MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo No. 1073 September 1988
C.B.I.P Memo No. 33

Seeing "Ghost" Solutions in Stereo Vision

Daphna Weinshall

Abstract

A unique matching is a stated objective of most computational theories of stereo vision. This report describes situations where humans perceive a small number of surfaces carried by non-unique matching of random dot patterns, although a unique solution exists and is observed unambiguously in the perception of isolated features. We find both cases where non-unique matchings compete and suppress each other and cases where they are all perceived as transparent surfaces. The circumstances under which each behavior occurs are discussed and a possible explanation is sketched. It appears that matching reduces many false targets to a few, but may still yield multiple solutions in some cases through a (possibly different) process of surface interpolation.

© Massachusetts Institute of Technology (1988)

This report describes research done at the Massachusetts Institute of Technology within the Artificial Intelligence Laboratory and the Center for Biological Information Processing in the Department of Brain and Cognitive Sciences and Whitaker College. The Center's research is sponsored by a grant from the Office of Naval Research (ONR), Cognitive and Neural Sciences Division; by the Alfred P. Sloan Foundation; and by the National Science Foundation. The Artificial Intelligence Laboratory's research is sponsored by the Advanced Research Projects Agency of the Department of Defense under Army contract DACA76-85-C-0010 and in part by ONR contract N00014-85-K-0124. Dr. Weinshall is supported by a Chaim Weizmann Postdoctoral Fellowship from the Weizmann Institute of Science.

Seeing "Ghost" Solutions in Stereo Vision

Biological stereo vision computes the depth of an object from the disparity in position of points matched between the left and right eye images of the object. Matching is a difficult computation, which humans apparently do well. Figure 1a illustrates this difficulty. There are four possible matches of the two points in the right image and the two in the left image. Humans see only the two matches (dark circles in figure 1a) that are order-preserving, namely, matching left to left and right to right (see [1] and [2]). Furthermore, stereograms with many random dots in each image (e.g. figure 3a) increase the ambiguity of the matching task: each point has many false targets in the other image. A matching algorithm is usually required to resolve such ambiguities and obtain a unique correspondence for each feature or patch (e.g. [3] and [4]). This report describes situations where humans perceive a small number of surfaces carried by non-unique matching of random dot patterns, although a unique solution exists and is observed unambiguously in the perception of isolated features. It appears that matching reduces many false targets to a few, but may still yield multiple solutions.

Braddick (1978, unpublished results) has extended Panum's limiting case, where one eye sees one vertical line and the other sees two (figure 1b), by copying a random pattern once in one eye and twice in the other with a horizontal gap of few pixels. In this case humans perceive two planes, the upper one transparent, and this depth perception is more robust than the single line limiting case. Grimson describes this experiment in his book ([5]), and argues that this result is consistent with a unique matching even though multiple matching takes place: if matching is done simultaneously from each image to the other, the matching from the double image to the other one is indeed unique. Note, also, that the perceived depth in the extended Panum's limiting case is consistent with the single feature perception. We have examined this case further and observed the following: first, the disparities of the two planes in the dense stereogram can be constant (two planes) or vary in any continuous way, like cos or sin (see figure 2a). Also, the same pattern can be copied more than once (e.g., three copies in three different disparities), in which case more than two planes are perceived (e.g. three, see figure 2b), though it becomes more difficult to make sense out of the stereogram. The effect is visible even for a very low density of points (0.001).

