A Computational and Evolutionary Perspective on the Role of Representation in Vision

Michael J. Tarr and Michael J. Black

Research Report YALEU/DCS/RR-899
October 1991

A Computational and Evolutionary Perspective on the Role of Representation in Vision*

Michael J. Tarr† and **Michael J. Black**‡§

†Department of Psychology, Yale University
P.O. Box 11A Yale Station, New Haven, CT 06520

Phone: (203) 432–4637, Fax: (203) 432–7172
Email: tarr-michael@cs.yale.edu

‡Department of Computer Science, Yale University
P.O. Box 2158 Yale Station, New Haven, CT 06520–2158

Phone: (203) 432–1223, Fax: (203) 432–0593
Email: black-michael@cs.yale.edu

Abstract

Recently, the assumed goal of computer vision, to reconstruct a representation of the scene, has been criticized as unproductive and impractical. Critics have suggested that the reconstructive approach should be supplanted by a new purposive approach that emphasizes functionality and task driven perception at the cost of general vision. In response to these arguments, we claim that the recovery paradigm central to the reconstructive approach is viable, and moreover, provides a promising framework for understanding and modeling general purpose vision in humans and machines. An examination of the goals of vision from an evolutionary perspective and a case study involving the recovery of optic flow support this hypothesis. In particular, while we acknowledge that there are instances where the purposive approach may be appropriate, these are insufficient for implementing the wide range of visual tasks exhibited by humans (the kind of flexible vision system presumed to be an end-goal of artificial intelligence). Furthermore, there are instances, such as recent work on the estimation of optic flow, where the recovery paradigm may yield useful and robust results. Thus, contrary to certain claims, the purposive approach does not obviate the need for recovery and the reconstruction of flexible representations of the world.

*We thank Greg Hager and David Kriegman for their helpful comments and advice.

§This author was supported by a grant from the National Aeronautics and Space Administration (NGT–50749)

1 Introduction

Young disciplines often experience moments of doubt, "Are we doing the right thing?," or, "Is such and such an approach viable?" (Banaji, 1991). No where is this better exemplified than in the study of computer vision (Jain & Binford, 1991). While progress has been made, the goal of general vision, on the order of human visual perception, remains elusive. Recently, this has led some to suggest that the entire endeavor is flawed, that we should discard the dominant paradigm, and that it should be replaced with a new, more practical alternative.[1] While this position may not qualify as a "paradigm shift" (Kuhn, 1970), it certainly advocates a substantial change in direction. To justify this radical deviation, proponents of the new, so-called *purposive approach* muster three lines of support: first, that machines fall far short of the visual capabilities of humans; second, that current computer vision systems can not actually do very much that is useful in the way of visual perception; and, third, that the purposive approach is consistent with the notion that biological organisms have evolved brain machinery composed of independent processes, each devoted to solving a particular visual task (Aloimonos, 1990).

Contrary to these arguments, we take an entirely conservative posture, suggesting that the presently dominant reconstructive approach is viable, and, moreover, that there are well-grounded computational and evolutionary reasons for its current, as well as possible future, successes. In support of these claims, two kinds of evidence will be presented: first, a general examination of the goals of vision in both artificial and biological systems; and, second, a case study of the recovery of structure from motion and optic flow. These arguments lead us to conclude that the reconstructive approach provides a framework for understanding

[1] These suggestions are somewhat reminiscent of proposals that the entire field of artificial intelligence should be scrapped; for instance, see Penrose, 1989, or Searle, 1981.

both human and machine vision, and that, in particular, there are already instances where successes have led to useful and robust vision systems. Moreover, this is not at the exclusion of the purposive approach, but rather suggests a common ground that we believe to be more fertile than either approach alone.

