  
800 N. Quincy Street  
Arlington, VA 22217-5000  
(3 copies)

Dr. Jesse Orlansky  
Inst. for Defense Analyses  
1801 N. Beauregard Street  
Alexandria, VA 22311

Lt. Dennis McBride, USN  
Pacific Missile Test Center  
Human Factors Branch  
Pt. Mugu, CA 93042

Dr. Howard E. Paschler  
Univ of Calif - San Diego  
Psychology Department  
La Jolla, CA 92093-0001

Dr. Richard Pew  
BBN Laboratories, Inc.  
Experimental Psychol. Dept.  
70 Fawcett Street  
Cambridge, MA 02238

Prof. Tomaso Poggio  
Massachusetts Institute  
of Technology  
Center for Biological  
Information Processing  
Cambridge, MA 02139

Dr. Michael I. Posner  
Washington University  
Medical Center  
St. Louis, MO 63130

## ONR Technical Reports Distribution List Perceptual Science Program

Dr. W. A. Rizzo  
Head, Human Factors Div.  
Naval Training Systems Ctr  
12350 Research Parkway  
Orlando, FL 32826-3224

Prof. Scott Robertson  
Department of Psychology  
Rutgers University  
Busch Campus  
New Brunswick, NJ 08903

Dr. Lawrence A. Rothblat  
George Washington University  
Department of Psychology  
Rice Hall, 6th Floor  
Washington, DC 20052

Dr. William B. Rouse  
Search Technology, Inc.  
4725 Peachtree Corners Circle  
Suite 200  
Norcross, GA 30092

Dr. Walter Schneider  
University of Pittsburgh  
LRDC Bldg.  
3939 O'Hara Street  
Pittsburgh, PA 15260

Prof. Thomas B. Sheridan  
Massachusetts Institute of  
Technology  
Dept. of Mechanical Engrg.  
Cambridge, MA 02139

Dr. Gordon L. Shulman  
Department of Neurology  
Wash Univ Schl of Medicine  
Box 8111  
St. Louis, MO 63110

Dr. Randall P. Shumaker  
Naval Research Laboratory  
Code 5510  
US Navy Center for ARAI  
Washington, DC 20375-5000

Prof. James A. Simmons  
Department of Psychology  
Brown University  
Providence, RI 02912

Dr. H. Wallace Sinaiko  
Manpower Research &  
Advisory Services  
Smithsonian Institution  
801 N. Pitt St. Suite 120  
Alexandria, VA 22314-1713

LCdr. Timothy Singer, USN  
Naval Air Development Ctr.  
Human Factors Engrg. Div.  
Warminster, PA 18974

Dr. Paul Slovic  
Decision Research, Inc.  
1201 Oak Street  
Eugene, OR 97401

Dr. Dana L. Small  
Naval Ocean Systems Ctr.  
Code 33  
San Diego, CA 92152-5000

Mr. James G. Smith  
Office of Naval Research  
Code 121  
800 N. Quincy Street  
Arlington, VA 22217-5000

Mr. Yale Smith  
Rome Air Development Ctr.  
RADC/COAD  
Griffis AFB, NY 13441-5700

Prof. Harry L. Snyder  
Dept of Industrial Engrg.  
VPISU  
Blacksburg, VA 24061

Capt. Frank Snyder, USN  
Naval War College  
Operations Department  
Newport, RI 02841

Cdr. S. Snyder  
Office of Naval Technology  
Code 222  
800 N. Quincy Street  
Arlington, VA 22217-5000

## ONR Technical Reports Distribution List Perceptual Science Program

Prof. Michael Sovereign  
Naval Postgraduate School  
Code 74  
Joint C3 Curriculum  
Monterey, CA 93943

Mr. Steven L. Speidel  
Naval Ocean Systems Ctr  
Code 632  
San Diego, CA 92152-5000

Dr. George Sperling  
New York University  
Department of Psychology  
6 Washington Place  
New York, NY 10003

Mr. Nelson F. Steele  
Advanced Res Dev Corp  
5457 Twin Knolls Road  
Columbia, MD 21035

Prof. Kent A. Stevens  
University of Oregon  
Dept. of Computer &  
Information Sciences  
Eugen, OR 97403

Dr. John Tangney  
AF Office of Sci. Res.  
Life Sciences Directorate  
Bldg. 410  
Bolling AFB, DC 20332-6448

Dr. James T. Todd  
Brandeis University  
Psychology Department  
Waltham, MA 02254

Dr. Martin A. Tolcott  
Decision Science Consortium  
1895 Preston White Drive  
Suite 300  
Reston, VA 22091-4369

Dr. Ronald J. Tusa  
Johns Hopkins University  
Neurology Department  
34th & Charles Place  
Baltimore, MD 21218

Dr. Amos Tversky  
Stanford University  
Dept. of Psychology  
Bldg. 420  
Stanford, CA 94305

Prof. William R. Uttal  
Department of Psychology  
Arizona State University  
Tempe, AZ 85287-1104

Dr. Harold P. Van Cott  
Committee on Human Factors  
NAS-National Res. Council  
2101 Constitution Ave, N.W.  
Washington, DC 20418

Dr. David C. Van Essen  
Calif Inst of Technology  
Biology Division  
1201 E. California Blvd.  
Pasadena, CA 91125

Dr. Charles S. Weaver  
MAXIM Technologies, Inc.  
3000 Patrick Henry Drive  
Santa Clara, CA 95054

Dr. John Weisz  
Technical Director  
US Army Human Engineering  
Laboratory  
Aberdeen P. G., MD 21005

Dr. Elizabeth M. Wenzel  
NASA-Ames Research Center  
MS # 239-3  
Moffett Field, CA 94035

Dr. Robert Wherry  
Analytics, Inc.  
2500 Maryland Road  
Willow Grove, PA 19090

Dr. Milton A. Whitcomb  
NAS-National Res. Council  
CHABA  
2101 Constitution Ave., N.W.  
Washington, DC 20418

## ONR Technical Reports Distribution List Perceptual Science Program

Prof. Christopher D. Wickens  
Department of Psychology  
University of Illinois  
Urbana, IL 61801

Dr. Diane Williams  
Naval Ocean Systems Ctr  
Code 4403  
Bldg. 334  
San Diego, CA 92152-6800

Dr. Dana Yoerger  
Woods Hole Oceanographic  
Institution  
Woods Hole, MA 02543

Dr. Wayne Zachary  
CHI Systems, Inc.  
Gwynedd Plaza III  
Bethlehem Pike  
Spring House, PA 19477

FINAL REPORT  
Computational and Psychophysical Study of  
Human Vision Using Neural Networks

Contract #N00014-85-K-0692

Office of Naval Research Report No. 1989F

Office of Naval Research

Prof. Donald A. Glaser, Principal Investigator  
Division of Neurobiology  
Department of Molecular and Cell Biology  
University of California, Berkeley  
Berkeley, CA 94720

Approved for public release; distribution unlimited

Reproduction in whole or in part is permitted for any  
purpose of the United States Government

## SUMMARY

The overall goal of our research program is to construct models of the human visual system that can be implemented on available computers and capture essential abilities of the real thing. These models should be useful in understanding how the human visual system works and for practical applications. In order to incorporate some of the known structural features of the brain in our models, we have chosen a neural net paradigm to mimic some aspects of the real nervous system. These networks contain nodes representing simplified nerve cells and can have an enormous variety of structures, some of which are the subjects of intensive study in many laboratories. Since so many different network structures are possible, it is necessary to use as much information as possible to limit the choice of nets to those most likely to be useful models of the human visual system. Our work in psychophysics is designed to provide limits on the choice of architectures for model nets by requiring them to satisfy certain general conditions indicated by these experiments.

Several experimental projects will be described below concerning perception of relative depth and motion. One generalization that emerges from all of them is that local visual judgments can be grossly influenced by information gleaned from quite distant parts of a scene. To mimic the operation of the human visual system, then, a neural net must collect information from sizeable areas of a scene and use it to influence outputs from local visual processes.

Analysis of these results have led to some unexpected physiological conclusions that will be described below. Our work has also raised doubts about the utility of the idea that the human visual system has an extensive "front end" which functions more or less automatically without being affected by higher level or cognitive processes nor by signals from other "low level" cortical areas. Since the results of many psychophysical experiments depend on verbal instructions given to the observers, on their prior experience, and also on information gathered from parts of the visual scene remote from the region of immediate attention, we have been led to consider neural nets which are consortia of many similar subnets whose outputs are combined in some probabilistic way. This combination process is a way to merge a number of pieces of incomplete or low confidence information to achieve high confidence in the reality of a total percept which depends partially on all of these "weak" clues.

## PERCEPTION OF RELATIVE DEPTH

Stereopsis is a task conventionally modelled with local operations. The problem of stereopsis is to determine the distance of objects from the images in the right and left eyes. The hard part of the problem is usually assumed to be the identification of matching constituents of the two images that both correspond to the same real feature of the scene being viewed. Random dot stereograms demonstrate that the correspondence does not have to be between recognizable contours in the two images. Accepting this, current models often begin by finding edges at different size scales in the images, usually by convolving the images with filters of different sizes. Each filter is often defined as the difference of two Gaussian windows of different widths. Edges defined by the coarsest DOG (difference of Gaussians) are sparsely distributed in many images and are least likely to yield mismatches. They are therefore usually found first to produce a rough correspondence between features of right and left images. Finer scale features are then easier to pair when this rough correspondence is known. In computer vision, the largest DOG used is about 0.5 degrees wide, which usually corresponds to an square about 34 pixels on a side for images containing 256x256 or 512x512 pixels. This procedure works fairly well for simple computer vision images, and can be parallelized for multiprocessor computers.

We have used psychophysical methods to probe human performance on particular tasks in order to learn how to construct stereopsis models with properties closer to human abilities. Our results show that human processing of depth related information is quite complex and not very similar to the usual stereo algorithms. Major failures of current stereo algorithms to mimic the performance of the human visual system are:

- 1) They cannot extract correct distance information from multiple transparent surfaces.
- 2) They have no mechanism for incorporating information from monocularly viewed areas into perceived images.
- 3) They do not do familiar problems faster than novel problems as do human observers who improve in speed with experience with a particular type of task.
- 4) They do not do "easy" stereo tasks faster than "difficult" ones as do human observers.
- 5) They ignore the influence of global image properties on local depth perception, unlike human observers.
- 6) They have no mechanism for using non-stereo cues to depth in achieving consistent depth perceptions in a scene.

APPENDIX A

Perception of Relative Depths of Features Moving in Depth

PERCEPTION OF RELATIVE DEPTHS  
OF FEATURES MOVING IN DEPTH

Tribhawan Kumar<sup>1</sup> and Donald A. Glaser<sup>2</sup>

<sup>1</sup>Department of Physics, University of California, Berkeley, CA 94720.  
Currently at the Department of Molecular Biology,  
University of California, Berkeley, CA 94720

<sup>2</sup>Departments of Physics and Molecular Biology  
University of California, Berkeley, CA 94720

ABSTRACT

When observers view a small test object in the center of the field view surrounded by a rectangular "picture frame", they always report that the object is moving in depth and never the frame if either or both are actually moving in depth. They perceive only motion in depth relative to the frame, which is perceived as stationary. However, the threshold for discrimination of depth differences within the central test pattern is significantly poorer for an actually moving pattern and fixed frame than for a fixed pattern and moving frame. Stereoacuity is shown not to be affected by offsetting the test pattern as much as 4 to 5 arc minutes of disparity from the fixation plane. It appears that mechanical convergence mechanisms need aiming accuracies no better than 4 to 5 arc minutes to allow the best stereoacuity thresholds to be achieved.

Acknowledgments. This work was supported by the U.S. Office of Naval Research, Contract No. N00014-85-K-0692.

Human automobile drivers have little difficulty in determining distances seen through a dirty cracked windshield on a rainy day, but no known depth algorithm can do this.

These conclusions are drawn from the psychophysical experiments briefly described below which provide requirements for any depth perception model intended to imitate the human visual system.

(1) Multiple plane random dot stereograms were generated as follows: A coarse black and white checkerboard-like pattern is imagined to be superimposed on the random dot stereogram to be built. A conventional random dot stereogram representing a plane floating **below** a flat sheet is built of small red dots and brown dots placed only in the "white" regions of the coarse checkerboard. Another conventional random dot stereogram representing a plane floating **above** a flat sheet is constructed using green and gray dots only in the "black" squares of the coarse checkerboard (other dot colors could have been used just as well). In this way two interleaved stereograms have been superimposed. Observers viewing this image report seeing all three planes having all four colors. No unique depth could be perceived for any particular dot of the random dot stereogram, and dot size could be varied from less than an arc minute to over 15 arc minutes without weakening the percept at all for most observers. Convolving the images with different sizes of DOG's produces no consistent edges. Since no dot is perceived to have a unique depth, it is not possible to assume "coherence of matter" to interpolate the depth between dots.

(2) To test for the incorporation of monocular images, we generated a random dot stereogram as follows: The image space was segmented with a rectangular grid whose cells were randomly colored black, light grey, or dark grey. One eye was shown this pattern and the other saw the same pattern with the light and dark grey cells interchanged and the black cells left the same. Observers report seeing two planes, one containing only grey squares and floating behind the other, which contained only the black squares. Next the observers were asked to pick a horizontal series of squares containing a black square followed by four or five grey squares and then another black square. They were then asked to call out the order of grey levels from left to right for the selected sequence, first with both eyes open and then with one eye closed. All the observers were quite surprised to find that they saw one rectangle less monocularly than they saw binocularly. When the observers reported the length of the selected sequence and the size of the cells, it was found that they were incorporating the monocular rectangle in the stereo percept without distorting the monocular distance scale. Each rectangle retained its size, and the length of the selected sequence of rectangles remained the same on monocular or binocular viewing, although there was one more rectangle binocularly than monocularly.