A similar extension of the double nail illusion, where the two eyes see two vertical lines with possibly different horizontal spacing (figure 1a), can be made to random dot stereograms. Thus the same random pattern in a middle square is copied twice in each image, with a horizontal gap of G_r pixels in the right image and G_l in the left image, see figure 3a and figure 4a ($G_r = 0$ or $G_l = 0$ give the previous case). Each pair of points in this configuration has four possible matchings, two mutually exclusive pairs if matching is unique (figure 3b). For single features, like lines, only the two matchings marked in figure 3b

with full circles are seen, as has been reported before (double nail illusion, see [6]). Surprisingly, in the extended case, observers with good stereo vision see four planes (with the help of vergence and memory). They are able to judge the depth of the "ghost" planes correctly, choosing the correct depth from a multiple choice scale. This perception, though, takes time to build and some concentration. Some people with reasonable stereo vision see only three planes. Others (including the author) do not see the "ghost" planes, but most seem to improve with practice. On the other hand, Prazdny's stereo algorithm ([7]) for example, an algorithm designed to handle transparent surfaces, will detect only two transparent surfaces in this case, those marked by full circles in figure 3b. If features, like lines, are added to the stereogram, usually only two of them are seen in the two middle planes as expected from the double nail illusion experiments. (Note that Kroll and van de Grind [6] also found that one observer occasionally saw a third match in the "double-nail" experiment.) This configuration, like the previous one, is not restricted to fronto-parallel planes only: one can construct tilted planes like in figure 4b, and other surfaces. We have tried similar configuration for motion, that is, images like in figure 3a are seen one after the other instead of in stereo. However, we couldn't detect similar effects, only two moving planes are seen.

This result seems to suggest that all disparities with sufficient support give rise to the perception of a distinct transparent surface. For convenience, we define the support of a given disparity at a given pixel to be the value of the correlation function between the two images. The correlation for disparity d is computed between a window around the pixel in one image (W in the caption of figure 6) and a window around the same location in the other image translated by d. A sufficient support may be a correlation value sufficiently above random, and the corresponding disparity will be called henceforth a "solution". To check the above hypothesis we use a special case of the same stereogram where $G_r = G_l$, see figure 5. In this case there are three possible solutions (disparities with obvious peaks in the support function), and thus three transparent planes can be seen in analogy to the previous case. One "strong" solution where all the points in one image are matched to all the points in the other image, and two "weak" solutions, in which half the points in one image are matched to half the points in the other image (see figure 6b). Surprisingly, in this case only the "strong" coherent solution is seen, and this perception is quite robust.

It now seems that only solutions with approximately equal support, i.e. comparable maxima in the correlation function, are detected, whereas weaker solutions are suppressed. To check this hypothesis we use the initial stereogram where $G_r \neq G_l$, but we double the amount of points in one plane, so that its peak in the correlation function will be twice as high as the others (figure 6c). In this case all four solutions are still seen, but the "strong" dense solution is seen darker. Even if the number of points in one surface is quadrupled, the other surfaces are visible. Figure 7a demonstrates a worse case from our

point of view, where we add new unambiguously matched points to the two external surfaces, which together constitute a complete solution of the matching problem. Consequently the support of these two solutions is doubled. Still, one can see from the figure that the other two solutions are readily seen. We conclude that it is not simply the relative strength (value of support) of the solutions that determines which of them will be perceived. If to a stereogram with $G_l = G_r$, where initially only one plane has been seen, we add an equal number of unmatched random dots to each image, the suppressed solutions may sometime (at least partially) reappear (see figure 7b). If we add correlated points to both images, forming a new plane, it and the dominant plane are seen (see figure 7c). Some of the suppressed solutions may then reappear. These effects have not been studied thoroughly enough to state conclusive results, though.

One way by which surface interpolation could lead to the perception of multiple transparent surfaces is if one or the other unique matching is chosen locally and randomly, and interpolation then smoothes across the "holes". But depth determination for isolated features is not random, the ordered matching is consistently seen. Moreover, it seems that random assignment can not explain the qualitative difference between the case with four possible solutions, when four planes can be seen (figure 3), and the case with three possible solutions, when only one opaque plane is always seen (figure 5). Finally, random assignment does not explain the perception of a single surface in periodic stereograms, which are characterized by repetitive patterns, like in wall-paper designs. Multiple possible matchings of the images exist here, each involving an almost complete matching of all points in the two images and creating different depth perceptions (see [8], pp. 187, for a summary). However, the different solutions are seen, but not simultaneously, one has to "flip" from one to the other using eyes vergence and cues from the surroundings (see figure 6.2-2* in [8]). Thus it seems possible that locally multiple matchings are detected, maintained and manipulated through the process of interpolation, even though eventually one unambiguous match is chosen for each local feature.

