



# HHS Public Access

Author manuscript

*Neurocomputing*. Author manuscript; available in PMC 2016 January 05.

Published in final edited form as:

*Neurocomputing*. 2015 January 5; 147: 485–491. doi:10.1016/j.neucom.2014.06.037.

## Motion sequence analysis in the presence of figural cues

Pawan Sinha<sup>1</sup> and Lucia M. Vaina<sup>1,2,3</sup>

<sup>1</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA

<sup>2</sup>Departments of Biomedical Engineering, Neuroscience and Neurology, Boston University

<sup>3</sup>Department of Neurology, Harvard Medical School, Boston, MA

### Abstract

The perception of 3D structure in dynamic sequences is believed to be subserved primarily through the use of motion cues. However, real-world sequences contain many figural shape cues besides the dynamic ones. We hypothesize that if figural cues are perceptually significant during sequence analysis, then inconsistencies in these cues over time would lead to percepts of non-rigidity in sequences showing physically rigid objects in motion. We develop an experimental paradigm to test this hypothesis and present results with two patients with impairments in motion perception due to focal neurological damage, as well as two control subjects. Consistent with our hypothesis, the data suggest that figural cues strongly influence the perception of structure in motion sequences, even to the extent of inducing non-rigid percepts in sequences where motion information alone would yield rigid structures. Beyond helping to probe the issue of shape perception, our experimental paradigm might also serve as a possible perceptual assessment tool in a clinical setting.

### Keywords

motion; form-cues; structure-from-motion; stroke patients; 3D-structure-from-motion; figural cues; motion-impaired-patients

### Introduction

Motion of objects in the environment induces complex transformations in their images. The human visual system can recover the 3-D structure of the viewed objects and their motion in space by interpreting these image transformations [5, 36]. As early as 1953, Wallach and O'Connell demonstrated humans' capacity to interpret structure from motion while studying what they termed the 'kinetic depth effect' [41]. In their experiments, an unfamiliar object

© 2014 Elsevier B.V. All rights reserved.

Correspondence should be addressed to: Professor Lucia M. Vaina, Brain and Vision Research Laboratory, Department of Biomedical Engineering, Boston University, Boston, MA, 02215, USA, Ph: 617-353-2455; Fax: 617-353-6766; vaina@bu.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

was rotated behind a translucent screen with its shadow being observed from the other side of the screen. In most cases, the viewers were able to describe correctly the 3-D shape of the hidden object and its motion, even when each static shadow projection of the object was unrecognizable and contained no 3-D information.

Any vision system that attempts to compute 3-D structure from motion must contend with the problem that the recovery of structure is under-constrained; there are infinitely many 3-D structures consistent with a given pattern of motion in the changing 2-D image. Additional constraint is required to establish a unique interpretation. Computational studies have used the rigidity assumption to derive a unique 3-D structure and motion; they assume that if it is possible to interpret the changing 2-D image as the projection of a rigid 3-D object in motion, then such an interpretation should be chosen [3,7, 11, 14,18,22,23,24,25,26, 35,36,37,45]. The rigidity assumption was suggested by perceptual studies that described a tendency for the human visual system to choose a rigid interpretation of moving elements [9,15,16,41].

The rigidity assumption has proven to be a powerful constraint, one that appears sufficient to explain how the human visual system solves the structure from motion problem in general settings. However, some interesting perceptual effects suggest that this notion of sufficiency might need to be revisited, at least insofar as modeling human performance is concerned. According to the rigidity assumption, a rigid object in motion should necessarily be perceived as rigid. But, a few studies have reported instances where displays of rigid objects in motion can give rise to the perception of distorting objects [4,6,10,42,46]. As detailed below, we suggest that these breakdowns of rigidity perception hint at a significant contribution from figural shape cues in the perceptual analysis of dynamic sequences. In this paper, we develop the hypothesis of a role for figural cues in sequence analysis, and present experiments designed to test this hypothesis.

The remainder of this paper is organized as follows: we first briefly review the current state of research related to the recovery of 3-D structure from motion and static image attributes; subsequently we present a hypothesis regarding interactions between shape information derived using motion and that derived from figural cues; and in Methods we describe the psychophysical experiment and in Results we present evidence from normal observers as well as stroke patients with perceptual impairments in support of the basic thesis of this paper.

### **Possible strategies for the analysis of dynamic sequences**

In a laboratory setting, it is possible to generate motion stimuli that cleanly dissociate between static and dynamic sources of 3-D information. For instance, the random clusters of dots often used in perceptual studies of recovering structure from motion are carefully controlled so that no single frame has any discernible static organization that may provide hints about the 3-D configuration of the dots. This has been the dominant paradigm of structure from motion research so far. Since there is no static 3-D information in such displays, the question of how statically and dynamically derived 3-D shape estimates interact with each other has been sidestepped.

In the real world, however, the static and dynamic shape cues are almost always confounded. The objects we see moving, e.g. cars, people and airplanes have non-random static configurations that may be used to derive good 3-D shape estimates. Since these estimates are available simultaneously with those from motion cues, we are faced with the question of whether they play a role in determining the eventual 3D percepts.