Under normal binocular viewing conditions, humans incorporate monocular features into the binocular space. The reader may easily verify

this by selecting a convenient edge (that of a door, a wall, or a desk; the rear view mirror of an automobile is ideal for this) and positioning himself so that some high contrast textured feature is occluded in one of the eyes. On binocular viewing, the feature is incorporated into the binocular space without any apparent spatial distortion of distances. This very simple observation raises questions about the nature of the "geometry" that is suitable in binocular "space." It is clearly not a simple matter of finding iso-disparity contours and leaving patches of "unmatched" (which, of course, implies unidentifiable) features on some convenient nearby surface. The incorporation is such that the dimensions of identifiable objects remains "undistorted."

3) For simple images in which feature correspondence is unique and not a problem, stereo psychophysical experiments show that depth perception in an area of local attention is greatly affected by elements of the image quite far away. If four dots on a horizontal axis are presented with different disparities, the relative perceived depth of any one of them depends on the perceived depth of the others, even for dots many degrees of arc apart. These effects are found even for viewing times as short as 10 milliseconds, but during extended viewing, perceived depths drift toward a final asymptotic value in about 250 milliseconds. Short viewing times consistently produce different relative depths than longer ones. If the viewer fixes attention on a particular "test dot," its apparent depth is dependent on whether the other dots are presented before or after the test dot. The perceived depth of a test dot is affected by a brief presentation of surrounding dots for about 500 milliseconds after they were shown. Surrounding dots presented after the test dot influence its perceived depth for about 200 milliseconds after the appearance of the test dot. Thus in a series of images, the apparent depth of an object is influenced both by other objects presented simultaneously with it and by objects which appear before and after it. This sluggish behavior of the depth perception system may help maintain coherence in the perceived scene during normal vision despite eye movements and the motion of objects. Neural nets designed to mimic depth perception by the human visual system must therefore have the correct temporal behavior as well as spatial behavior.

4) Images including additional cues to depth such as shading, perspective, and extended shape and orientation were also found experimentally to exhibit global influences on the perceived depth of test objects. For example the stereo image of a tilted plane generated by lines or dots of different disparities grossly influenced apparent depth of test dots as far as 25 degrees of arc away from the elements defining the tilted plane. It is well known that it is difficult to see a single tilted plane in depth correctly, but it is surprisingly easy to perceive the tilt if an oppositely tilted plane is also in the same scene, even when the two planes are tens of degrees apart. It is obvious that purely local mechanisms cannot account for these facts.

Experiments with a variety of background elements showed that these "induced depth" effects could not be explained by simple transformation of

coordinate frames or even the use of curved spaces. Here again we see the requirement for a neural net model to have the ability to implement the gathering and application of global information to predict correctly the nature of a local depth experience.

Detailed evidence for these results and conclusions are presented in the Appendices (drafts of papers to be submitted for publication: Appendix A, "Perception of Relative Depths of Features Moving in Depth," Tribhawan Kumar and Donald A. Glaser; and Appendix B, "Long Range Effects on Discrimination of Local Depth," Tribhawan Kumar and Donald A. Glaser).

### SPEED JUDGMENTS IN APPARENT AND "REAL" MOTION

Our apparent motion studies using 2 dots presented sequentially showed that just 2 additional "distractor" dots could interfere with judgments of speed in a way not predictable by any published version of the dual receptor motion detection schemes first proposed by Hassenstein and Reichardt for the house fly. Therefore we have learned that a more complex system, probably involving higher levels in a functional hierarchy, is required.

### ORIENTATIONAL ASYMMETRIES IN THE PERCEPTION OF APPARENT MOTION

An asymmetry was discovered when observers perceived vertical motion more often than horizontal motion in an ambiguous apparent motion display for which the vertical and horizontal interpretation were equally plausible. Detailed experimentation and analysis led to the conclusion that there is a vertical strip of human retina centered on the fovea that projects to both right and left visual cortices unlike the rest of the retina. It is concluded that those motion percepts studied which require correlation of information from both the right and left hemispheres are "weaker" than those involving only one hemisphere in the sense that the "weaker" percept is experienced less frequently than the "stronger" one. This result is interpreted to mean that signals transmitted by the corpus callosum produce percepts of lower confidence level than intrahemispheric percepts, perhaps because of noise or other signal degradation introduced by callosal transmission or due to dispersion of signals resulting from non-uniform speed of transmission.

### SCIENTIFIC PERSONNEL

Donald A. Glaser (principal investigator), Samuel Bowne (post-doctoral researcher), Wang Liang Zhong (visiting scholar); graduate research assistants Brent Beutter, Avijit Chaudhuri, Leonid Kruglyak, Tribhawan Kumar, Mark Kwong, Jim Lewis, Mark Liebman, Terry Regier, Bob Reid,

Joseph Weber, Jason Weisman, David Zingmond; and undergraduate research assistant Keith Eaton.

## PUBLICATIONS

- Bowne, S.F., and Glaser, D.A. (1987a). Speed discrimination using simple sampled-motion stimuli, 59th Ann. Meet. ARVO, Sarasota, Fl.
- Bowne, S.F., and Glaser, D.A. (1987b). Speed discrimination using simple sampled-motion stimuli, Suppl. Invest. Ophthal. & Vis. Sci., 28, 237.
- Bowne, S.F., McKee, S.P., and Glaser, D.A. (1989). Motion interference in speed discrimination, Jour. Opt. Soc. Am. A, in press.
- Chaudhuri, A. (1987). Aspects of metastable apparent motion, Ph.D. dissertation, School of Optometry, Univ. of California, Berkeley.
- Chaudhuri, A., and Glaser, D.A. (1986a). A new anisotropy in apparent motion, 58th Ann. Meeting ARVO, Sarasota, Fl.
- Chaudhuri, A., and Glaser, D.A. (1986b). A new anisotropy in apparent motion, Suppl. Invest. Ophthal. & Vis. Sci., 27, 345.
- Chaudhuri, A., and Glaser, D.A. (1987a). Differences between vertical and horizontal apparent motion, 59th Ann. Meeting ARVO, Sarasota, Fl.
- Chaudhuri, A., and Glaser, D.A. (1987b). Differences between vertical and horizontal apparent motion, Suppl. Invest. Ophthal. & Vis. Sci., 28, 251.
- Chaudhuri, A., and Glaser, D.A. (1988). Observations on "visual inertia" in apparent motion, in preparation.
- Chaudhuri, A., and Glaser, D.A. (1989). Psychophysical evidence for a bilateral projection for retina to cortex in humans, J. Physiol., submitted.
- Kumar, T. (1988). Local and global aspects of stereopsis, Ph.D. dissertation, Dept. of Physics, Univ. of California, Berkeley.
- Kumar, T., and Glaser, D.A. (1988). Distortion of disparity-induced depth perceptions by distant cues, in preparation.
- Kumar, T., and Glaser, D.A. (1989a). Long range effects on discrimination of local depth, in preparation.
- Kumar, T., and Glaser, D.A. (1989b). Perception of relative depths of features moving in depth, in preparation.

## INTRODUCTION

The perceived depth profile of the world does not change drastically as one turns his or her head even though the relationship between the two retinal images changes. With what accuracy must the position of the eyes be controlled to extract precise depths from the two eyes' views?

The relative change of retinal images during head motion and the role of absolute and relative disparity in determining depth has been discussed in the literature. Steinmen and Collewijn (1980) reported that head rotations cause retinal image motions between the eyes of about 2 to 3 degrees/sec for side to side head rotations of about 2Hz. (See their Fig. 1) Erkelens and Collewijn (1985) reported observation of version and vergence movements during dichoptic viewing of random dot stereograms moving sinusoidally toward and away from the viewer. Without a reference frame, the random dot stereograms appeared stationary despite the fact that the absolute convergence was changing. They concluded that only the relative disparities within an image determine the relative depth in different regions of a random-dot stereogram. Convergence is driven by absolute disparities but plays no role in determining relative depth in a random-dot stereogram. Fender and Julesz (1967) found that for retinally stabilized images, the observer required the images to come within 6 to 12 arc minutes of each other for fusion to take place. However, once fusion had taken place, they could pull the two retinal images apart by as much as two degrees before fusion

would be lost. This experiment was replicated by Piantinada (1986), who confirmed the results of Fender and Julesz except that he found that the initial distance for fusion could be significantly greater than the 6 to 12 arc minutes they reported. He found that the fusion and diplopia limits were much closer in size, and larger than about 40 arc minutes. Westheimer and McKee (1978) showed that stereoscopic acuity in the human fovea remains unimpaired during retinal image lateral motions of up to 2 deg/sec. However, the stereoscopic acuity degraded somewhat if the test features moved in depth. The experiments were conducted in a dimly lit room which provided a reference frame for the test features. Patterson and Fox (1984) reported that for stereopsis during oscillatory head motion of frequencies of about 2 Hz, the frequencies known to disrupt binocular correspondence, all observers reported no change in either task difficulty nor apparent depth visibility during head movements. This result was consistent with the hypothesis that precise binocular correspondence is unnecessary for processing stereoscopic information. Blakemore (1970) measured the relative disparity threshold for a two line task as a function of absolute disparity of the two lines. He reported that stereo-threshold degraded exponentially as the absolute disparity increased. In these experiments the smallest absolute disparity off the fixation

plane that was reported was 40 arc minutes.<sup>1</sup>

Initial fusion requires that the eyes be correctly aimed to within at least 40 arc minutes if one accepts Piantinida's estimate, or to within 10 arc minutes according to Julesz and Fender's estimate. However, once fusion has been achieved there appears to be little input required from absolute convergence to see depth. In this study we investigated the precision of control of absolute convergence required to perceive fine differences of depth. First the performance for stimuli moving in depth was determined qualitatively, and then the threshold for discriminating depth for the moving stimuli was measured. Finally the stereo threshold as a function of absolute disparity was measured around the fixation plane.

#### METHODS

Stimuli were presented on two identical Hewlett Packard vector oscilloscopes (HP 1345's) with P4 phosphor, which were set up so that the observer saw one with one eye and the other with the other eye. The arrangement consisted of orthogonally oriented polarizing sheets on the oscilloscope faces and in front of the observer's eyes, allowing only one scope screen to be

---

<sup>1</sup> Absolute disparity of a point is the disparity of that point with respect to the fixation point. Relative disparity of two points is the difference of the absolute disparities of those two points. This can easily be shown to be the angle subtended by the difference of distances between the two points in the left and right eye images. The absolute disparity can only be specified as accurately as the convergence on the fixation point. Relative disparity can be specified as accurately as the distance between the two points in the stimulus and the viewing distance.

visible to each eye. A beam splitting pellicle was used to superimpose the images of the two screens. The active areas of the screens were 11.2 centimeters wide and 8.5 centimeters high. There were 2048 addressable positions available on the oscilloscopes both horizontally and vertically. The room was too dark to detect anything but the stimuli when data were being collected. The observer was shown a pattern for a fixed length of time, every 2 or 3 seconds. The durations of presentation of the stimuli are specified for each experiment in the results section below. The pattern was preceded and followed by a blank screen for 250 milliseconds. For the rest of the time interval an inter-stimulus pattern was shown. The observer was asked to indicate whether a particular feature was nearer or farther than the reference features. For example, if the test features were two lines then the observer was to report whether the line to his or her left was nearer or farther than the line on the right. If the test features were three lines then the task was to decide whether the line in the middle was nearer or farther than the lines on either side ( these two lines always had equal disparities). The observer had to set a switch and press a button for the judgement to be recorded. For each presentation the test features had a relative disparity randomly selected from a set of several different disparities. Sessions were run so that the average number of responses per level of disparity of the test features was at least 30. For example, if seven different values of disparities of the test features were

selected to be shown, then at least 210 responses were collected in a session. The stimulus pattern, the inter-stimulus pattern and the makeup and size of the set from which the stimulus was randomly selected is specified for each experiment below. There were at least three sessions for each experiment. The results were tabulated as percent of responses for which the particular feature was judged nearer than the reference features. A psychometric curve was fitted through the points using the probit technique (Finney). The curve is specified by its mean (i.e. the 50% point), a slope and their associated errors. The 'threshold' was calculated from the fitted curve. The threshold is defined here as half the incremental relative disparity of the test dots required to go from 25% response to 75% response. The threshold reported is the average of at least three threshold values obtained for a given experiment. The convention used in this paper is that crossed disparities are positive and uncrossed disparities are negative.

## RESULTS

### Experiment 1

The stimulus was a rectangular frame 1.5 degree wide by 1.5 degrees high. In the center of the frame were two bright lines 10 arc minutes long, 15 arc minutes apart placed in the center of the frame symmetrically around the central vertical axis. A trial began by presenting an empty field for 250 msec, and then the above stimulus would be shown with either the frame or the

test lines executing an approximately sinusoidal motion in depth at a frequency of 1 or 3 Hz. The amplitude of the motion was 10 arc minutes of disparity behind, and 10 arc minutes in front of the fixation plane. These two frequencies are equivalent to an average disparity "velocity" of 40 arc minutes per second, and 2 arc degrees per second, respectively. Three complete cycles were shown, with the stimulus starting from the fixation plane, and the moving feature either coming out of or going behind the fixation plane. At the end of the three cycles, an empty field was presented until the observer reported whether the lines or the frame had been moving. When the experimental room was pitch black, all four observers always saw only the test lines move in depth regardless of whether the test lines or the surrounding frame were actually in motion. The observers saw the lines moving away when the disparity changes corresponded to the frame moving towards the observer. With the room dimly lit, the observers easily perceived the frame move when it was actually in motion. Various other test stimuli were tested with the same result; the observers never reported seeing the surrounding frame move in a pitch black room. The movement of the eyes was not monitored. Perhaps the eyes track to keep the frame stationary on the retina and all features move with respect to it.

To address this issue, the above experiment was repeated for presentation times of 100 milliseconds, allowing only part of the first cycle to be shown. The latency of human tracking and convergence is such that no eye movement is expected in this

amount of time; the motion in the stimulus is essentially equivalent to the motion on the retina (see Westheimer, 1954). For such briefly presented stimuli, there was very little sensation of movement in depth, though on closing one eye the lateral motion being executed by the monocular stimulus appeared to be very pronounced. Despite this all the observers reported that it was the enclosed features that appeared to have changed position, while the surrounding frame appeared stationary.