A possible surface interpolation procedure that agrees with the above results will now be briefly discussed, though by no means this is the only possible explanation. Initially surfaces are constructed that take into account all possible matches of all the pixels. We define the support set of a surface to be the set of points, say within a window of size W, that contribute to its construction. We assume W is large enough so that random matches corresponding to all disparities are abundant. A surface is maintained if it has a sufficiently large support set that is not included in the support set of a different surface. With simple transparencies (see [7] and [9]) all possible solutions will survive, as their support sets are disjoint. In figure 3 all four solutions will survive due to random matches, whereas in figure 5 the "strong" solution will dominate. Finally, all the solutions in a periodic stereogram have overlapping support sets (of all the points in the window) almost everywhere. In agreement with the

observations, the above scheme predicts that each solution can suppress the other, depending upon additional cues or vergence (see [10]).

The main implication of the above results is the conclusion that the resolution of ambiguities is not as simple as practically all computational theories of stereo vision assume (e.g. [4]). A unique solution is a stated objective of these theories. The above results will require their extension to describe surface interpolation which extends multiple possible matchings. Possibly, feature matching and disambiguation and surface interpolation are different processes (see similar suggestion by Mitchison in [11]).

Acknowledgements: I thank Scott Kirkpatrick for his invaluable help in discussion, advice and criticism. For very helpful comments and advice I thank Graeme Mitchison, Tomaso Poggio and Heinrich Bülthoff.

References

[1] Alan L. Yuille and Tomaso Poggio. A generalized ordering constraint for stereo correspondence. Technical Report AIM–777, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1984.

[2] P. Burt and B. Julesz. The disparity gradient limit for binocular fusion: an answer to J D Krol and W A van de Grind. *Perception*, 11:621–624, 1982.

[3] David Marr and Tomaso Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London B*, 204:301–328, 1979.

[4] David Marr and Tomaso Poggio. Cooperative computation of stereo disparity. *Science*, 194:283–287, 1976.

[5] W. Eric L. Grimson. *From Images to Surfaces.* MIT Press, Cambridge, Mass., 1981.

[6] J. D. Kroll and W. A. van de Grind. The double nail illusion: experiments on binocular vision with nail, needles, and pins. *Perception*, 9:651–669, 1980.

[7] K. Prazdny. Detection of binocular disparities. *Biological Cybernetics*, 52:93–99, 1985.

[8] B. Julesz. *Foundations of cyclopean perception.* The University of Chicago Press, Chicago, 1971.

[9] J. E. W. Mayhew S. B. Pollard and J. P. Frisby. A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14:449–470, 1985.

[10] S. P. McKee and G. J. Mitchison. The role of retinal correspondence in stereoscopic matching. *Vision Res.*, 28:1001–1012, 1988.

[11] G. J. Mitchison. Planarity and segmentation in stereoscopic matching. Technical Report 8, King's College Research Centre, 1988.