Let us consider two extreme scenarios. In scenario 1, the visual system uses only motion trajectories of feature points to estimate structure ('features' are defined as points of high curvature or other punctuate discontinuities). This has been the typical approach for sequence analysis. The reason for the popularity of this approach is its ability to yield unique structure interpretations based on motion information from only a few frames and features. For instance, Ullman [36] has shown that under orthographic projection, three views of four non-coplanar points are sufficient to guarantee a unique 3-D solution. Longuet-Higgins and Prazdny [18] proved that the instantaneous velocity field and its first and second spatial derivatives at a point admit at most three different 3-D interpretations. Tsai and Huang [35] showed that two perspective views of seven points are also usually sufficient to guarantee uniqueness.

While scenario 1 is mathematically elegant and powerful, we should consider whether the human visual system does in fact adopt such a strategy, or perhaps it might incorporate other, non motion-based, cues as well in its analysis of dynamic sequences. This leads us to scenario 2. Here, we treat each frame of a motion sequence as an entity to be analyzed on its own, in terms of the figural cues it contains. These cues provide 3D estimates on a frame-by-frame basis, rather than requiring the use of feature motion trajectories, as prescribed by scenario 1. Several figural cues, such as shading, or texture gradients can provide 3D structure information [7,12, 33,47]. Even in the absence of such gradients, global contour based cues provide powerful constraints for 3-D shape recovery [8,13,17,19,43,44] as demonstrated by computational schemes for the recovery of 3D structures from simple 2D line drawings [20,29,30].

Is human analysis of dynamic sequences more akin to scenario 1 (predominant use of motion information), or scenario 2 (predominant use of figural information, when such information is available)? Addressing this question presents a challenge in that for most dynamic sequences, both scenarios tend to produce identical results. For instance, a moving wire-frame cube would be seen as a cube irrespective of whether one uses structure from motion algorithms on the vertex trajectories, or applies shape from contour algorithms on individual frames. In order to overcome this constraint, we need dynamic sequences where the motion based and figural content-based strategies yield different results.

Here we use dynamic sequences showing rigid wire-frames in motion, where the wire-frame objects are specially constructed so that their different views suggest different 3D shapes. Conventional structure from motion algorithms would easily recover the true rigid 3D structure of these objects. However, the use of figural cues on a frame by frame basis would suggest that the underlying 3D shape was changing over time. Thus scenario 1 would predict the percept of a rigid object, while under scenario 2, a non-rigid percept would result. Figure 1 illustrates our proposal. The object depicted across all of the frames is a rigid wire-frame

but its different views appear to derive from different underlying 3D geometries. Thus, if scenario 2 is correct, then even though there is a unique rigid interpretation for the whole sequence, observers would opt for a strange non-rigid one. If, however, motion information takes precedence over figural cues, then the veridical rigid structure will be perceived. In the next section, we develop an experimental paradigm to test this hypothesis.

## Methods

### Experimental tests of the contribution of motion and figural cues in the analysis of dynamic sequences

We describe our experimental paradigm by means of a specific example. Consider a 2-D image showing a cube. Up to a depth reversal, there is a unique set of depth values for all of the vertices that would be precisely consistent with a 3-D cube shape. Let us label this set of depth values as defining the ‘base’ structure. Now, we can add increasing amounts of random depth noise to the vertices. Since the noise is constrained to be only along the depth axis, the 2-D projections of the resulting objects are always the same (from the original viewpoint). However, the 3-D structures are now no longer consistent with an observer’s expectations based on the original projection, since the expectation is still that of a cube. Consequently, we would expect that the noise-distorted variants will, upon rotation, look non-rigid (since the 3D structures suggested by the initial view is different from that suggested by a slightly rotated view). Furthermore, the perceived non-rigidity will increase with the amount of depth noise added. However, the original, no-noise base structure will look rigid, since there is no inconsistency between the 3D structures suggested by the different frames. If the objects are presented simply as vertices, without the connecting segments, then figural cues are highly impoverished, thus largely doing away with the inconsistencies between motion and figural cues. The prediction would be that such vertex-only motion sequences will be perceived as rigid, irrespective of the amount of noise added.

These hypotheses are illustrated in figure 2. Since the sequence in figure 2 (a) has a seemingly random collection of dots, the static shape recovery processes do not yield a strong shape estimate, and the structure from motion mechanism, working under the constraint of the rigidity assumption, would be expected to produce a rigid interpretation for the sequence. However, if we link the dots in the previous motion sequence in a manner that permits some estimates to be made about their 3D configuration statically (see figure 2 (b)), then since (in this particular example) the figurally estimated 3D configuration of the dots appears to be varying across the frames, a percept of non-rigidity would be expected when the object begins moving. The inclusion or exclusion of the inter-dot lines, in effect, is expected to act as a perceptual switch that renders active or inactive the static shape recovery process with measurable perceptual consequences.

In summary, the experimental paradigm is designed to study whether the magnitude of the perceived non-rigidity changes as increasing amounts of depth noise are added to a perceptually correct 3D structure. In previous studies, [31,32], we have shown that depending upon the figural cues used, the motion percept can be strongly biased to be one of extreme rigidity or non-rigidity. We expect that if figural information plays a role in the analysis of dynamic sequences, then the original 3D structure will look the most rigid

relative to the noisy versions and there will be a systematic increase in perceived non-rigidity with increasing noise levels (as schematized in figure 3a). If, on the other hand figural information does not influence sequence analysis, the original 3D structure will not be at a minimum of perceived non-rigidity and no systematic trends relating non-rigidity and added noise-level will be observed (figure 3b).