When the brief presentation experiment was repeated with the amplitude of motion being 2 degrees, then in 100 msec at average speeds of 2 deg/sec, the displacement was 12 arc minutes. Under these conditions, even though the surrounding frame still appeared to be stationary, and the enclosed test lines appeared to have moved, the observers could discriminate whether it was the frame or the enclosed test lines that had moved on the CRT face. The cue that allowed the discrimination was that when the test lines moved, they were diplopic near the end of the presentation, while for the motion of the frame the test lines appeared fused for the entire presentation. This result is consistent with the range of Panum's fusional range reported in the literature.

### Experiment 2

This experiment was intended to test whether observers show any differences in visual performance that depend on whether the test lines or the frame actually moved, even though they are unable to consciously perceive and report which elements moved.

It was designed to measure the thresholds for discriminating the relative depths of the two test lines during apparent motion events. The geometrical configuration of the stimulus was the same as for experiment 1. The presentation time was 75 msec, the trial time was 2.5 seconds, and the average speed of the moving feature was 2 deg/sec (i.e. a change of 9 arc min in depth in 75 msec). The stimuli shown had the frame moving with the test lines stationary, or the lines moving with the frame stationary, or both the frame and the test lines stationary. The moving feature could either be moving towards (starting 4.5 arc min behind and finishing 4.5 arc min in front of the fixation plane) or away (starting 4.5 arc min in front and finishing 4.5 arc min behind the fixation plane) from the observer. This gave five conditions (one for both lines and frame stationary, two for only the frame moving, and two for only the test lines moving), and for each condition seven relative disparity values between the two test lines were used to generate a set of 35 different stimuli. A stimulus from this set was randomly selected and presented to the observer, and his or her response recorded. The observers were asked to decide, "is the left line closer or farther away from you?". An error signal, a brief flash 15 minutes vertically below the test lines, was shown if the observer reported incorrectly; either the left line closer and the relative disparity was negative (crossed), or the left line farther and the relative disparity was positive (uncrossed). The results for the frame moving toward the observer and moving

It is clear from these results that although the observers could only report the relative movement or displacement of the frame and the test lines, their fine depth judgment depended on which feature was moving on the retina. It is obvious that the relative disparities in the two cases (only the frame moving or only the test lines moving) are similar, but the fine depth discrimination measurements can distinguish between these two cases on the basis of disparity changes which depend on retinal coordinates. In the above experiments the stimulus is presented in the fixation plane at some instance during each trial, and the improvement of threshold for stationary lines may be attributed to the longer time the test lines spend near the fixation plane. How is the fine depth discrimination task affected when the stimulus and the reference features are presented off the fixation plane? The next experiment addresses this question.

### Experiment 3

The stimulus pattern was three vertical lines 12 arc min. apart. The lines were 10 arc min long and less than a min wide. The inter-stimulus pattern was a rectangular bracket (see fig. 5) 48 min wide by 30 min high with a fixation dot in the center. The inter-stimulus pattern was designed to aid convergence and to specify a fixation plane. The three line stimulus was presented for 100 msec, and a trial time was 2.5 sec. The three lines were

away from the observer were averaged together to measure the effect of the frame motion. Since the results showed no difference between the conditions when the test lines were moving away or towards the observer, the results for these two conditions were averaged together to measure the effect of the motion of the test lines. The responses as fraction of times the left line was seen closer were fitted to the integral of a gaussian using the probit technique (Finney). For comparison data were also collected for the situation in which the frame and the test lines moved laterally within the fixation plane. The presentation time, and the stimulus configuration was kept the same as for the other cases. Figures 1 through 4 show the results for three different observers. As seen in Figure 1, the threshold for discriminating depth was unaffected by speeds up to 2 deg/sec for lateral motion across the eyes (the frame and the lines moved together either left or right, and there was no motion in depth). This is in accordance with the results reported by Westheimer and Mckee (1978). Figures 2 through 4 show the threshold for discriminating relative depth between the two test lines when either only the frame was moving in depth or only the test lines were moving in depth. The observers clearly found discriminating the relative depth of the moving lines harder than either of the other two tasks, and for at least two of the observers the motion of the frame appears to have provided enough of a "distraction" to make the discrimination task of the central test lines somewhat more difficult.

shown with an offset disparity<sup>2</sup> relative to the fixation plane defined by the inter-stimulus pattern. The value of the offset disparity was randomly selected from five preselected values. The data were collected for offset disparities (crossed and uncrossed) of 0, 1, 2, 3, 4 and 5 arc min. These values were grouped into three different groups with each group having five offset disparity values. The three groups were (0,-1,+1,-3,+3 arc min), (0,-2,+2,-4,+4 arc min) and (0,-3,+3,-5,+5 arc min). For each value of offset disparity of the three lines, seven disparity values for the middle line were selected. This generated a set of 35 stimuli for each group (7 relative disparities of the middle line relative to the flanking outer two lines for each of the 5 offset disparities of all the three lines with respect to the fixation plane), and the stimulus presented was randomly selected from one of these 35. Each group was run independently in separate sessions and data were collected from two observers. As can be seen in Fig 6 stereoacuity thresholds were unaffected for both the observers for offset disparities of less than 4 arc min, and showed some deterioration for offset disparities of 5 arc min. The experiment was repeated with the separation between the three vertical lines reduced to 5 arc min.

---

<sup>2</sup> Offset relative disparity is essentially the same as absolute disparity if one accepts that the observer was fixating on the fixation plane and not slightly off it. Since we were not measuring either the fixation disparity or eye position, we cannot specify absolute disparity of the lines. However, the two values should be fairly close to each other since fixation disparity is usually reported to be less than 2 arc minutes anyway.

The other parameters of the stimulus, the inter-stimulus pattern, the presentation time and the time for a trial were kept the same as before. The stereoacuity thresholds were a little higher (see Fig 7), but the effect of the offset disparity was the same: no change for offset disparities of less than 4 arc min, and a slight deterioration for offset disparities of 5 arc min. The observers were asked to report if they ever saw more than three lines in the stimulus, which would indicate incomplete fusion, or the possibility of matching two lines in each eye leaving an unmatched line on either end giving a perception of four lines in the stimulus. This would be expected if there were a point for point matching for the retinally "corresponding points" for the given fixation. The observers always reported only three lines.

#### DISCUSSION

Observers in these experiments qualitatively perceive only relative motion in depth and are unable to specify whether the test features or the reference features are moving. They are better able to discriminate fine differences in depth for features that are stationary with respect to the retina. For fine depth discrimination tasks convergence has to bring the two eyes' views into registration to within 3 to 4 arc min. The rules describing which feature is seen moving during relative motion of features were not investigated in detail. Preliminary investigations lead us to suspect that these rules will be similar to the induced motion effects seen in the frontal plane

(see Day, Millar & Dickinson 1979), and that there will be similar difficulty in defining precise attributes of the "frame of reference" with respect to which the motion is perceived. The induced motion in the frontal plane should also show a similar quantitative difference between the stationary and the moving test feature when measuring discrimination thresholds as we have found for motion in depth.

The ability to discriminate depth better for stationary objects cannot be attributed to the fact that stationary objects spend more time near the fixation plane since the motion spanned 4.5 arc min in front of and behind the fixation plane and as shown the ability to discriminate depth within this range of offset from the fixation plane is hardly affected. The difference is probably due to the sluggish response of the stereo system to changes in depth in time (see Kumar, 1988). If one postulates special units for detecting motion in depth (Beverly and Regan) then our results indicate that these units perhaps do not discriminate as fine depths as does the system responsible for stationary stereopsis.

Westheimer reported that stereothresholds for discriminating depths degrade even when the disparity "pedestal" is as little as 1 arc min. His stimulus was three lines with the middle line having some disparity with respect to the outer two lines. The task was to determine the just noticeable difference in the disparity of the middle line from the initial disparity with respect to the outer two lines. Our results are not in conflict

with this, and address a different aspect of stereo-processing. In our experiments the disparity of the middle line being discriminated is the difference from zero disparity of the middle line with respect to the outer two lines. The offset disparity is given to all the three lines which leaves the relative disparities of the three lines unchanged.

These experiments didn't explore the effects of convergence control on the choice of what is "matched" between the two eye's views. Discrimination of fine depth differences in ambiguous stimuli may require finer convergence control than 3 to 4 arc minutes. Grouping and configurational effects do override spatial proximity of retinal "corresponding points" to determine the matching of the two eyes views (see Ramachandran and Nelson, 1976; this is similar to the observation above where the observers never reported more than three lines though two lines out of three in each image were on corresponding points, and the outermost lines were "unmatched"). This overriding of spatial proximity of retinal correspondence takes place for presentation times as brief as 100 msec, which rule out any role for eye movements and change of convergence. It is likely that the visual system relies on processing techniques for finer control of correspondence rather than the mechanical alignment of the eyes. However, McKee and Mitchison (1988) report some possible role of eye position control finer than 3 to 4 arc min. in determining "matches" in stimuli consisting of regularly spaced dots when viewing times were much longer (up to several seconds).

## REFERENCES

- Beverly, K.I., and Regan, D. (1973) Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *J. Physiol., Lond.* 235, 17-29.
- Blakemore, C. (1970) Binocular depth discrimination and the nasotemporal division. *J. Physiol. Lond.* 205, 471.
- Day, R.H., Millar, J.H., and Dickinson, R.G. (1979) Induced movement as nonveridical resolution of displacement ambiguity: Effect of enclosure and number of field elements. *Percept. and Psychophys.* 25, 23-28.
- Erkelens, C. J., and Collewijn (1985) Motion perception during dichoptic viewing of moving random dot stereograms. *Vision Res.* 25, 583-588.
- Fender, D.H., and Julesz, B. (1967) Extension of Panum's fusional area in binocularly stabilized vision. *J. Opt. Soc. Am.* 57, 819-830.
- Kumar, T. (1988) Local and global aspects of stereopsis. Ph.D. dissertation, Univ. of California, Berkeley.
- McKee, S.P., and Mitchison, G.J. (1988) The role of retinal correspondence in stereoscopic matching. *Vision Res.* 28, 1001-1012.
- Patterson, R., and Fox, R. (1984) Stereopsis during continuous head motion. *Vision Res.* 24, 2001-2003.
- Piantinada, T.P. (1986) Stereo hysteresis revisited. *Vision Res.* 26, 431-437.
- Ramachandran, V.S., and Nelson, J.I. (1976) Global grouping overrides point-to-point disparities, *Perception* 5, 125-128.
- Steinmen, R.M., and Collewijn, H. (1980) Binocular retinal image motion during head rotation. *Vision Res.* 20, 415-429.
- Westheimer, G. (1954) Eye movement responses to a horizontally moving visual stimulus. *Archs Ophthal.* 52, 932-941.
- Westheimer, G., and McKee, S.P. (1978) Stereoscopic acuity for moving retinal images. *J. Opt. Soc. Am.*, 68, 4, 450-455.

FIGURE 1  
Lateral motion

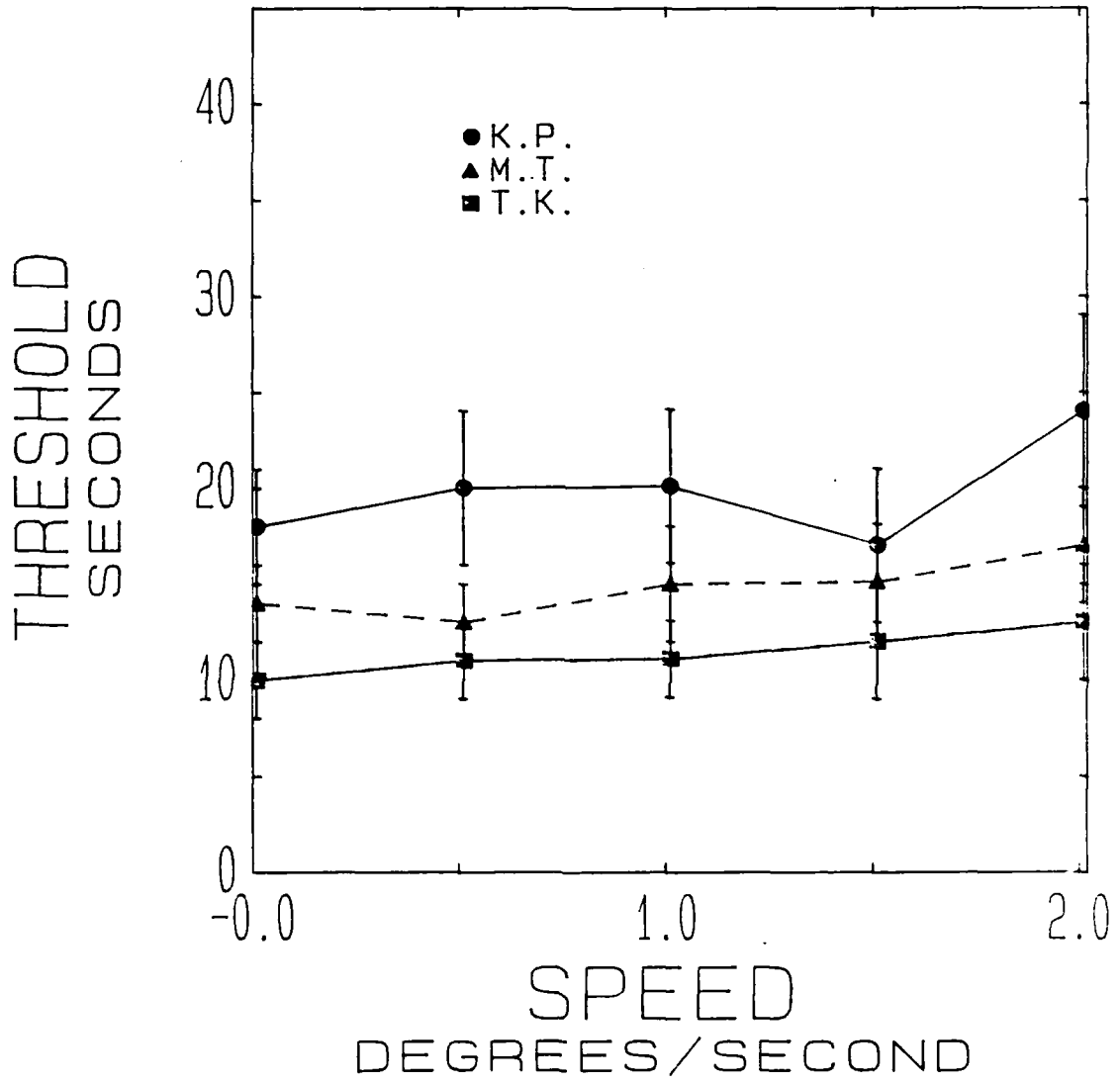


FIGURE 2  
Observer K.P.