2 Understanding Vision

In understanding vision, one must begin with the fact that many problems in visual perception are considered to be "ill-posed" (Marroquin, Mitter, & Poggio, 1987) and that in order to find a solution inference and constraint must be introduced. While these constraints may be phrased in general terms, for instance the "rigidity" constraint of Ullman's (1979) structure from motion algorithm, it is also possible to narrow the domain, positing specialized constraints dependent upon the visual task at hand. This constitutes the crux of the purposive approach. Moreover, there are instances in the natural world where the notion of narrow constraint is obviously applied. In a now classic (but once ridiculed) paper, Lettvin, Maturana, McCulloch, and Pitts (1959) demonstrated that a frog's retina contains special purpose hardware sensitive to small moving black spots; commonly understood to be "bug detectors". Clearly, this adaptation provides great advantages to the amphibian so equipped: frogs mostly eat bugs and their effective detection, even at the cost of an occasional false alarm, is presumably a key to frog survival. However, what is good for frogs is not necessarily good for humans or "intelligent" machines. In particular, it may be impossible to identify the specific tasks appropriate to this level of constraint in such complex systems. Marr (1982) raises precisely this point in reference to the level at which knowledge or hypothesis is brought to bear on visual perception. For while such constraints may not be "general but particular and true only of the scene in question," this suggests that "any very general vision system must command a very large number of such hypotheses and be able to find and deploy just the one or two demanded by the particular situation". Moreover, "this prospect casts a whole (new) complexion on the vision problem" (Marr, 1982, p. 271). We contend that couched in terms of the purposive approach, the number of functionally independent human visual behaviors, as well as their consequential constraints, is too large a number to represent. Indeed, we doubt whether human visual behavior, or for that matter the operation of any general purpose vision system, can be understood in such a narrow context.

Even when one considers higher primates such as vervet monkeys, there is little evidence that their mental representations are generally as abstract and as flexible as our own (Cheney & Seyfarth, 1990). While it is true that "many animals are specialists, performing skills with much greater sophistication in some contexts than in others" (Cheney & Seyfarth, 1990, p. 310), this is not the hallmark of human cognition. For example, many species of birds acquire knowledge of bird song through domain specific "tunable blueprints," an innate special-purpose acquisition mechanism, while human children seem to acquire domain specific knowledge through the operation of general acquisition mechanisms that are rooted in flexible representational structures (Carey, 1985). Without such representations, knowledge will remain compartmentalized and inaccessible, leaving the mental system without the capacity to extend knowledge from one context to another (Cheney & Seyfarth, 1990). In particular, it is this ability that distinguishes human information-processing from that of other species.

3 Religious Reconstructionism[2] and Fanatical Purposivism

It is also this flexibility that distinguishes the reconstructive approach from the purposive approach. Reconstruction, or the recovery paradigm, focuses on deriving a functional description of the visible world including its geometric properties and the physical properties of the visible surfaces. The goal then is to build a symbolic description of the scene. Once derived, symbolic descriptions may be used in a variety of "cognitive" operations, such as visual reasoning, planning, or propositional thought. Stated succinctly, "the goal of a perception system, whether biological or machine, is to create a model of the real world and to use this model for interacting with the real world" (Jain & Binford, 1991, p. 116). What this paradigm means, from the artificial intelligence perspective, is that vision can be effectively ignored; that vision is a self contained problem which will produce symbolic input to AI programs.

In contrast, the goal of the purposive approach is to build systems that will accomplish particular domain specific tasks, the output of which is successful task completion. The study of vision in general is reduced to the study of the "tasks that organisms possessing vision can accomplish" (Aloimonos, 1990, p. 349); independent of such tasks the study of the general problem of vision is not even thought to be possible. For instance, the standard goal of model-based object recognition is supplanted by a framework in which objects are viewed in terms of the their roles, functions, or purposes. It is these properties, not the object's geometry, that serve as the basis for its visual recognition (Aloimonos & Rosenfeld, 1991). Citing a specific example, "a chair is an object on which a person can sit.... To recognize a

[2]We thank Jitendra Malik for suggesting this wonderful phrase.

chair, we should check for the presence of the functional primitive (the surface patch) just defined" (Aloimonos & Rosenfeld, 1991, p. 124; this example is reminiscent of Gibson's, 1979, idea of objects "affording" their functions). But these arguments belie the nature of complex visual information-processing and the stated goals of computer vision/artificial intelligence.[3] For there are no a priori reasons for supposing that general purpose vision is impossible: unquestionably the evolution of vision in humans offers an existence proof for the development of precisely the kind of flexible, reconstructive system to which the discipline of artificial intelligence aspires.