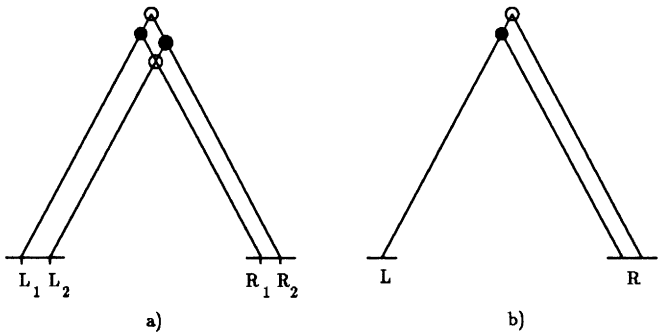
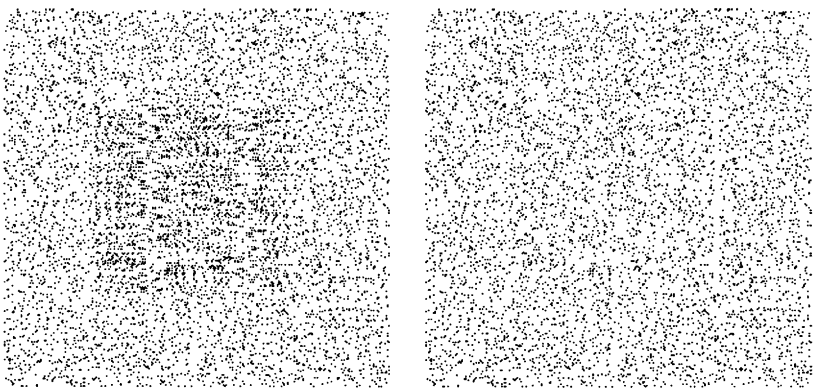
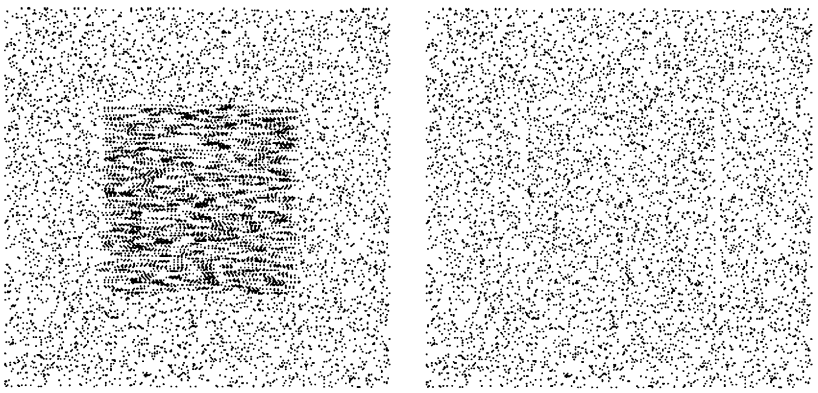


Figure 1: a) Ambiguous matching or the double nail illusion: there is one "natural" order preserving matching of L_1 to R_1 and L_2 to R_2. This solution is marked in full circles, and is always perceived, even when the correct matching is L_1 to R_2 and L_2 to R_1 (the "ghost" solution). This effect is demonstrated in an experiment where two nails are put one behind the other with respect to the viewer. A depth illusion is then created where the two nails are seen one beside the other in the same depth ([6]). b) Panum's limiting case: here one eye sees two lines and the other sees only one. Most people perceive the two lines differing in depth, as if they match the single line in one image simultaneously with the two lines in the other (if the distance between them is sufficiently small).



a)



b)

Figure 2: Variations on the extended Panum's limiting case: a) cos and its mirror image, b) three planes.