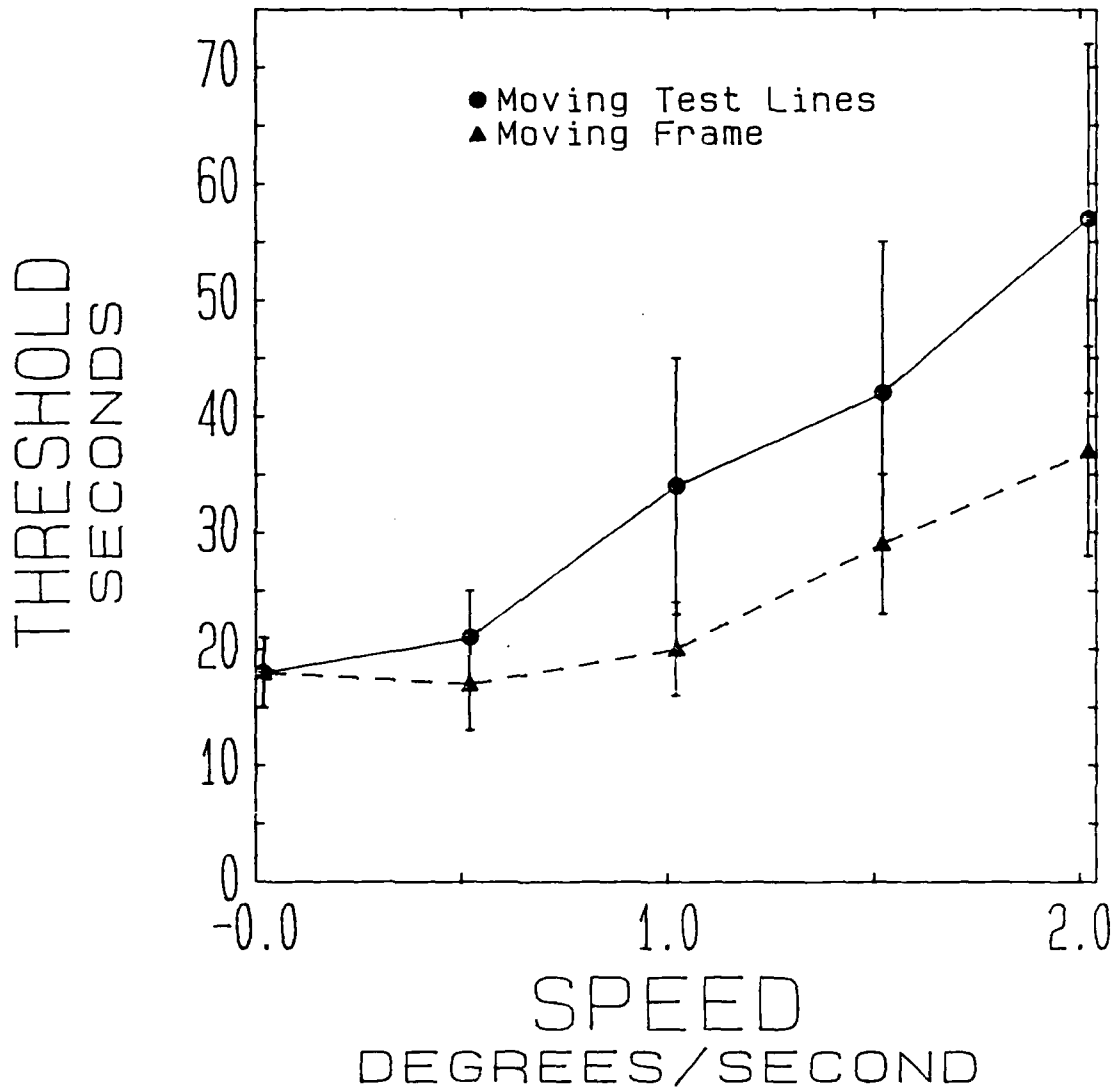


FIGURE 3  
Observer M.T.

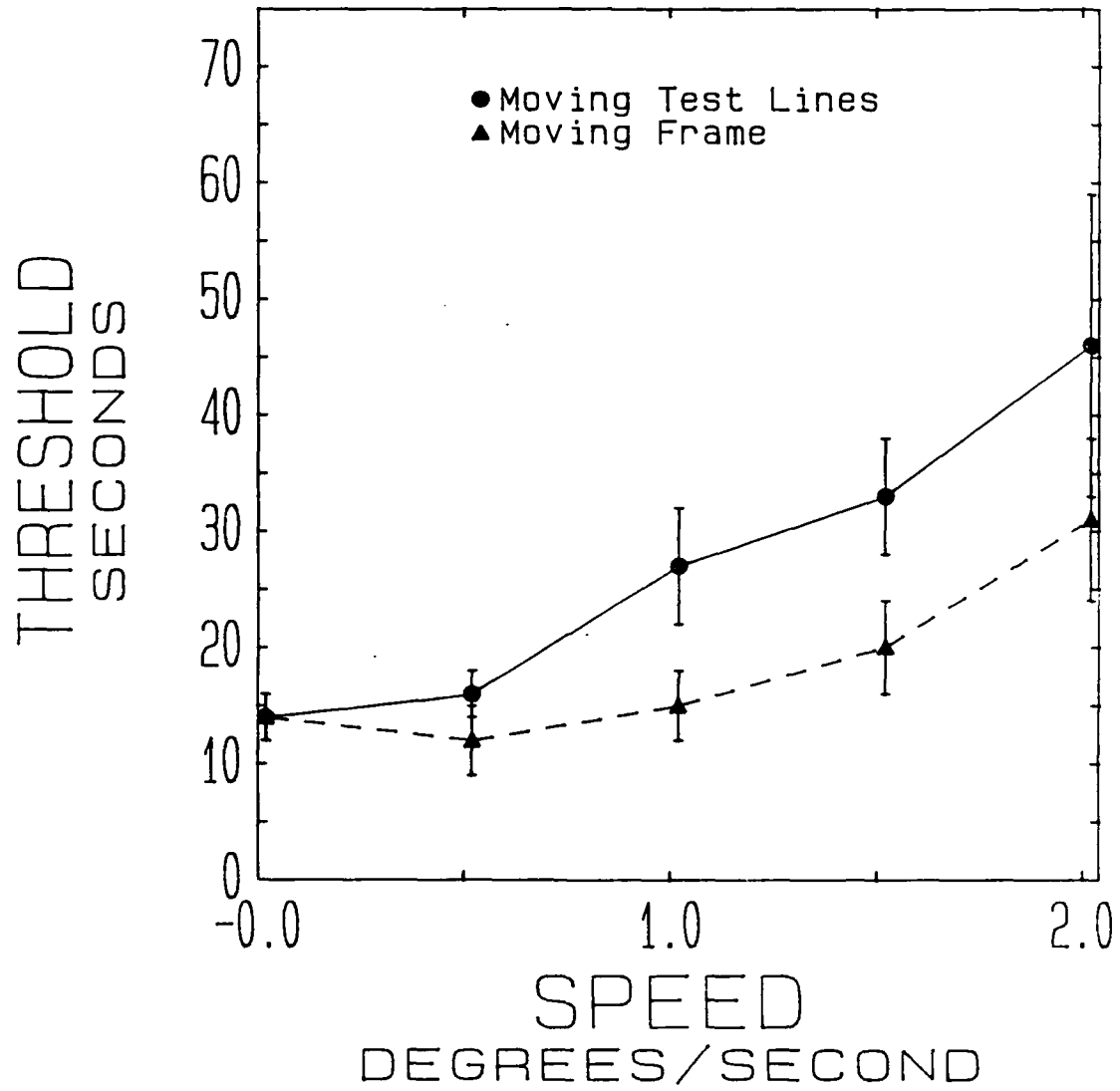
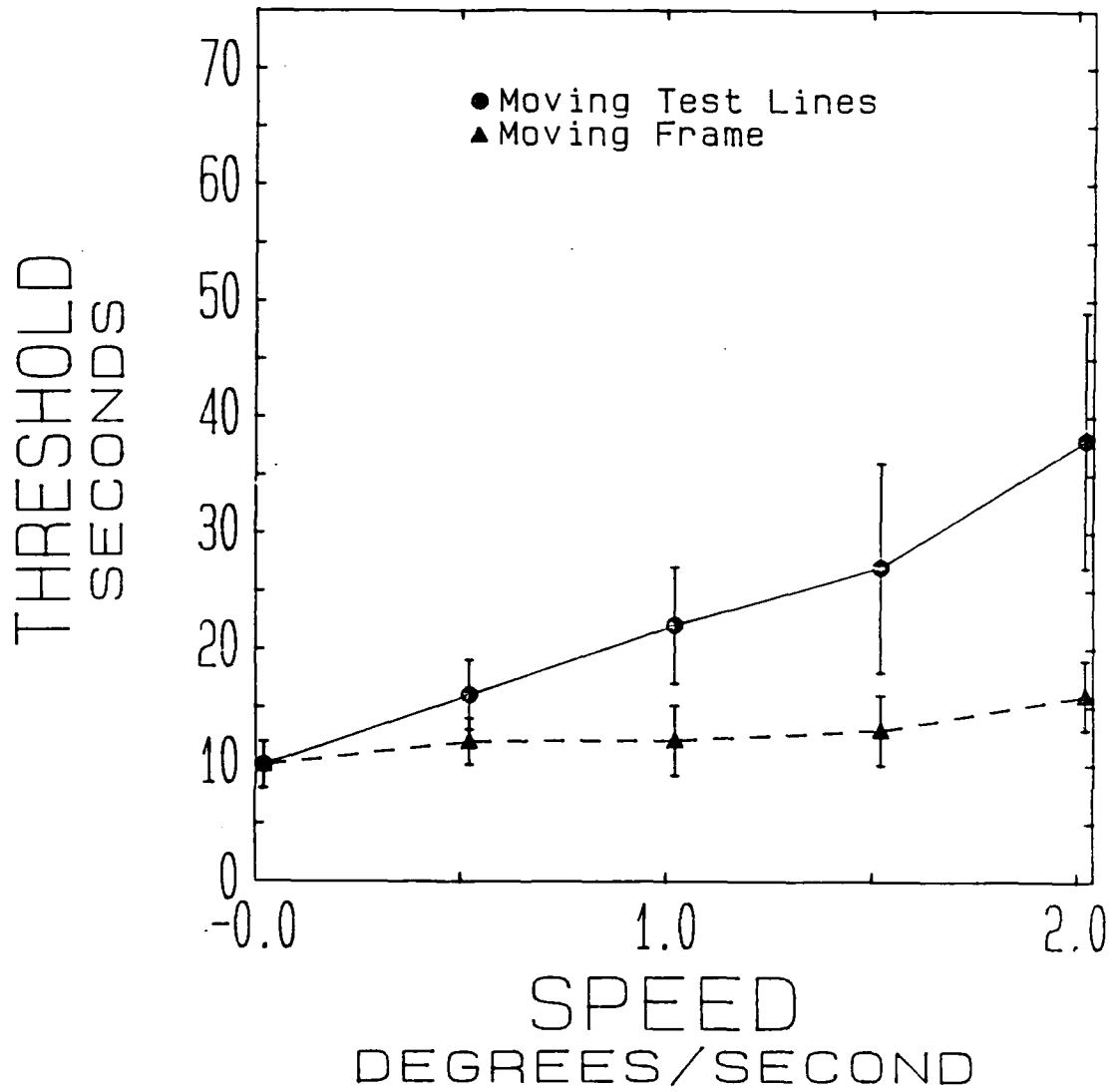
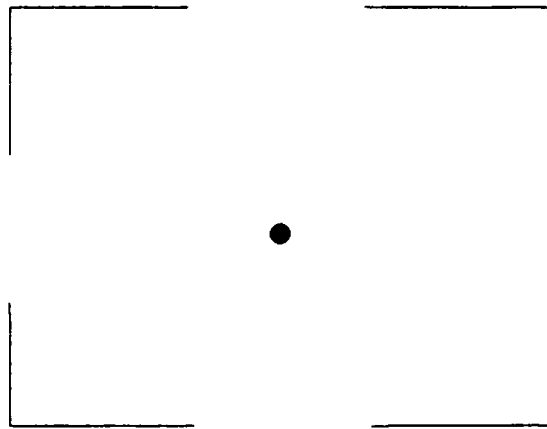


FIGURE 4  
Observer T.K.