Of course, in specifying end goals that entail the reconstruction of the scene one could argue that computer vision is barking up the wrong tree altogether.[4] Aloimonos (1990) does just that, asserting that the goal of computer vision should not be to build systems that mimic human vision or to serve as general purpose perceptual systems, but to provide answers to the question, "What am I going to use this visual ability for?" (p. 348). Yet, this conception is at odds with one of the major tenets of the purposive approach, that machine perception is not up to snuff with human visual capabilities. For while there is no denying that this is the current state of affairs, this comparison leads to a different research agenda than does the notion that we should give up trying to build "intelligent" vision systems and instead concentrate on simpler domain specific problems.[5] The commonly understood goal of the reconstructive approach is to both explain and implement complex visual behaviors; changing the goal does not solve this problem, it simply avoids it. Taken together, these arguments

[3] Any attempt to understand recognition at this level is also plagued by the fact that even seemingly well-defined concepts such as "even number" appear to be neither definitionally or prototypically represented (Armstrong, Gleitman, & Gleitman, 1983). Thus, for more complex concepts, such as "chair" or "fruit," it may be difficult, if not impossible, to operationalize their core functions or purposes.

[4] As well as much of perceptual psychology; although Gibson, 1979, would most likely concur with the revisionists.

[5] Of course, this ignores the less extreme viewpoint that "intelligent" behavior may be understood as an emergent property of simpler processes.

suggest that the purposive approach does not really provide an alternative explanation to the reconstructive approach, but rather simply offers an alternative goal for computer vision - one that can not hope to explain or accomplish many of the commonly held objectives of artificial intelligence, cognitive psychology, or neuroscience.

4 Evolutionary Perspectives

It would be iniquitous to suggest that advocates of the purposive approach completely ignore evolutionary considerations. Indeed, Aloimonos (1990) suggests that the purposive approach is "consistent with evolution" (p. 348) in that individual visual abilities, such as avoiding danger, locating food, and recognizing kin, would seem to have been selected for independently of each other. This notion is reiterated by Cheney and Seyfarth (1990) when they state that "natural selection, it appears, has acted not on general skills but on behavior in more narrowly defined ecological domains" (p. 310). In this instance, we concur and set forth the hypothesis that the characterization of natural selection as a "tinker" (e.g. Dawkins, 1986) provides strong reasons to believe in some version of information-processing modularity (Fodor, 1983) in the evolution of complex systems. However, this conception of independent mechanisms should not be confused with the purposive agenda of decomposing visual problems into continually simpler tasks. Bequeathing modularity upon a particular subsystem in no way entails that it is in any way purposive, but rather that it may be generally characterized as modality specific, innately specified, hard wired, autonomous, and not assembled (Fodor, 1983). Notice that all of these properties are orthogonal to the information-processing goal of the module. Computational objectives need to be specified independently and may take almost any form, including the recovery of scene attributes or the purposeful execution of a specific visual task.

Recent studies on the evolution of complex information-processing systems in humans underscore this point (Cosmides & Tooby, 1987). For instance, Pinker and Bloom (1990) have argued that natural language is the result of traditional Darwinian selective pressures. In particular, they suggest that human language satisfies two important criteria for when a trait should be attributed to natural selection: first, complex design for some function (e.g. the computational objective), in this instance "the communication of propositional structures over a serial channel" (Pinker & Bloom, 1990, p. 712); and second, the absence of alternative explanations for such complexity. Similarly, we surmise that when human vision is judged by the same two criteria it too should be considered the product of selective evolutionary pressures.