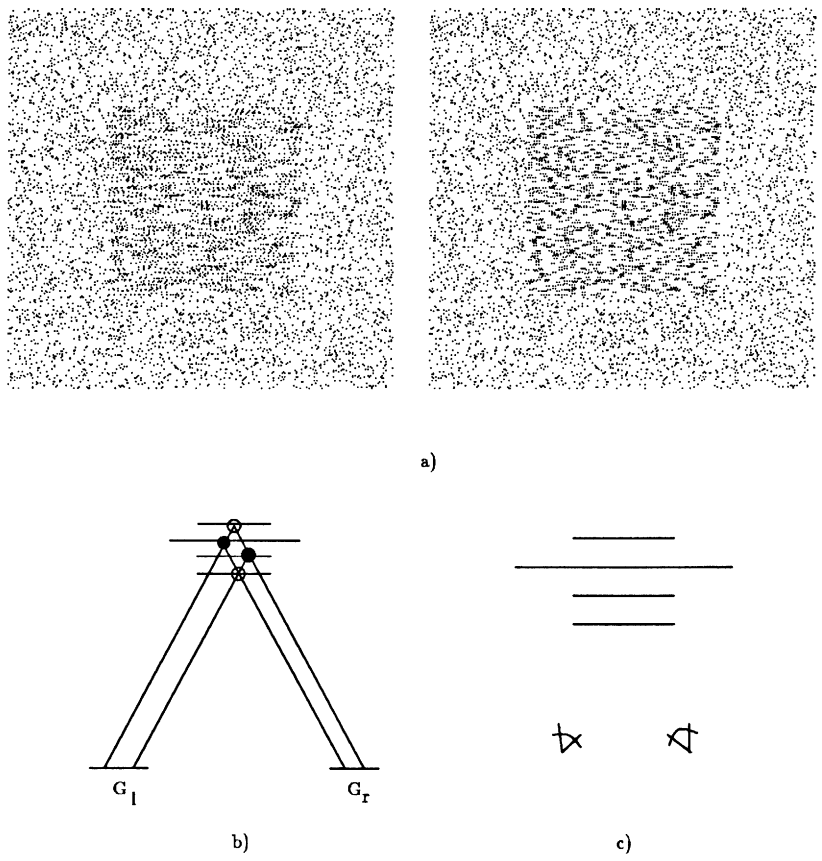
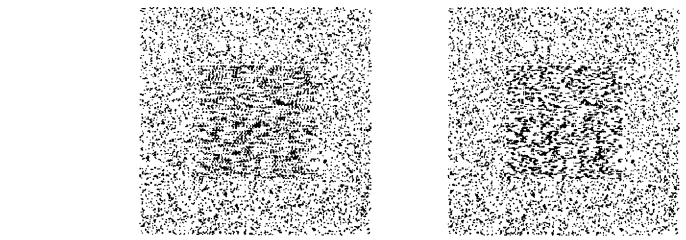
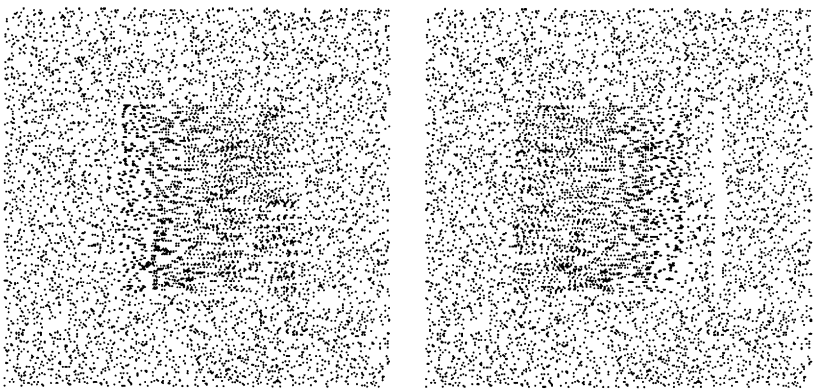


Figure 3: a) An ambiguous stereogram with $G_r = 2$ pixels and $G_l = 4$ pixels. After some staring one can see four planes: two in front of the background, the background (all three transparent), and one behind the background. The deepest plane is usually the most difficult to see, and it often helps to slightly diverge the eyes to capture it. Note the two lines ("nails") in the middle of the stereogram, which are usually seen on the two middle surfaces only, and which are hard to flip to the other surfaces. Because it takes some time for the impression to build, it is recommended to use a stereo viewer about 5.5 inches high. b) A graphic illustration of the projections of two points from the stereogram in a. The two solutions that are mutually exclusive if matching is unique are separately marked by filled and hollow circles. c) The depth profile of the four possible solutions of the stereogram in a.



a)



b)

Figure 4: Variations on figure 3a: a) smaller disparities, 5-6 minutes of arc only, are separating two nearby planes, so that it is easier to see the four planes together but more difficult to distinguish them; b) two fronto-parallel planes and two tilted planes that create an "X" shape between them in depth can be seen.