# FIGURE 5



The angular bracket was 48' by 30'

The gap in the horizontal side was 24'

The gap in the vertical side was 10'

FIGURE 6

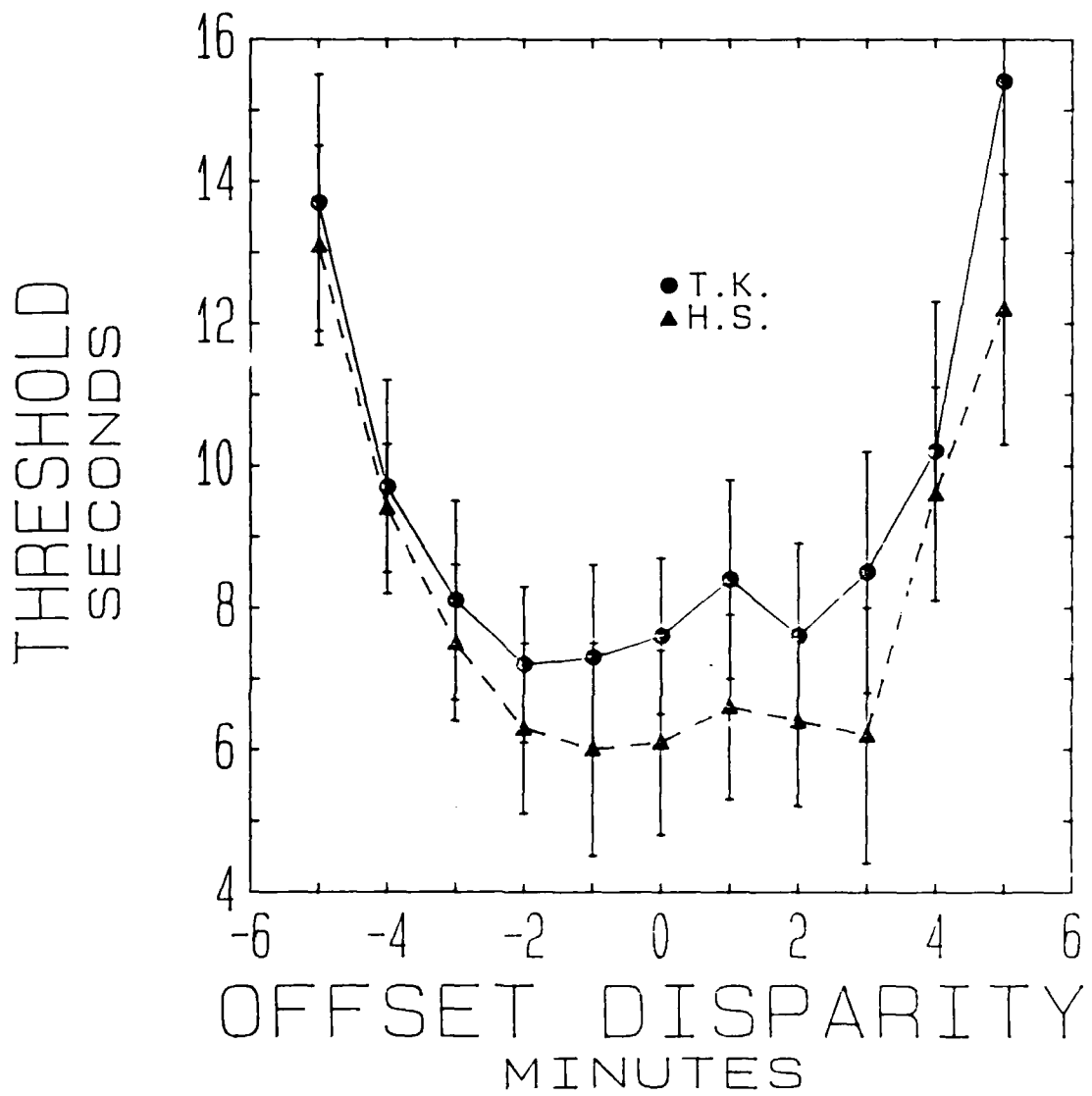
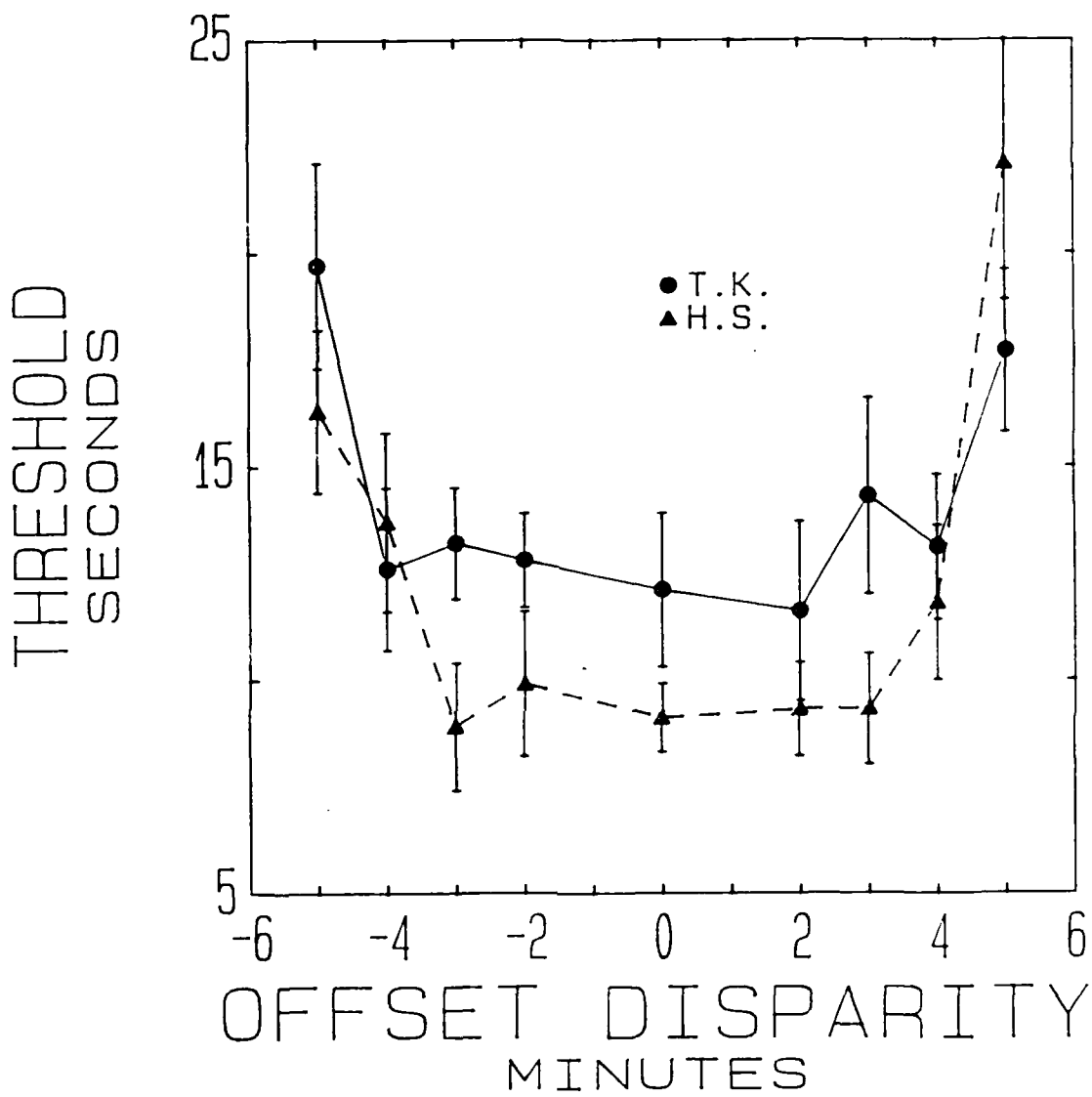


FIGURE 7



APPENDIX B

Long Range Effects on Discrimination of Local Depth

LONG RANGE EFFECTS ON  
DISCRIMINATION OF LOCAL DEPTH

Tribhawan Kumar<sup>1</sup> and Donald A. Glaser<sup>2</sup>

<sup>1</sup>Department of Physics, University of California, Berkeley, CA 94720.  
Currently at the Department of Molecular Biology,  
University of California, Berkeley, CA 94720

<sup>2</sup>Departments of Physics and Molecular Biology  
University of California, Berkeley, CA 94720

ABSTRACT

Observers judgement of relative depth between two test dots shown foveally is shown to be influenced by surrounding test dots that are many degrees away (51 degrees was the largest separation that was tested) and have much larger disparities (largest relative disparity of the outlying features tested was 20 degrees) than the disparities of the test dots. This influence is seen for briefly presented (100 msec or less) stimuli which rule out explanations requiring changing eye positions. The measured influence is inversely proportional to the spatial separation between the features, and shows saturation when relative disparities of remote features are greater than 2 degrees. Different observers show almost qualitative differences for various configuration of the outlying features. This rules out explanations based on only the spatial positions and disparities of the features in the stimuli.

Acknowledgments. This work was supported by the U.S. Office of Naval Research, Contract No. N00014-85-K-0692.

## INTRODUCTION

That the depth of a feature is influenced by the depth of other features in the visual field has been shown by many researchers (see Jaensch 1911; Wallach and Lindauer 1961; Pastore 1964; Werner 1937; Koffka 1935; Gogel 1956; Richards 1971; Mitchison and Westheimer, 1984; Westheimer 1986). Stereoacuity has also been shown to degrade progressively when an unrestricted view is increasingly restricted (Luria 1969). A few observers tested by Luria showed a slight degradation in stereoacuity when an unrestricted view was restricted to 45 degrees. This research explored spatial separations between features of no more than a degree or so, although Richards (1971) had previously shown that disparities of the ends of a vertical bar as large as 16 degrees masked the depth of the central part of the same feature (the vertical bar).

At present there is an implicit concept of interaction among the disparity-detecting units responsible for stereopsis. How far do these interactions extend spatially and in the disparity domain? This report examines the interaction of disparity signals across large spatial ranges (up to 51 degrees) and over large disparities (up to 20 degrees). The interest lies in determining the nature of interactions giving rise to incorporation of depth of surrounding features and the nature of disparity masking over large spatial distances. The results show that perceived differences in the depth of items near fixation are grossly influenced by image elements up to 25 degrees away

and that large disparities of remote features do influence the relatively small disparities of the central features. The results also show significant differences among the observers, raising the possibility that the influence of depth may not be due solely to disparity of surrounding features. The findings demonstrate that naive subjects are unaware that the image elements in the periphery are being manipulated and that subjects' local relative depth judgments are affected by peripheral image elements. Although observers are able to report the relative depth of the peripheral image elements when asked, they are unable to report the relative depth of peripheral image items if also required to report the depth of central test features.

#### METHODS

Stimuli were presented on two identical Hewlett-Packard vector oscilloscopes (HP 1345's) with P4 phosphor, appropriately arranged using polarizing filters to present stereo images. The active area of the oscilloscopes was 11.2 centimeters horizontally and 8.5 centimeters vertically, limiting the field to about 17 degrees by 13 degrees at 36 centimeters viewing distance. The addressability of the oscilloscopes was 2048 positions horizontally and vertically. The luminance of the patterns was 100 cd/m<sup>2</sup>. The room was dark enough so that only the stimulus was visible when data were being collected, eliminating any incidental cues that may have been available from objects in the room.

The observer was shown a pattern for a fixed length of time,

usually 50 milliseconds, every 1.5 or 2 seconds. The pattern was preceded and followed by a blank screen for 250 milliseconds. For the rest of the time interval a fixation dot was shown at the center of the screen. Two test dots were shown symmetrical to the fixation dot during the presentation time of the pattern. During this time the fixation dot was not visible. The observer was asked to respond whether the dot on the left was nearer or farther away than the test dot on the right. To record his or her judgment, the observer set a switch and pressed a button. For each demonstration, the test dots had a relative disparity randomly selected from a set of six or seven different disparities. No error feedback was provided.

Experiments were run in sessions of 300 responses. Each experiment consisted of two cases. The second case stimuli was the same as the first case except all of the surrounding features around the test dots had their disparities reversed in sign. Following convention, crossed disparities were labeled positive and uncrossed disparities, negative.

The results were tabulated as percent of responses reporting the left dot as farther away than the right dot versus the different relative disparity of the two test dots. A psychometric curve was fitted through the points using the probit technique (Finney). The curve is specified by its mean (i.e., the 50% point), a slope and their associated errors. The "threshold" was calculated from the fitted curve and is defined here as half the incremental relative disparity of the test dots

required to go from 25% correct response to 75% correct response. The difference between the means of the two cases is the parameter referred to as the observed shift of the mean (DELTA). The error in DELTA is the square root of the sum of squares of the errors in the two means used to calculate the shift. For most points DELTA was measured at least three times over a period of several weeks. The DELTA reported is the mean of the measured DELTA's. The error reported is either the weighted error of the standard error associated with each DELTA, or the standard deviation of the spread in the measured means, whichever is larger.

For a field of view larger than 17 degrees by 13 degrees, the stimulus was projected on a screen using two slide projectors. Polaroid sheets were placed in front of the projector lenses, and appropriately aligned polaroid glasses worn by the observer allowed her to see only one of the projections with each eye. Electronic shutters mounted on the lenses of the projectors were controlled by a personal computer. This setup allowed a presentation time of 100 milliseconds. A set of mounted slides was placed in the slide carousels, and the positioning of the carousels was done manually.

All of the observers, except one (one of the authors) were naive as to the purpose of the experiment. The naive observers were undergraduate students who had been tested for stereoacuity using a three-line task. The lines were 15 minutes in length, and 15 minutes apart. The outer two lines had zero disparity,

and the middle line's disparity was varied. Subjects were selected if their initial stereoacuity threshold was less than 40 arc seconds. All of the subjects showed decreasing thresholds with practice, and in time all of them approached a very respectable level of less than 10 arc seconds.

## RESULTS

### Experiment 1

The depth contrast effect was explored parametrically in this experiment. The underlying stimulus used was larger in all dimensions but similar to the configuration used by Werner (1938) to demonstrate depth contrast. The stimulus consisted of a rectangular frame 12.44 degrees wide and 9.36 degrees high in one eye and of varying smaller width and same height in the other eye. There were two symmetrically placed test dots 21.6 arc minutes apart in the center. The stimulus was shown for 100 milliseconds preceded and followed by an empty visual field which lasted 250 milliseconds. The difference in the widths of the rectangular frame in the two eyes is the disparity of the frame (D) referred to in Figure 1. The two cases consisted of a set of twelve or fourteen different displays. Each case had half of these displays, with each display corresponding to a different relative disparity of the test dots. The first case had the wider frame shown to the left eye and the narrower frame shown to the right eye (this is selected to be negative frame disparity by convention). The second case had the wider frame shown to the right eye and the narrower frame shown to the left eye (this case

being the positive frame disparity). For each case six or seven different relative disparities of the test dots were selected. The stimulus that was shown was randomly selected from this set of twelve or fourteen displays.

After recording 300 responses, a psychometric function was fitted to each case giving the mean, the slope and their associated errors. The mean shift for the above configuration for four different observers is given in Figure 1. The associated thresholds for three of these observers is given in Figure 2. Figure 3 and 4 gives the mean shift for two other observers when the dimensions of the stimulus were changed. The dimensions of the stimulus for Figure 2 are 9.8 degrees for the wider frame and 7.5 degrees high. The test features are now two vertical lines 10 minutes long, 49.2 minutes apart.