We can make an additional prediction in the domain of neuropsychology. If the percepts of non-rigidity have their genesis in figural analysis, as we hypothesize, then we would expect that these percepts will be obtained even with patients who have compromised motion perception due to neurological damage. Such observers will experience increased percepts of non-rigidity with increasing noise in sequences that have figural information, but not in sequences devoid of figural information. If, on the other hand, the percepts are a consequence of motion analysis, then the results with such patients will be different from those of observers with normal motion perception.

## Procedure

Our experimental display comprised two simultaneously presented variants of a base 3-D object (figure 4). Subjects viewed the monitor from a distance of 80 cm. Each of the stimuli subtended, on average, 3 degrees of visual angle at this distance. Subjects were asked to maintain fixation at a mark placed between the two objects in the display. Their task was to indicate with a mouse click on either one of two screen buttons, which of the two objects looked more non-rigid. Each of the objects underwent rotation around a horizontal axis passing through its centroid. The object rotated by  $\pm 10$  degrees about the starting position with an angular speed of 15 degrees/second. To ensure that subjects understood the meanings of the terms 'rigid' and 'non-rigid', we preceded the experiment with a short session wherein they were shown a real rigid object like a plastic cube and a non-rigid one like a squeezable ball and two flaps on a hinge (the opening and closing of a book).

Two different polyhedra were used as base objects, and were presented in two modes: as vertices only and as vertices connected with line segments. For each base object (figure 4b), five additional variants were created by adding positional jitter of varying magnitudes along the depth axis. This 'depth noise' could range upto 20%, 40%, 60%, 80% and 100% of the full depth extent of the base object. These five levels corresponded to the five variants for each of the base objects. For each collection of six configurations (base object + five depth distorted variants) all possible pairings were shown to the subjects. The presentation was self-paced and no feedback was provided.

Data from the forced alternative design were compiled to obtain rank-orders of the six configurations along the dimension of perceived non-rigidity. This was done by incrementing the 'non-rigidity score' associated with a given configuration by unity every time it was rated to be more non-rigid than another. Thus, a configuration that was always rated more non-rigid in every pair that it was displayed in (there would be five such pairs in all) would achieve a score of 5 while a configuration that always was seen as being more rigid than any other would end up with a score of 0. After all 15 pairs had been presented, the relative scores would allow us to place the six configurations on an ordinal scale of perceived non-rigidity.

## Subjects

Two healthy adult subjects and two neurological patients who had suffered a single ischemic stroke participated in our experiments. The healthy subjects were drawn from the Boston University subject pool and age matched to the patients. They had no visual acuity impairments. The patients' neurological damage affected their motion perception abilities. A brief description of the two patients is provided below. All subjects gave informed consent according to Boston University Institutional Review Board.

**Patient AMG**—At the time of testing, patient AMG, was a 53-year-old right handed woman who underwent an acute left hemisphere cerebro-vascular accident. MRI demonstrated involvement of the left posterior parietal lobe extending into the posterior-medial part of the temporal lobe into the central and lateral portion of the occipital lobe but sparing the primary visual cortex. The lesion also extended into the white matter by the trigone of the lateral ventricle and around the margin of the body of the lateral ventricle (Figure 5a). On neuropsychological evaluation on the Wechsler Memory Scale-Revised (WAIS-R), her verbal IQ was in the average range (VIQ = 96), the Performance I.Q. was in the low average range (PIQ=80), and the full scale I.Q. in the low average range (FIQ=88) compared to her peers. Snellen acuity was 20/20 in both eyes. Visual fields assessed shortly after the stroke by Goldmann perimetry revealed a right inferior quadrantanopsia for very small targets. This improved by the time of the testing reported here (12 months after the infarct). As described in [39,40] AMG was severely impaired on several low-level motion tasks (speed and direction discriminations). Her performance was normal on tasks involving spatial integration, but she could not integrate motion information across scales of resolution (she failed on the plaid test). AMG' performance on visual motion perception was reported in detail previously [38,39,40], here we only summarized the results to demonstrate that she was not “motion blind”, but her motion deficits were selective.

**Patient JT**—At the time of testing JT was a 52-yr-old right-handed woman who underwent a large infarction in the region of the left middle and posterior cerebral arteries extending from the left temporal pole posteriorly to the occipital lobe and superiorly to the most posterior aspect of the parietal lobe. There is also involvement of the inferior left temporal lobe (Figure 5b). Visual fields assessed by Goldmann perimetry showed a complete right homonymous hemianopsia. Her score on the Performance IQ test of the Wechsler Adult Intelligence Scale, Revised (WAIS-R) was below average for her age and education-level (PIQ =79). JT performed within normal range on the discrimination of spatial relations. Language comprehension was excellent and when asked to describe a scene visually presented she first needed to scan it for several seconds and then she described the scene reporting single items but not the global scene being represented. When asked to match the correct action corresponding to a verb auditorily presented she could easily indicate the correct corresponding picture. Her naming ability on visual presentation was severely impaired, she was presented with 15 black and white outline drawings (taken from the Birmingham Object Recognition Battery, BORB [27]) and she scored 3/15 ( $z = -4.4$ , compared to age-matched control group). Her short term memory was mildly impaired (digit span = 3). She had alexia without agraphia. On repeated testing, she was unable to read words longer than 2 letters, while she could write words of different length under dictation.