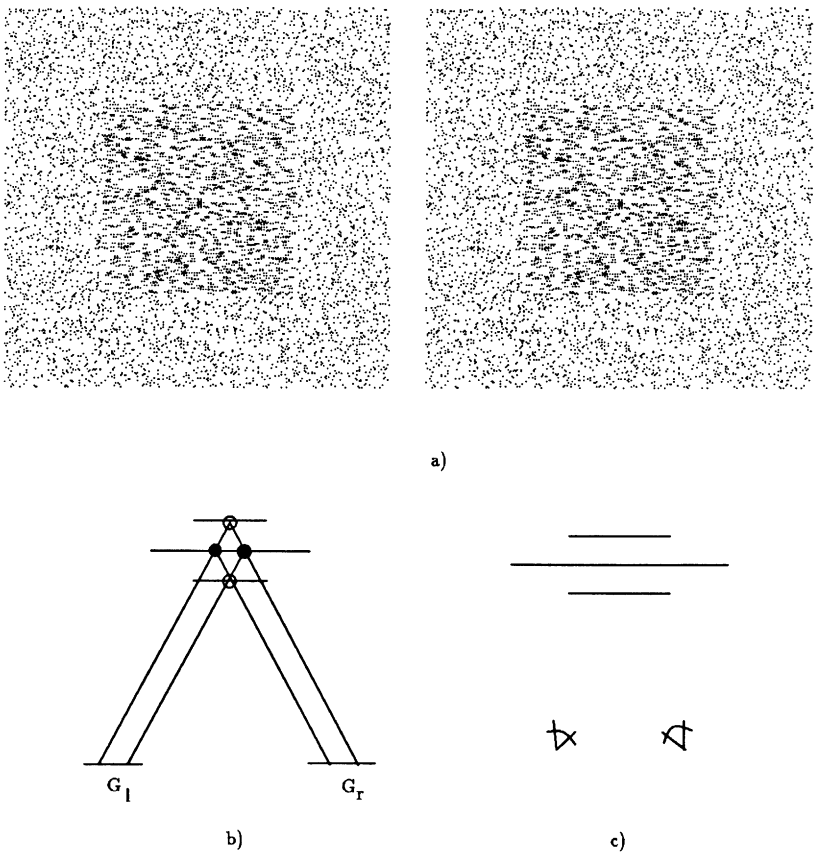


Figure 5: a) An ambiguous stereogram with $G_r = G_l = 2$ pixels. Here only one opaque plane is seen, the background, and no vergence can help detect the other two planes (one above the background and one below it). b) A graphic illustration of the projections of two points from the stereogram in a. c) The depth profile of the three possible solutions of the stereogram in a.