All subjects showed the same qualitative behavior, although the actual numbers for the shift differed by a factor of two for the same stimulus between observers. Even though the stimulus was diplopic, frame disparities of 8 degrees yielded consistent and reliable shifts. The measured shift is proportional to the frame disparity until the disparity is about 2 degrees. Beyond that the shift is roughly constant or slightly decreasing and may indicate processing of the frame's edges in qualitative stereopsis range (see Ogle 1950). The shift was measured for different frame widths and test dot separations. The results for two observers are given in Figures 5 and 6. The measured shift is inversely proportional to the frame width over the range

measured. These results demonstrate remarkable support for Gogel's adjacency principle and are in agreement with Mitchison and Westheimer's (1984) implementation in their Saliency concept. Although the measured shift generally increases with test dot separation, it is not proportional to the test dot separation.

Table 1 lists the shifts for three subjects when the outer frame was 17 degrees in one eye and 13.6 degree in the other eye; the test dots were 15 minutes apart. The presentation time of the stimulus was 100 milliseconds. For this stimulus, after the data had been collected three times over a period of two weeks, the naive observers were asked if they had noticed whether the frame was flat and in the fronto parallel plane or tilted out of the plane. They had paid no attention to the frame and did not know. The same stimulus was shown again, and the observers were asked to respond as to whether the left side of the frame was in front or behind the right side of the frame. All of them correctly identified the depth 100 percent of the time at the above frame disparity. The observers were then asked to respond only when the left side of the frame was in front of the right side (positive frame disparity case). For these responses they were to indicate whether the left central dot was in front or behind the right central dot. The average of the thresholds obtained for the runs when the number of responses for the negative frame disparity to the total number of responses was less than 10 percent is also given in Table 1. It is clear that the ability of the observers to report the relative depth of the

central features is considerably impaired when they are required to be simultaneously aware of the depth of the peripheral features.

### Experiment 2

To test for frame wider than 17 degrees slide projectors setup was used. The test features were two lines one degree long and 3 degrees apart and had zero relative disparity. The larger frame width was 51 degrees and the frame disparity was 12 degrees. The width of the test lines and the lines of frame was 10 minutes. The observer was asked to report whether the left central line was in front of or behind the right test line. The set of stimuli consisted of zero degree frame disparity, +/- 12 degree frame disparity, +/- 20 degree frame disparity, a case with no frame, and another one in which the frame was shown to only one eye (monocular frame). The observer was shown each stimulus in random order. Each stimulus was shown a minimum of 30 times. The results, the fraction of the time the left line was reported in front of the right line, for three naive observers who did not know about the purpose of the experiment, and one experienced observer (T.K., one of the authors) are given in Table 2. It is clear that all observers except one were influenced by the disparity of the frame even when the frame disparity was 20 degrees.

### Experiment 3

In the above experiments the vertical edges of the rectangular frame had relative disparity with respect to each

other. The authors wished to test whether there was any interaction over spatial distance if the frame edges, while having no relative disparity with respect to each other, had an offset disparity with respect to the central test lines. The offset disparity showed the frame in the fronto parallel plane, but either in front of or behind the central test lines. A frame of dimensions 6 degrees by 9 degrees was shown with varying offset disparity with respect to the central test lines. The central test lines were shown with a relative disparity with respect to each other selected randomly from a set of seven disparities. The question was the same as before: Is the left line in front of or behind the right line? The two cases that were compared were (1) the offset disparity of the frame was crossed, and (2) the offset disparity of the frame was uncrossed. There was no mean shift observed for any of the observers, but there was a degrading of the threshold values. There was no significant difference in the threshold values for the crossed and uncrossed offset disparities cases for any of the observers tested. Figure 7 gives the average threshold for the crossed and the uncrossed offset disparity cases for three different observers.

#### Experiment 4

Is it feasible to seek an explanation of depth contrast only in terms of disparity and position of features being shown? If the observers showed judgements of relative depth significantly different from each other as the surrounding features were

changed, then it is likely that information other than what was shown in the position and disparity of the visual features was involved. The relative weights of the surrounding features was investigated by showing only a certain select portion of the rectangular frame. Figure 8 gives the front view of the case one of the stimulus that was shown. The right and the left eyes' views are superimposed. Filled circles mark the position of the dots in the right eye's view, and crosses mark the position of the dots in the left eye's view. If the feature was a line, then a solid line marks the position of the line in the right eye's view, and a dashed line marks the position of the line in the left eye's view. Only one of the disparity values out of the set of six or seven that were used is shown in the figure. The values obtained for the observed shifts for the stimulus shown is listed for four observers (Table 3). Observers B.H. and T.K. were given frame disparities of 8 degrees for the configuration, and observers V.M. and H.S. had frame disparities of 2 degrees. The test dots were 21.6 minutes apart for all the observers.

The observed shift is largest for the complete frame (Figure 8a), and smallest for the single dot (Figure 8i) on the side. There are significant differences among the observers. Comparison of the shifts obtained for configurations 8h and 8i indicates a doubling of the shift for observers B.H. and V.M., a 20% increase for T.K., and tripling for the observer H.S. Comparison of the configurations 8f and 8g similarly show that doubling the vertical dimension of the frame produces a

negligible change in the shift for observer B.H., but close to doubling for observer H.S. Comparing 8d and 8e shifts shows B.H. and T.K.'s shifts decreasing, while those of H.S. and V.M., increasing. All observers show a decrease in the shift between configurations 8b and 8c. These differences among the observers cannot be attributed to the difference in the frame disparity for the observers B.H. and T.K. and observers H.S. and V.M., since there are differences between the observers shown the same disparity. The authors were unable to find a simple scaling rule that could bring results of different observers into agreement.

#### DISCUSSION

The dynamics of stereo processing suggested by these results indicate that the interaction of disparity processing can extend over degrees of visual space, and from small disparities to fairly large ones. The relative depth of two foveally seen dots is influenced by the disparity of a frame degrees away, given presentation times of 100 milliseconds. The short presentation times rule out explanations requiring changing eye positions to minimize some function of corresponding points. The almost qualitative difference among the observers for different configurations of the frame rules out any attempt to explain stereo processing as a simple addition or subtraction of disparity of individual features over the whole of the visual scene. These individual differences probably indicate that the brain does not have a reliable technique for extracting depth from disparity alone. Frame disparities of twenty degrees or ten

degrees for either side of the frame do influence the depth judgement of the central test lines. The influence appears to saturate when the disparity of the frame exceeds a few degrees (i.e., one degree for either vertical side of the frame). Although the offset disparity of the frame does not change the relative depth of the central test features, it does degrade the stereoacuity of the central features. Although this degradation was measured only for a frame 6 degrees wide, the effect on stereo acuity is expected to extend to much larger distances.

Westheimer and Tanzman (1956) showed that observers could discriminate between 10 degrees crossed and 10 degrees uncrossed disparity for foveal bars. Richards and Foley (1971) showed that the detectability of disparity when the binocular components are presented to different hemispheres extended to 16 degrees, and probably to 32 degrees. This study agrees with their results in detecting such large disparities and demonstrates further that large disparity detectors interact and influence the disparity detectors responsible for the determination of the much smaller disparity difference of the central test lines; this interaction extends over large spatial distances. These results also suggest that this interaction is not likely to be a simple algebraic operation on the disparity and position of the features in the viewing field.

Luria's conclusion that stereoacuity deteriorates because of relative lack of stimulation in the periphery of the visual field is difficult to reconcile with these authors' experience of

measuring stereoacuity in a pitch dark room. The value of stereoacuity was found to be independent whether it was measured in the pitch dark room or in a dimly lit room in which the various objects (instrumentation, etc.) in the room were visible to the observer. In light of these data, it is likely that Luria's results were a consequence of the way he restricted the field of view. According to Luria, he placed "15.24 cm in front of the subject's eyes, a sheet of curved white bairbridge board with two circular holes of appropriate size; one hole was fixed, and the other could be moved horizontally to adjust for differences in interpupillary distances." The "frame disparity" provided by the edges of the holes, and not the lack of stimulation of the periphery, may have caused the deterioration of the threshold that Luria measured.

To summarize, these results indicate that (1) disparities of 20 degrees can be detected; (2) these large disparities interact with the units responsible for detecting much smaller disparities (on the order of seconds of arc); (3) these interactions extend over large visual angles; (4) these interactions do not require the observers to be either attending to or aware of the depth of the peripheral features; and (5) these interactions cannot be explained in terms of algebraic operations on disparity and position of features alone. These results indicate that there may not be a clean separation between local and global stereopsis. If there are local disparity units, then their responses are likely to be influenced by units responsible for

global stereopsis.

#### REFERENCES

- Gogel, W.C. (1956) Relative visual direction as a factor in relative distance perceptions. Psychol. Monographs 70, 1-19.
- Jaensch, E.R. (1911) Uber die Wahrnehmung des Raumes. Zeitschrift fur Psychologie Ergänzungsband 6.
- Koffka, K. (1935) Principles of Gestalt psychology. New York: Harcourt, Brace.
- Luria, S.M. (1969) Stereoscopic and resolution acuity with various fields of view. Science, N.Y. 164, 452-453.
- Mitchison, G.J. and Westheimer, G. (1984) The perception of depth in simple figures. Vision Res. 24, 1063-1073.
- Ogle, K.N. (1950) Researches in binocular vision. Philadelphia, W.B. Saunders.
- Pastore, N. (1964) Induction of stereoscopic depth effect. Science, N.Y. 144, 888.
- Richards, W. (1971) Anomalous stereoscopic depth perception. J. opt. Soc. Am. 61, 410-414.
- Richards, W. and Foley, J. M. (1971) Interhemispheric processing of binocular disparity. J. Opt. Soc. Am. 61, 3, 419-421.
- Wallach, H. and Lindauer, J. (1961) On the definition of retinal disparity. Psychologische Beitrage 6, 521-530.
- Werner, H. (1937) Dynamics in binocular depth perception. Psychol. Monog. 42, 129pp.
- Werner, H. (1938) Binocular depth contrast and the conditions of the binocular field. Am. J. Psychol. 51, 489-497.
- Westheimer, G. (1986) Spatial interaction in the domain of disparity signals in human stereoscopic vision. J. Physiol. 370, 619-629.
- Westheimer, G. and Tanzman, I.J. (1956) Qualitative Depth localization with Diplopic Images. Jour. Opt. Soc. Amr. 46, 116-117.