She could copy outline drawings, 2-D shapes and 3-D shapes and overlapping geometrical figures. She also copied correctly drawings which required switching attention between objects (a house surrounded by three trees). However, Patient JT exhibited a total inability to perceive 2-D form from relative motion and 3-D form from motion.

## Results

Relative scores of perceived non-rigidity were obtained as described in the ‘Procedure’ section. The results from the healthy observers and the patients are shown in figure 6. They are plotted as perceived non-rigidity scores as a function of amount of depth noise added. For all of the plotted points, we determined the coefficients of correlation (between amount of depth-distortion and perceived non-rigidity score) and their statistical significance. Healthy observers exhibited an increased percept of non-rigidity with higher levels of added noise when the sequences included figural cues (i.e. objects were presented with line segments). The correlation between these two variables was high and statistically significant. However, no statistically significant correlation was evident when the objects were presented only as collections of vertices. Patients AMG and JT also failed on the 3D-Structure from motion test. Similar to other motion impaired patients, who in separate tests had been found to be highly compromised on their structure from motion abilities [38,39] also exhibited the same pattern of results – a strong positive correlation between added noise and perceived non-rigidity in the presence of figural cues, but not without them.

## Discussion

The experimental results suggest that, when available, figural information plays a significant role in the analysis of dynamic sequences. This is in contrast to conventional theories of structure from motion, wherein motion cues are used exclusively for structure recovery in sequences [35,36]. Recent evidence from other research groups also highlights the role of figural information in the analysis of motion sequences [2,21]. In the light of these results, we now discuss a revised theoretical formulation of the processes responsible for the perception of 3D structure in dynamic sequences.

Conceptually, we can assume that two processes operate concurrently - one concerned with 3-D shape recovery from dynamic cues and the other from figural ones. The 3-D estimates they produce as outputs need to be combined in order to produce the eventual percept. Exactly how this combination is accomplished is currently unknown. The results of the experiments presented here suggest that when strong figural cues are available, they can over-ride the motion derived shape estimates. However, our experiments leave open the issue of the combination rules when the figural cues are weakened, for instance by having unfamiliar shapes which are not associated with any specific 3D structures. Confining our attention to the case where the figural cues are strongly evocative of specific 3D geometries, we can outline a simple processing strategy for the analysis of dynamic sequences. The strategy involves verifying whether the figurally derived shape at different time instants is the same. A rigid object is expected to maintain the same shape at all times. Changes (over time) in the figurally derived 3-D shapes would suggest non-rigidity even though the SFM (Structure- From-Motion) process might recover a rigid structure. Although adequate to

account for our present experimental results, this is probably an oversimplification. Both the static and dynamic shape recovery processes are likely to be important for the analysis of arbitrary motion sequences. Briefly, the simple model these results suggest requires determining the 3-D shape of the object seen in the motion sequence from static image attributes at frequent intervals. Depending on the constancy (or lack thereof) of the recovered shape at different times, the model predicts whether the object would be perceived as rigid by human observers.

## Conclusion

We have considered the perceptual analysis of dynamic sequences that provide not merely motion information of a few feature points, but also figural information in each frame. We hypothesized that the figural cues would contribute to the eventual percept experienced by an observer. The results from experiments with normal and neurologically impaired subjects support this hypothesis. They suggest that figural cues can overwhelm motion cues, even to the extent of over-riding the rigidity bias, resulting in objects that are physically rigid being perceived as non-rigid by observers. Furthermore, even participants whose neurological damage impaired their motion perception skills, reported percepts of non-rigidity, suggesting that the percepts were likely being engendered by figural analysis. Based on these results, we have outlined an admittedly oversimplified theoretical account of shape recovery in motion sequences.

Several important questions remain open. Primary among these is a finer grained analysis of interactions between motion-based and figural-based 3D shape recovery processes. Investigating this issue will require a way to parametrically change the strengths of these two types of cues, and assessing the effects of these changes on the eventual percept. This study has contributed evidence showing that such interactions do exist and that the analysis of motion sequences involves a significant contribution from figural cues.

## Acknowledgments

The authors wish to thank all observers who participated in the experiments reported here. This research and the preparation of this manuscript was supported by the National Institutes of Health-RO1 NS064100 grant to LMV.

## References

1. Barrow HG, Tenenbaum JM. Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*. 1981; 17:75–116.
2. Beintema JA, Lappe M. Perception of biological motion without local image motion. *Proc Natl Acad Sci U S A*. 2002; 99(8):5661–3. [PubMed: 11960019]
3. Bobick, A. In: Badler, N.; Tsotsos, J., editors. *A Hybrid Approach to Structure-From-Motion; Motion: Representation and Perception*, Proceedings of Association of Computing Machinery Siggraph Workshop on Motion; Toronto, Canada. 1983. p. 91-109.
4. Braunstein ML. Depth Perception in rotating dot patterns. *J Exp Psych*. 1962; 64:415–420.
5. Braunstein, ML. *Depth Perception through Motion*. Academic Press, Inc; New York: 1976.
6. Braunstein ML, Andersen GJ. A counterexample to the rigidity assumption in the perception of structure from motion. *Perception*. 1984; 13:213–217. [PubMed: 6504681]
7. Clocksin WF. Perception of surface slant and edge labels from optical flow: A computational approach. *Perception*. 1980; 9(1):253–269. [PubMed: 7454508]