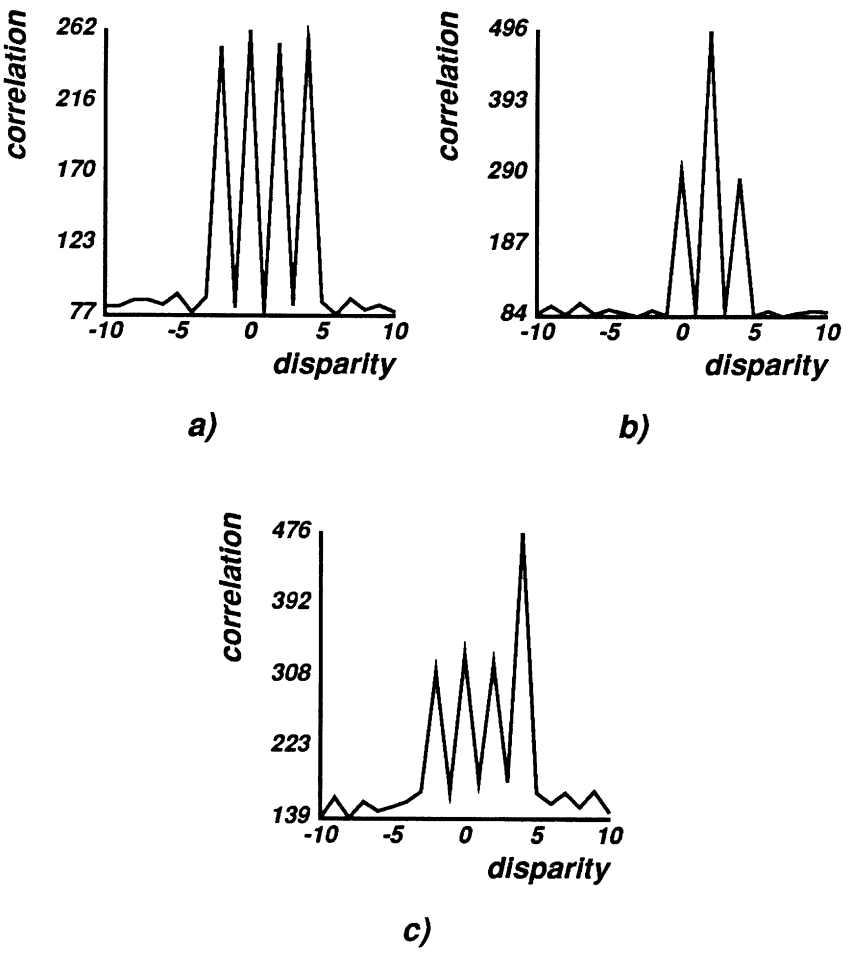
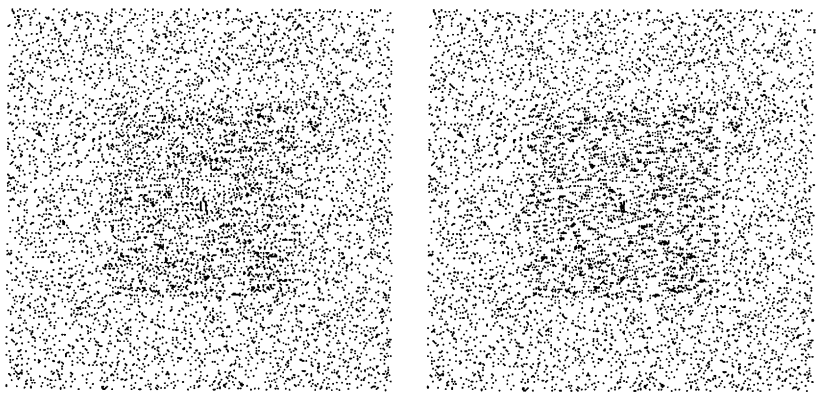
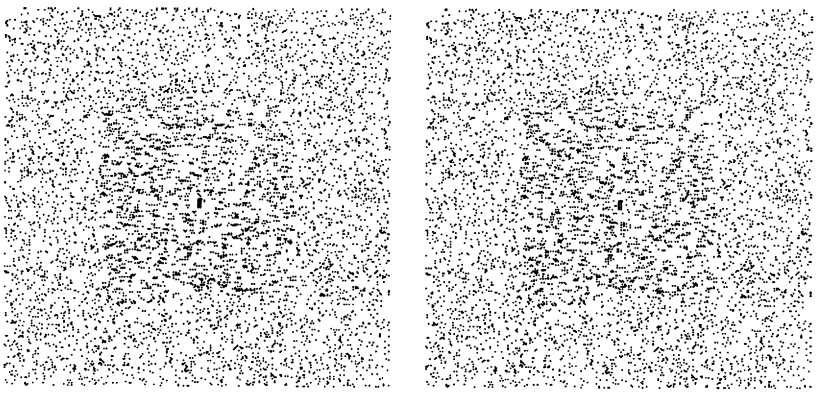