FIGURE 1

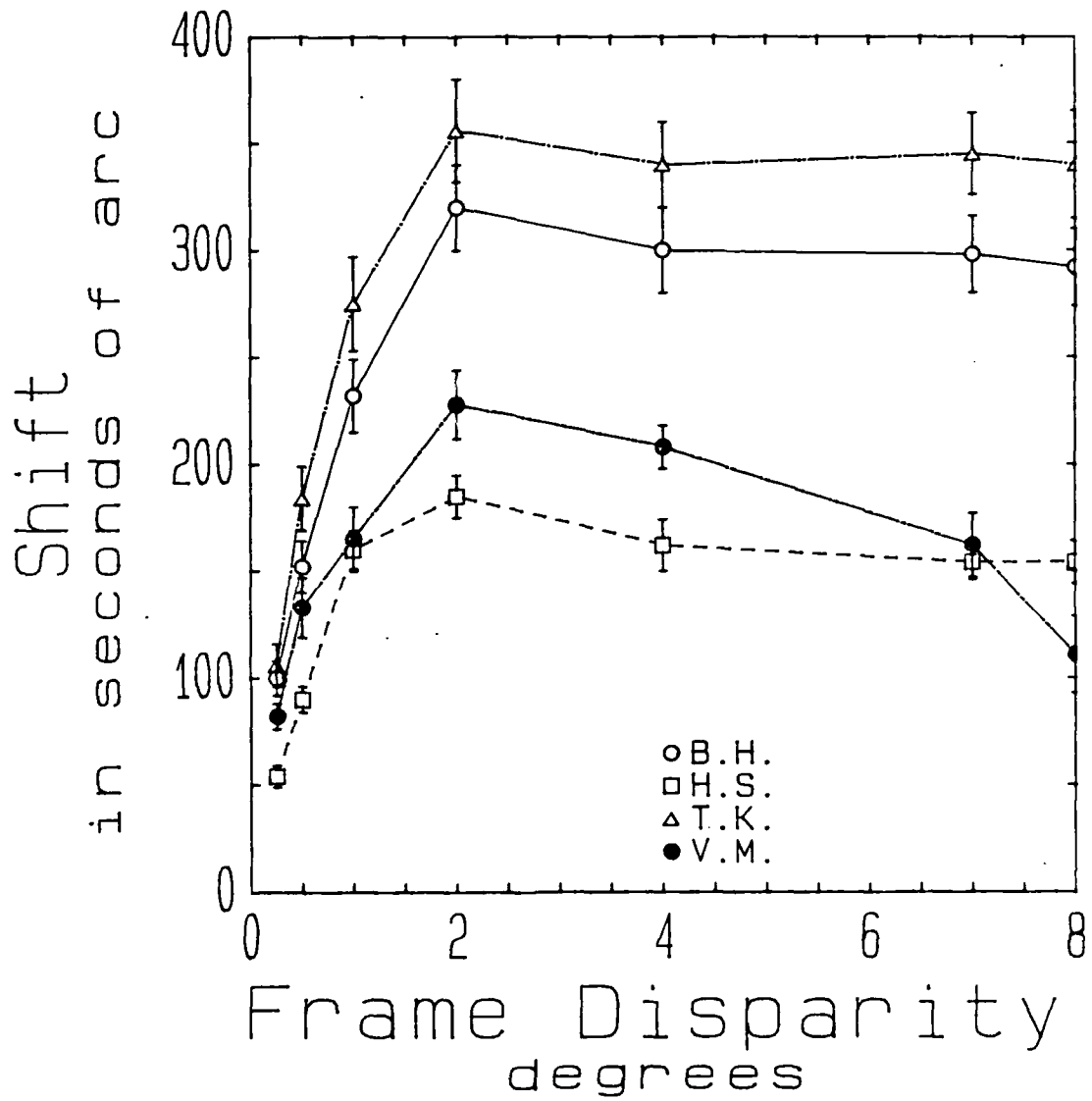


FIGURE 2

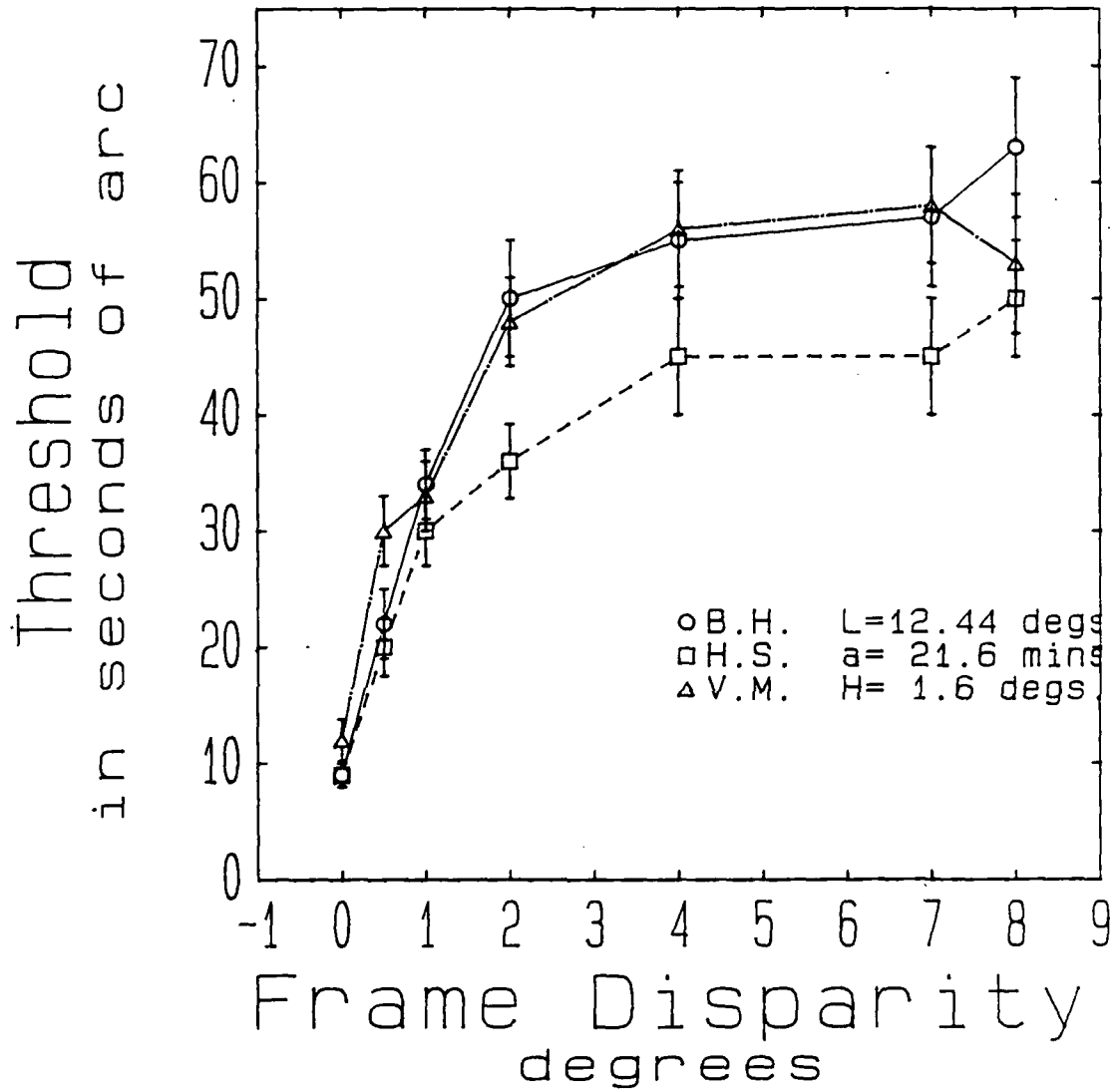


TABLE 1. Threshold and shifts for three different observers for a frame approximately 15 degrees wide and disparity of 3.4 degrees when the observers are paying no attention to the disparity of the frame ( $N_R > 0.9$ ), and thresholds when observers were trying to respond only to the positive frame disparity ( $N_R < 0.1$ ).

OBSERVER	$N_R > 0.9$		$N_R < 0.1$
	SHIFT (seconds)	THRESHOLD (seconds)	THRESHOLD (seconds)
B.H.	110 +/- 15	35 +/- 10	107 +/- 28
H.S.	52 +/- 12	38 +/- 4	> 120
V.M.	52 +/- 18	40 +/- 6	> 120

$$N_R = \frac{\text{Number of responses for negative frame disparity}}{\text{Number of responses for positive frame disparity}}$$

TABLE 2. Fraction of responses when left line called closer for frame 51 degrees wide and 40 degree high. The test lines always had zero relative disparity.

F R A M E DISPARITY	OBSERVER			
	B.H.	H.S.	T.K.	V.M.
MONOCULAR	0.50	0.53	0.47	0.44
NO FRAME	0.57	0.47	0.44	0.53
0°	0.40	0.47	0.50	0.57
+12°	0.17	0.13	0.10	0.23
-12°	0.77	0.70	0.83	0.67
+20°	0.20	0.53	0.17	0.27
-20°	0.80	0.53	0.83	0.77

FIGURE 3  
Observer: B.H.

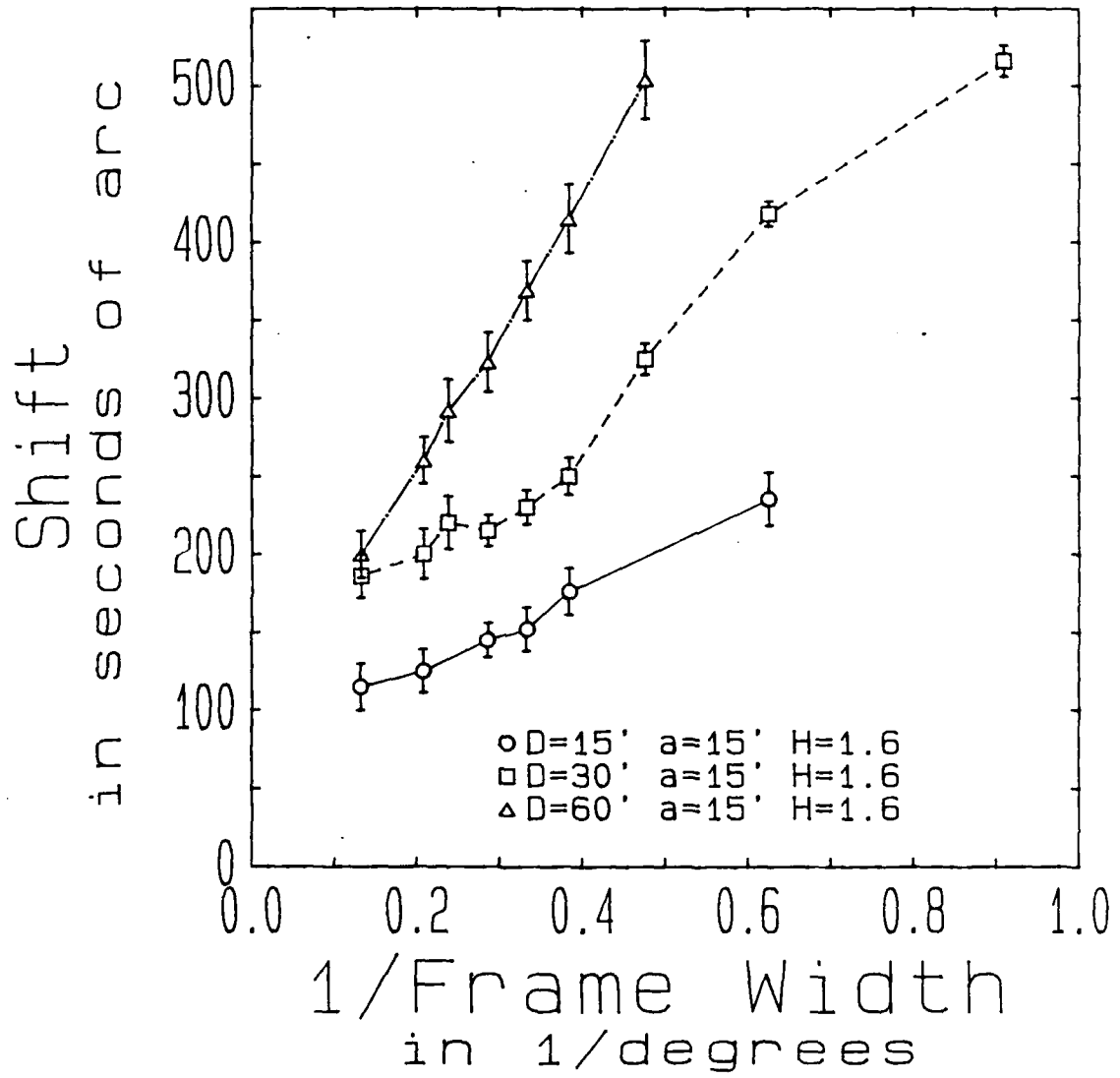


FIGURE 4  
Observer: V.M.

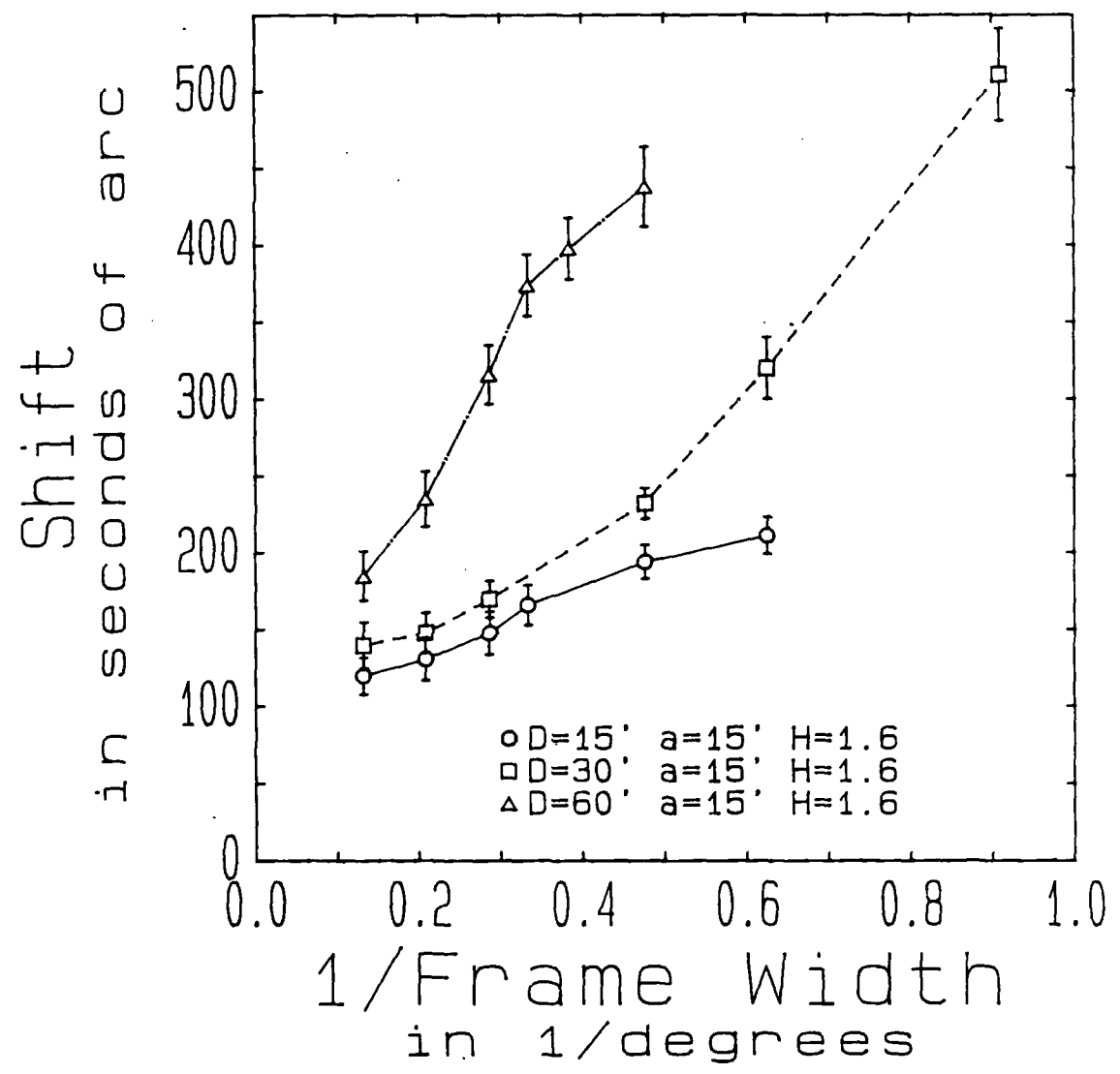


FIGURE 5  
Observer: B.H.