8. Clowes MB. On Seeing Things. *Artificial Intelligence*. 1971; 2(1):79–116.
9. Gibson JJ, Gibson EJ. Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology*. 1957; 54:129–138. [PubMed: 13475637]
10. Hildreth, EC. *The measurement of visual motion*. Cambridge, MA: MIT Press; 1984.
11. Hoffman DD, Flinchbaugh BE. The interpretation of biological motion. *Biological Cybernetics*. 1982; 42:195–204. [PubMed: 7059621]
12. Horn, BKP. Obtaining Shape from Shading Information. In: Winston, PH., editor. *The Psychology of Computer Vision*. Vol. chapter 4. McGraw-Hill; New York: 1975. p. 115-155.
13. Huffman, DA. Impossible objects as non-sense sentences. In: Meltzer; Michie, editors. *Machine Intelligence 6*. Edinburgh Univ. Press; 1971. p. 295-323.
14. Longuet-Higgins HC, Prazdny K. The interpretation of a moving retinal image. *Proceedings of the Royal Society*. 1980; B208:358–397.
15. Jansson G, Johansson G. Visual perception of bending motion. *Perception*. 1973; 2:321–326. [PubMed: 4794128]
16. Johansson, G. *Visual motion perception*. WH Freeman; San Francisco: 1975.
17. Kanade T. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*. 1981; 17(1–3):409–460.
18. Longuet-Higgins HC. A computer algorithm for reconstructing a scene from two projections. *Nature*. 1981; 293:133–135.
19. Mackworth AK. Interpreting pictures of polyhedral scenes. *Artificial Intelligence*. 1973; 4:121–137.
20. Marill T. Emulating the human interpretation of line-drawings as three-dimensional objects. *Intl J of Comp Vis*. 1991; 6:147–161.
21. Michels L, Lappe M, Vaina LM. Visual areas involved in the perception of human movement from dynamic form analysis. *Neuroreport*. 2005; 16(10):1037–1041. [PubMed: 15973144]
22. Mitiche, A. Computation of optical flow and rigid motion. *Proc. Workshop Comput. Vision: Representation and Contr*; Annapolis, MD. 1984. p. 63-71.
23. Mitiche A. On kineopsis and computation of structure and motion. *IEEE Trans on Pattern Analysis and Machine Intelligence*. 1986; 8(1):109–112.
24. Mitiche, A.; Seida, S.; Aggarwal, JK. Determining position and displacement in space from images. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*; San Francisco, CA. 1985. p. 504-509.
25. Prazdny K. Egomotion and relative depthmap from optical flow. *Biological Cybernetics*. 1980; 36:87–102. [PubMed: 7353067]
26. Prazdny K. On the information in optical flows. *Computer Vision, Graphics, and Image Processing*. 1983; 23:239–259.
27. Riddoch, MJ.; Humphreys, GW. *BORB: the Birmingham Object Recognition Battery*. Hove; Erlbaum UK: 1993.
28. Sinha, P. Masters Dissertation. MIT Department of Electrical Engineering and Computer Science; 1992. *The Perception of Shading and Reflectance*.
29. Sinha, P. Doctoral Dissertation. MIT Department of Electrical Engineering and Computer Science; 1995. *Perceiving and recognizing three-dimensional forms*.
30. Sinha, P.; Adelson, EH. Recovering reflectance in a world of painted polyhedra. *Proc. Fourth Intl. Conf. Comp. Vision*; Berlin. 1993. p. 156-163.
31. Sinha P, Poggio T. The role of learning in 3-D form perception. *Nature*. 1996; 384(6608):460–463. [PubMed: 8945472]
32. Sinha, P.; Vaina, LM. Perception of 3D structure from static form and/or motion cues in a patient with visual motion deficits. *Proceedings of the Annual Meeting of the Optical Society of America*; Albuquerque. New Mexico: 1992. p. 1331-1332.
33. Stevens KA. The visual interpretation of surface contours. *Artificial Intelligence*. 1981; 17:47–73.
34. Talairach, J.; Tournoux, P. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional*. Thieme; Germany: 1988.