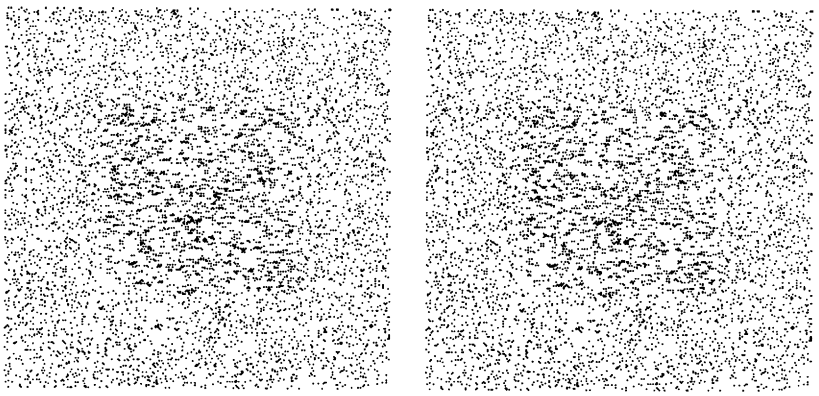
Figure 6: The correlation between the left and the right images at their center: $\sum_i^W \sum_j^W f(x_i^R, y_j^R) f(x_i^L + D, y_j^L)$. The X-axis is the disparity D: a) stereogram of figure 3, b) stereogram of figure 5, c) stereogram of figure 3 with additional points to one solution.



a)



b)



c)

Figure 7: a) like figure 3a, where the number of points in the two external planes have been doubled with new unambiguous points; b) like figure 5a, where the number of points in each image is doubled with new random unmatched points (noise); c) like figure 5a, with an additional new uncorrelated plane at disparity 4, double the disparity of one of the suppressed planes (2 and -2).

CS-TR Scanning Project
Document Control Form

Date : 4/19/95

Report # AIM-1073

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 12(18 images)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages(by page number):_____

Photographs/Tonal Material (by page number):_____

Other (note description/page number):

Description :	Page Number:
IMAGE MAP: (1) UN#'ED TITLE PAGE	
(2-12) PAGES #'ED 1-11	
(13-15) SCANCONTROL, DOD(2)	
(16-18) TRGTS (3)	

Scanning Agent Signoff:

Date Received: 4/19/95 Date Scanned: 4/20/95 Date Returned: 4/27/95

Scanning Agent Signature: Michael W. Cook

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AI Memo 1073	2. GOVT ACCESSION NO. A203581	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Seeing "Ghost" Solutions in Stereo Vision		5. TYPE OF REPORT & PERIOD COVERED memorandum 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR*(s)* Daphna Weinshall		8. CONTRACT OR GRANT NUMBER*(s)* N00014-85-K-0124 DACA76-85-C-0010
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd Arlington, Virginia 22209		12. REPORT DATE September 1988 13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* Office of Naval Research Information Systems Arlington, Virginia 22217		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

stereo vision
periodic sterograms
unique matching

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A unique matching is a stated objective of most computational theories of stereo vision. This report describes situations where humans perceive a small number of surfaces carried by non-unique matching of random dot patterns, although a unique solution exists and is observed unambiguously in the perception of isolated features. We find both cases where non-unique matchings compete and suppress each other and cases where they are all perceived as transparent surfaces. The circumstances

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
AI Memo 1073

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

Seeing "Shape" Solutions [illegible]

under which each behavior occurs are discussed and a possible explanation is sketched. It appears that matching reduces many false targets to a few, but may still yield multiple solutions in some cases through a (possibly different) process of surface interpolation.

REPORT NUMBER

7. AUTHO[illegible]

Daphn[illegible]

[illegible] NUMBER(s)

[illegible]24

DACA76-85-C-0010

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Massachusetts 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Advanced Research Projects Agency
1400 Wilson Blvd
Arlington, Virginia 22209

12. REPORT DATE
September 1988

13. NUMBER OF PAGES
11

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)
Office of Naval Research
Information Systems
Arlington, Virginia 22217

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

stereo vision
periodic sterograms
unique matching

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**

Scanned

Date: 4/20/1995

M.I.T. Libraries
Document Services