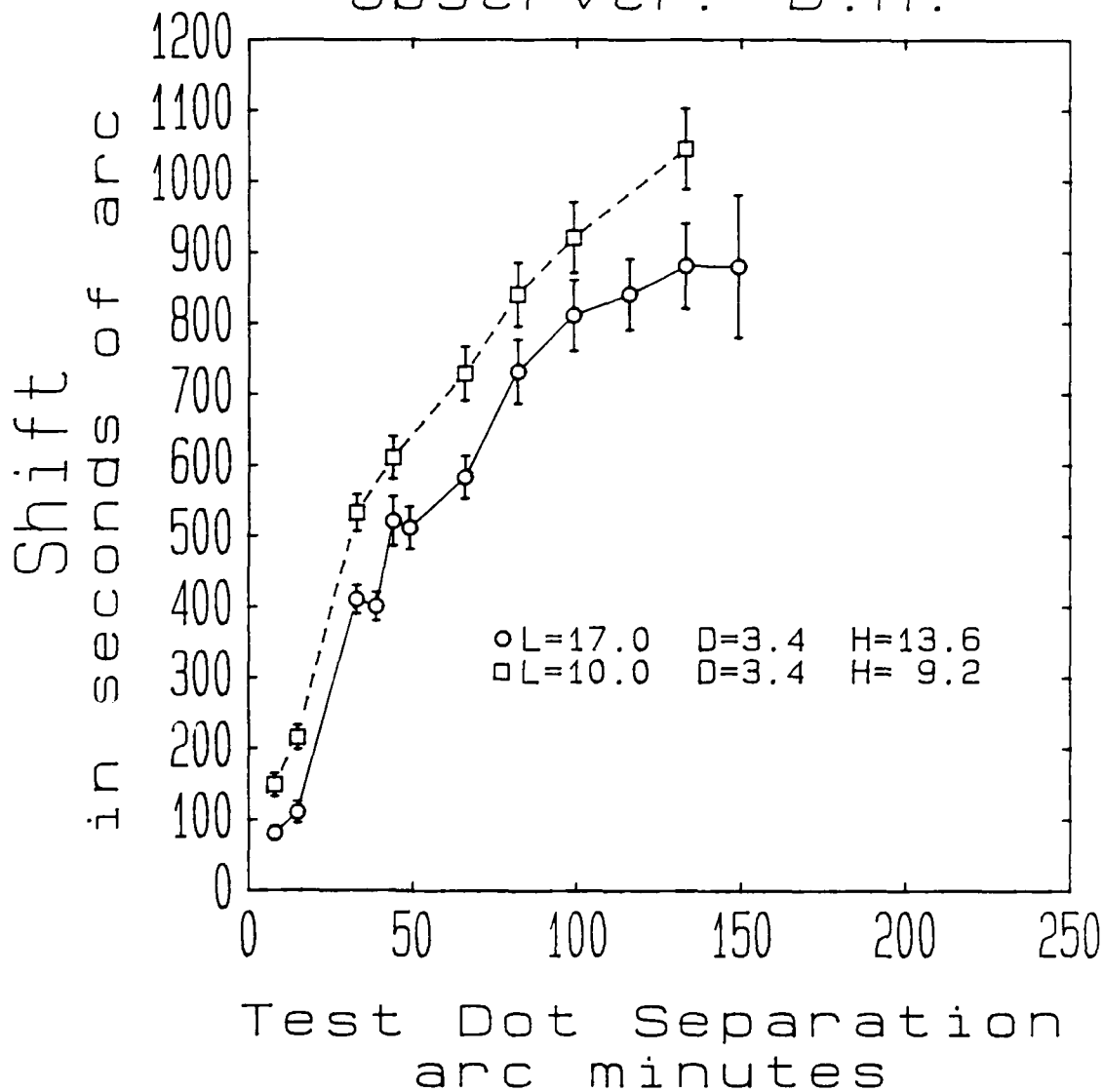


FIGURE 6  
Observer: H.S.

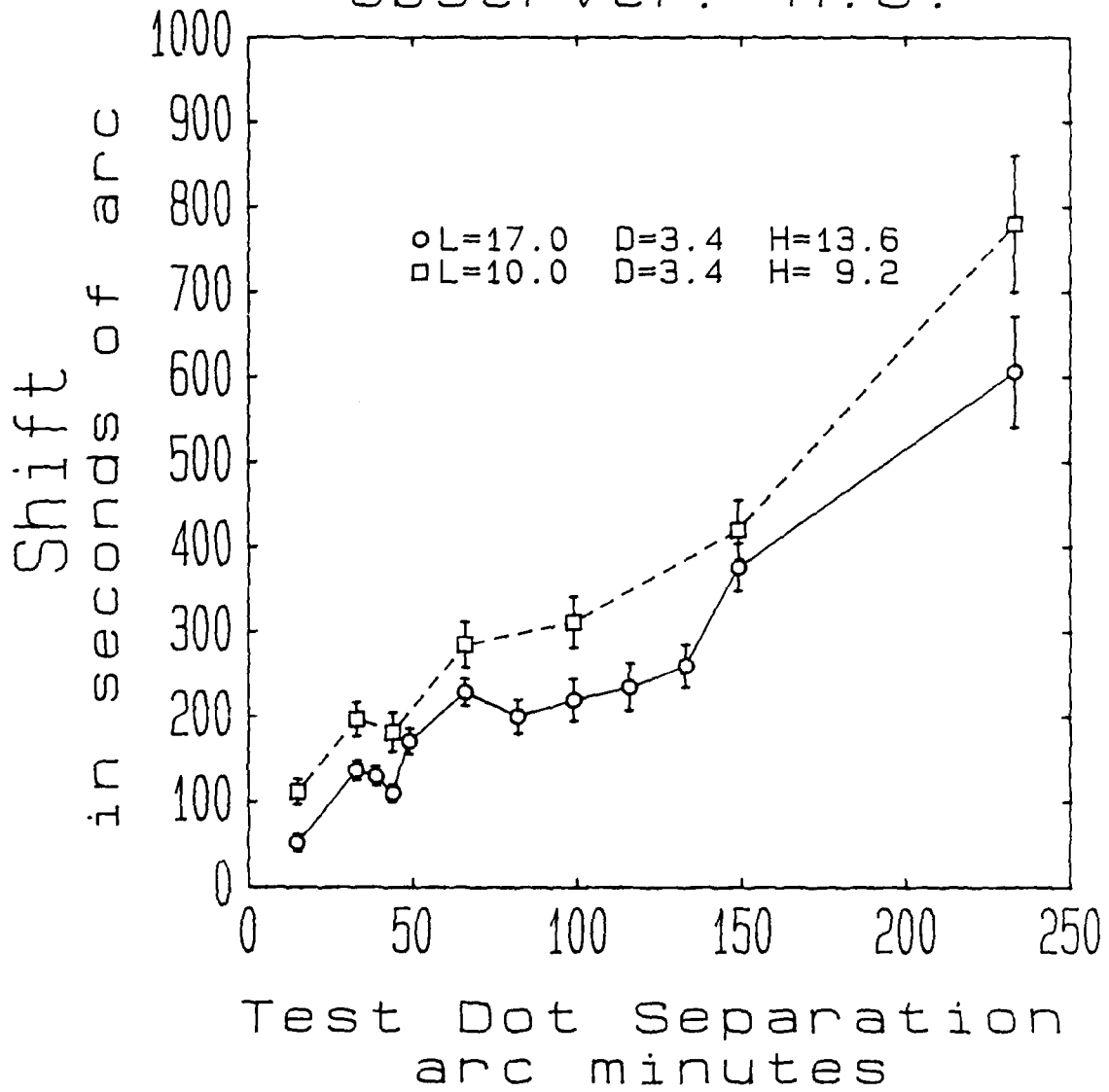
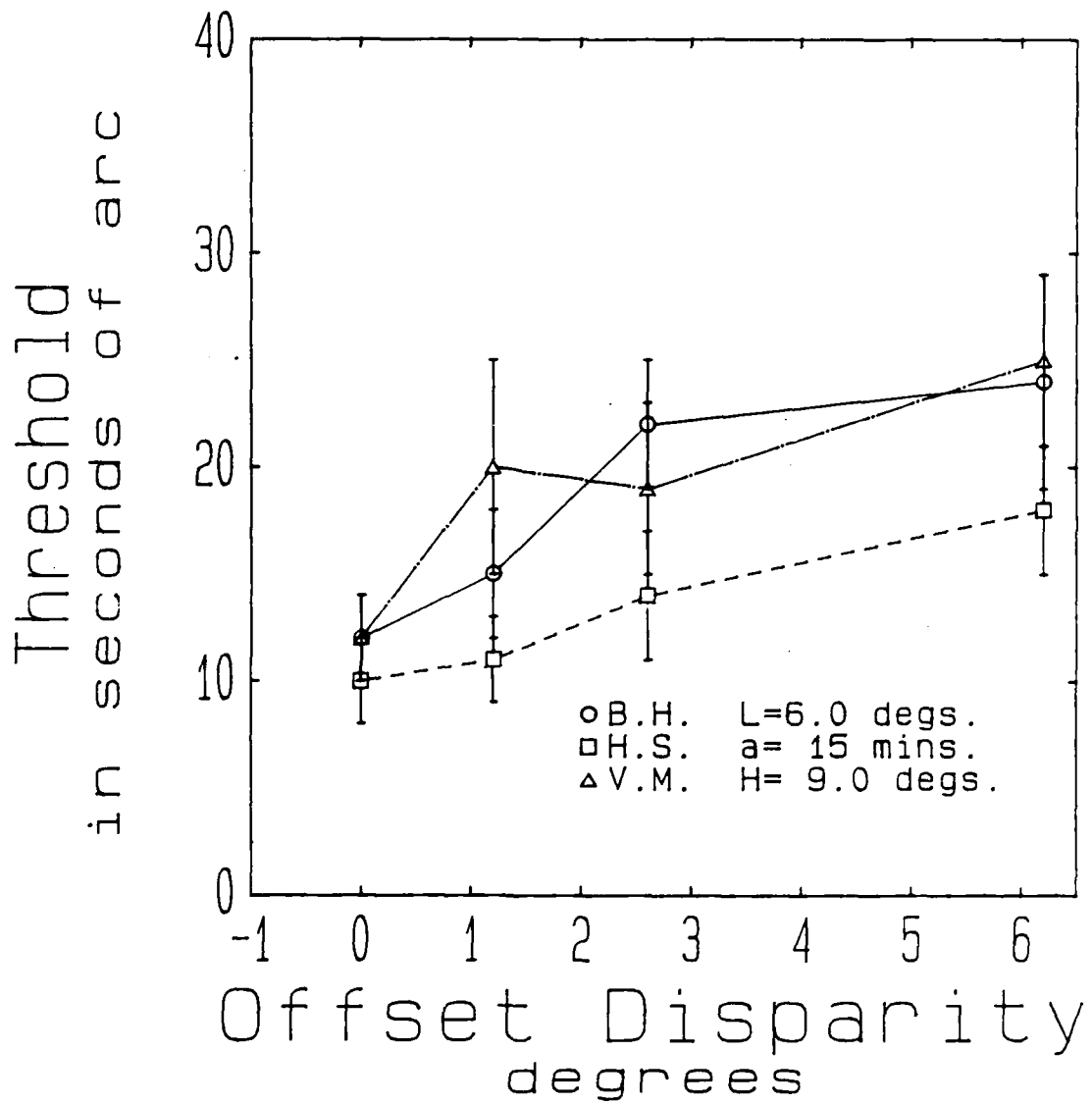


FIGURE 7



# FIGURE 8

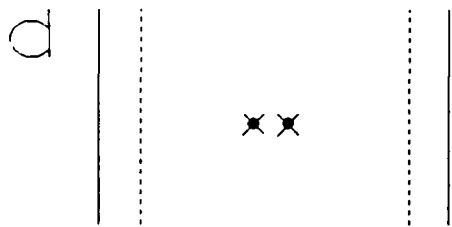
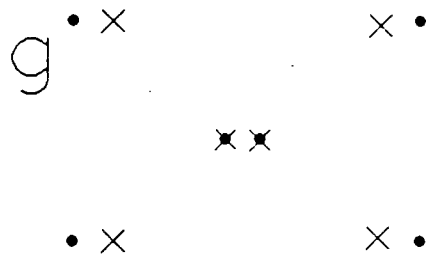
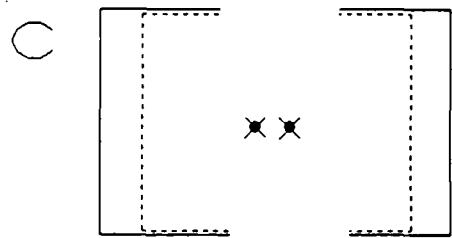
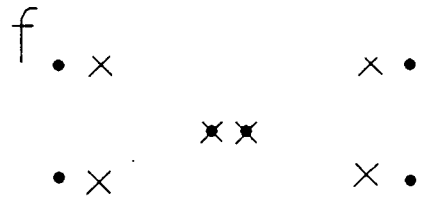
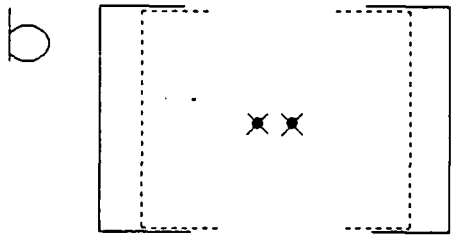
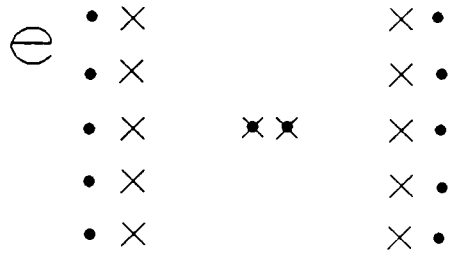
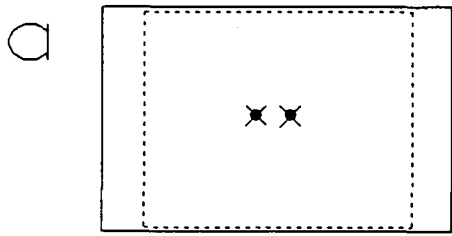


TABLE 3: The shift in arc seconds is listed for stimuli shown in Figure 26A through 26I. Frame disparity for B.H. and T.K. was 8° and for H.S. and V.M. was 2°.

STIMULUS	OBSERVERS			
	B.H.	T.K.	H.S.	V.M.
5A	240+/-30	364+/-15	184+/-18	228+/-10
5B	226+/-30	352+/-13	162+/-18	174+/-35
5C	123+/-25	266+/-14	93+/-10	140+/-13
5D	130+/-10	266+/-14	40+/-10	150+/-15
5E	93+/-15	167+/-13	125+/-10	222+/-10
5F	88+/-15	205+/-12	90+/-10	188+/-10
5G	95+/-15	148+/-10	46+/-10	152+/-8
5H	60+/-15	166+/-15	92+/-10	146+/-10
5I	30+/-10	133+/-16	28+/-10	60+/-10