35. Tsai R, Huang T. Estimating three-dimensional motion parameters of a rigid planar patch. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1981; 29:1147–1152.
36. Ullman, S. *The interpretation of visual motion*. Cambridge, MA: The MIT Press; 1979.
37. Ullman, S. Recent computational studies in the interpretation of structure from motion. In: Rosenfeld, A.; Beck, J., editors. *Human and Machine Vision*. New York: Academic; 1983.
38. Vaina LM, Gryzwacz NM, Saiviroonporn P, LeMay MM, Bienfang DC, Cowey A. Can spatial and temporal motion integration compensate for deficits in local motion mechanisms? *Neuropsychologia*. 2003; 41:1817–36. [PubMed: 14527545]
39. Vaina LM, Cowey A, Jakab M, Kikinis R. Deficits of motion integration and segregation in patients with unilateral extrastriate lesions. *Brain*. 2005; 128:2134–2145. [PubMed: 15975945]
40. Vaina LM, Grzywacz NM, Kikinis R. Segregation of Computations underlying perception of motion discontinuity and coherence. *Neuroreport*. 1994; 5:2289–2294. [PubMed: 7881048]
41. Wallach H, O'Connell DN. The kinetic depth effect. *Journal of Experimental Psychology*. 1953; 45(4):205–217. [PubMed: 13052853]
42. Wallach H, Weisz A, Adams PA. Circles and derived figures in rotation. *American Journal of Psychology*. 1956; 69:48–59. [PubMed: 13302498]
43. Waltz, DL. Generating semantic descriptions from drawings of scenes with shadows. In: Winston, PH., editor. *The Psychology of Computer Vision*. New York: McGraw-Hill Book Co; 1972.
44. Waltz, D. Understanding line drawings of scenes with shadows. In: Winston, PH., editor. *The Psychology of Computer Vision*. New York: McGraw-Hill Book Co; 1975. p. 19-91.
45. Waxman AM, Ullman S. Surface Structure And Three-Dimensional Motion From Image Flow Kinematics. *IJRR*. 1985; 4:72–94.
46. White BW, Mueser GE. Accuracy in reconstructing the arrangement of elements generating kinetic depth displays. *J Exp Psychol*. 1960; 60:1–11. [PubMed: 13844335]
47. Witkin AP. Recovering surface shape and orientation from texture. *Artificial Intelligence*. 1981; 17:17–45.

## Biographies



### Lucia Maria Vaina

Harvard Medical School, Department of Neurology ; Boston University; Departments of Biomedical Engineering, Neurology and Neuroscience; Massachusetts Institute of technology, Center for Biological and Computational Learning, Department of Brain and Cognitive Science

Lucia Maria Vaina is a professor of biomedical engineering and she is a clinical and computational research faculty at Harvard University and Massachusetts Institute of Technology. She received her PhD in mathematics from the Sorbonne University, Paris, France, and her MDPHD at Toulouse.

Using a combination of experimental, computational and magnetoencephalography approaches she studies visual –cognitive functions in the normal and damaged human brain

in order to provide an explanatory model of the spatiotemporal orchestration of high level sensory brain functions and to understand how they fail as a result of focal brain lesions due to stroke.

Professor Lucia Maria Vaina is a fellow of the American Institute for Medicine and Biomedical Engineering. In 2005, she received a Doctorate Honoris Causa from the University V Goldis for her contributions to medical science. She received the Alfred Sloan Foundation award, and the Ford Foundation Award for postdoctoral studies that she carried out at University of California, at Berkeley, Stanford University and Massachusetts Institute of Technology.

Professor Vaina serves as a charter member on the Perception and Cognition study section of the National Institute of Health, and she also served as an ad hoc member on many NIH and NSF study sections, including the Panel reviewing the ‘Human Connectome Project’. She serves on the editorial board of several journals, among which ‘Synthese’, ‘Journal of Neurophysiological and Neurological Disorders’, ‘Medical Science Monitor’, and the ‘American Journal of Case reports’



### **Pawan Sinha**

Department of Brain and Cognitive Sciences

Massachusetts Institute of Technology

Pawan Sinha is a professor of computational and visual neuroscience in the Department of Brain and Cognitive Sciences at MIT. He received his undergraduate degree in computer science from the Indian Institute of Technology, New Delhi and his Masters and doctoral degrees from the Department of Computer Science at MIT.

Using a combination of experimental and computational modeling techniques, research in Prof. Sinha’s laboratory focuses on understanding how the human brain learns to recognize objects through visual experience and how objects are encoded in memory. Prof. Sinha’s experimental work on these issues involves studying healthy individuals and also those with neurological disorders such as autism. The goal is not only to derive clues regarding the nature and development of high level visual skills, but also to create better therapeutic routines to help children overcome visual impairments. Prof. Sinha founded Project Prakash in 2005 with the twin objectives of providing treatment to children with disabilities and also understanding mechanisms of learning and plasticity in the brain. Project Prakash, which is a collaboration between MIT researchers and ophthalmologists in India, has provided insights into some fundamental questions about brain function while also transforming the lives of many blind children by bringing them the gift of sight.

Prof. Sinha is a recipient of the PECASE – the highest US Government award for young scientists, the Alfred P. Sloan Foundation Fellowship in Neuroscience, the John Merck Scholars Award for research on developmental disorders, the Jephtha and Emily Wade Award for creative research, the Troland Award from the National Academies, and the Distinguished Alumnus Award from IIT Delhi

Prof. Sinha has served on the program committees for prominent scientific conferences on object and face recognition and is currently a member of the editorial board of ACM’s Journal of Applied Perception. He is a founder of Imagen Inc, a company that applies insights regarding human image processing to challenging real world machine vision problems. Imagen was the winner of the MIT \$50K Entrepreneurship competition. Prof. Sinha was named a Global Indus Technovator, and was also inducted into the Guinness Book of World Records for creating the world’s smallest

Author Manuscript

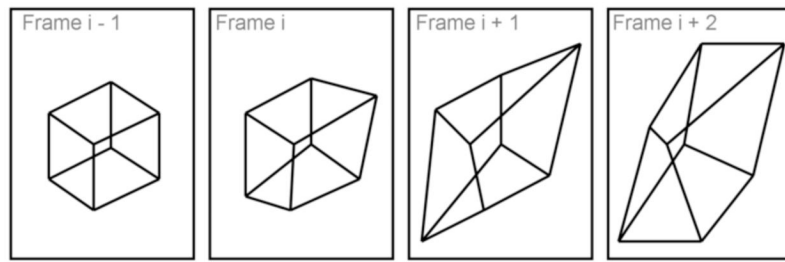
Author Manuscript

Author Manuscript

Author Manuscript

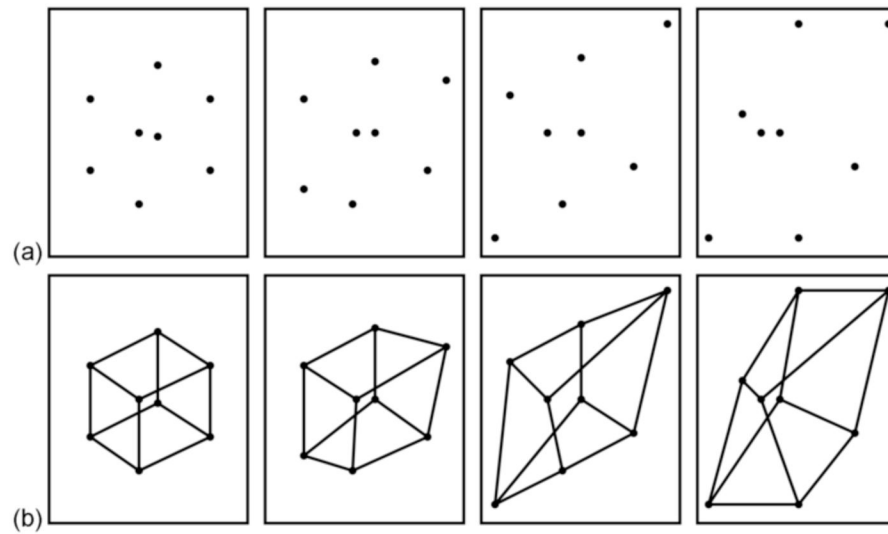
### Highlights

1. Addressing how dynamic and figural cues interact in visual sequences
2. Novel paradigm to dissociate contributions of figural and dynamic cues
3. Assessment of healthy subjects and those with brain lesions
4. Empirical results point towards a computational model of visual sequence analysis.



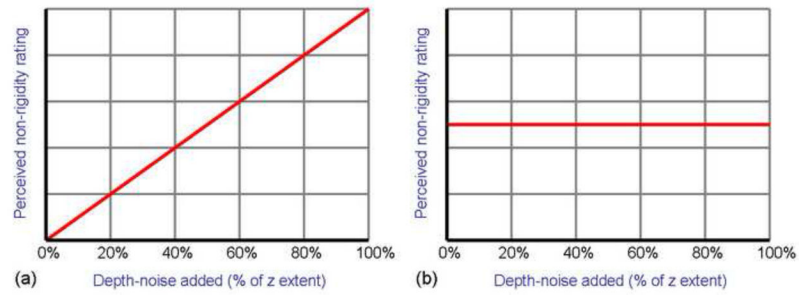
**Figure 1.**

A few frames from a motion sequence that shows a specially constructed rigid 3D object whose different views suggest different 3D structures. Analysis of the motion trajectories of the vertices would yield the veridical rigid 3D structure. However, an analysis based on figural cues within each frame would suggest non-rigidity.



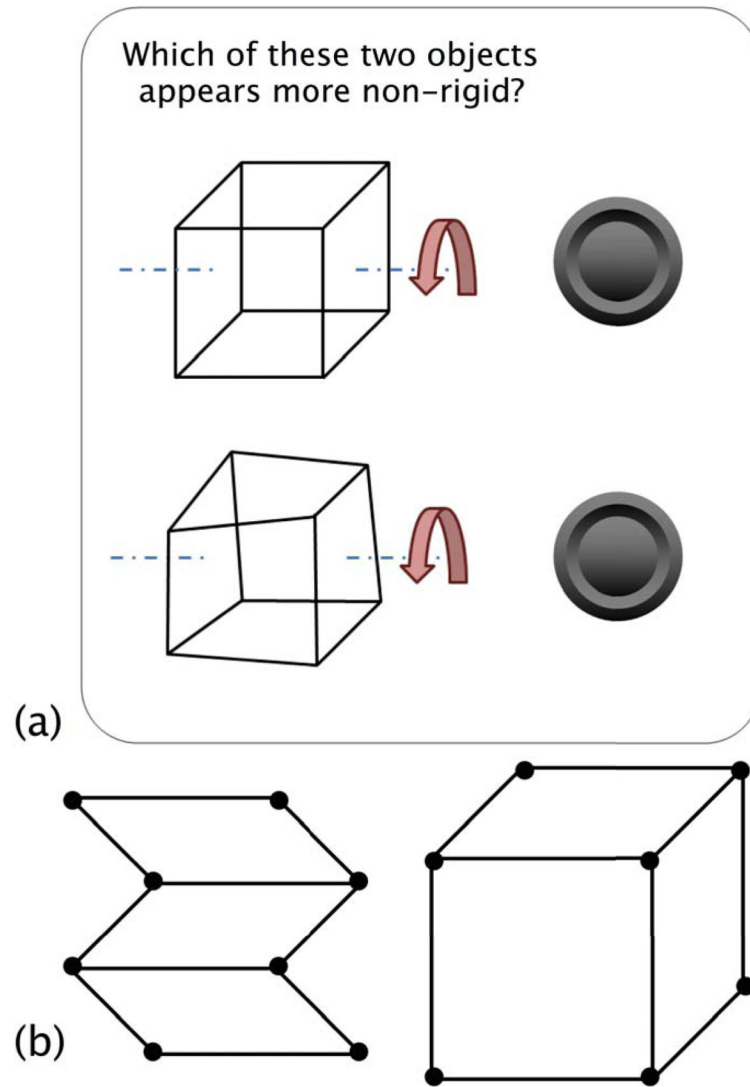
**Figure 2.**

If figural cues participate in 3D shape analysis with dynamic sequences, the percepts obtained with a sequence are expected to be different depending on whether or not static 3D figural cues are available. Under this hypothesis, the top sequence would be seen as a rigid configuration of dots rotating. The bottom sequence, however, would be perceived as a non-rigidly distorting object even though the underlying vertex structure and the motion parameters are identical to those in the top row.



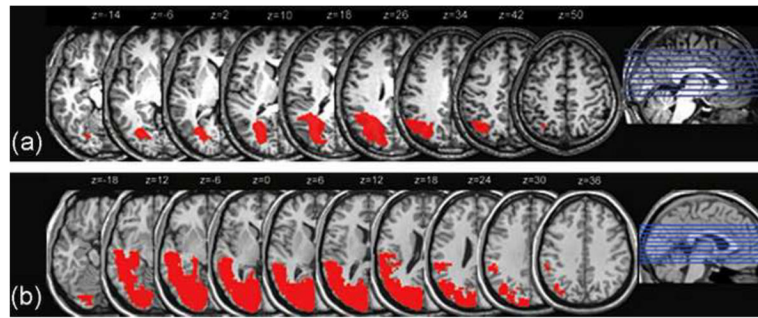
**Figure 3.** Schematic depiction of expected pattern of results under the hypothesis of the use of static figural cues in motion sequence analysis. The ordinate represents ratings of perceived non-rigidity. (a) Expected results when figural cues are present in displays. (b) Expected results in the absence of figural cues (vertices only).





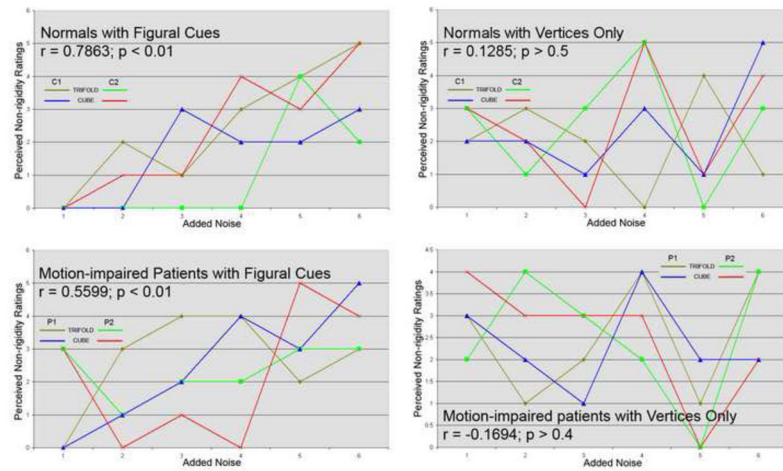
**Figure 4.**

(a) Layout of our experimental display. Subjects were asked to maintain fixation at a mark placed between the two objects in the display. Each of the objects underwent rotation around a horizontal axis passing through its centroid. The object rotated by  $\pm 10$  degrees about the starting position with an angular speed of 15 degrees/second. Observers' task was to indicate with a mouse click on either one of two screen buttons, which of the two objects looked more non-rigid. Such paired comparisons were repeated for all possible pairings of the six noise variants corresponding to each of two objects. (b) The objects we used as stimuli in our experiments. The object on the left is referred to as the 'trifold' and that on the right as 'cube'.



**Figure 5.**

(a,b) shows axial MR slices registered into the Talarach space [34] by using MRIcro (<http://www.mricro.com>) illustrating in red the left hemisphere infarctions in both patients: AMG in (a) and JT in (b). The sagittal view (the right most figure) illustrates the level of the slices shown from the lowest to the highest, and “z” refers to the axial slice number in the Talarach coordinates. MR images were obtained with a Magnetom 1.5T unit (6mm thick slices, with 1.2mm gap). (a) Patient’s AMG lesion is in the left hemisphere and it extends laterally by the margins of the occipital horns of the lateral ventricle into the lateral portion of the occipital lobe; (b) Patient JT has a large infarction of the left hemisphere, extending from the occipital pole to the temporal pole with a significant involvement of the occipital lobe. The medial temporal lobe, however, is preserved to a great extent, although the parahippocampal gyrus is likely involved in the lesion.



**Figure 6.**

Plots of perceived non-rigidity as a function of added noise for two healthy observers (C1 and C2) and two motion-perception impaired participants (P1: AMG and P2: JT). The four curves in each plot correspond to data from two subjects (either normal or motion impaired, indicated in each plot) on each of two objects (in each panel, the brown and green plots depict data from one subject and the red and blue plots depict data from the second subject). Also shown are the correlation-coefficients and their levels of significance.