First, it seems incontrovertible that the human visual system exhibits complex design. But for what functions? It is here that we believe the traditional goals of the reconstructive approach come into play, setting forth two clear objectives: the reconstruction of the scene and the recognition of objects within the scene.[6] There are numerous lines of admittedly introspective evidence that human vision is adapted for fulfilling precisely these functions: for example, the recovery of object properties for recognition has implications for kin recognition, social interaction, visual communication, predator avoidance, tool making, and food identification; likewise, the reconstruction of a symbolic representation of the visual scene has implications for danger avoidance, navigation, food location, tool use, and visual reasoning.

Second, we are dubious as to whether the purposive approach provides an alternative

[6]Contrary to the view espoused by several prominent theories of object recognition, e.g. Biederman, 1987, recognition does not entail reconstruction, see Tarr, 1989; the converse is also true - Aloimonos, 1990, points out that reconstruction does not entail recognition, and moreover, that the tasks accomplished by each may be considered independently.

explanation for the evolution of such complex behaviors. Essentially, the purposive approach offers a "divide and conquer" explanation in which "the machinery of the brain devoted to vision consists of various independent processes that are devoted to the solution of specific visual tasks" (Aloimonos, 1990, p. 348). However, while these abilities may be based on common principles, there are hypothesized to have evolved at separate times and in within different neural hardware. Therefore, the purposive approach requires not one, but many "miracles" of luck (for genetic change is merely a matter of chance) - repeatedly arriving at common solutions for an almost infinite variety of purposive tasks beneficial to the survival of our ancestors. In contrast, because the reconstructive approach presupposes that each independent recovery mechanism contributes to a common representation that suffices for general purpose vision, the development of complex visual behaviors requires only a single chance occurrence for each adaptive visual principle.[7]

Notice that human visual cognition clearly displays some of the attributes that one would expect to find in a flexible, general purpose vision system. For instance, it is well documented that humans use mental imagery - a sophisticated subsystem for performing visual reasoning via symbol manipulation over spatial representations (Kosslyn, 1980). The use of mental imagery has been implicated in a variety of problem solving domains; not only is it useful for solving the piano mover's problem, but there is evidence that it is used in scientific reasoning, creative discovery, and navigation (Finke, 1989). Anthropologists have also speculated that an "increased ability to think in - and communicate by means of - specific visual images" and "an emerging consideration of design possibilities by way of two- and three-dimensional images" may have helped to spur the rapid development of new tools and weapons in hu-

[7]Although it may be argued the the likelihood of a single "miracle" is only slightly less than that of several such "miracles".

man evolution. Furthermore, while the actual initiation of image-based representations may not be attributable to "the crossing of a neurological threshold", there is no doubt that certain types of neural hardware (and coincident information-processing capabilities) are a prerequisite (White, 1989, pp. 98-99). Thus, without the presence of such flexible visual representations, one of the most distinctive signatures of human behavior, tool use, might have been impossible.

5 Computational Considerations

Where then does the necessary inference and constraint arise if not from the purposive approach? In actual fact, the seeds for solving this problem may be found in Marr's (1982) original formulation of the problem of vision. This perspective has been reiterated by Marr's collaborator Whitman Richards in a chapter entitled "The Approach" where he states that, "The success of the perceptual act is intimately coupled with the observer's ability to build internal representations whose assumptions reflect the proper structure and regularities present in the world.... Fundamental to perception is thus the notion that there is indeed structure in the world" (Richards, 1988, p. 11). The reconstructive approach is no more mired in the ill-posed nature of visual perception than is the purposive approach! Constraints are introduced, but at the level of the physical world.

Indeed, many of the assumptions of the purposive approach appear to be restatements of constraints found in recovery algorithms. For instance, structure from motion algorithms have often introduced the concept of multiple views (Braunstein, Hoffman, Shaprio, Andersen, & Bennett, 1987), a constraint that is often construed in a manner akin to the idea of "active" vision (Aloimonos, Weiss, & Bandyopadhyay, 1987). In particular, using the types of assumptions found in active vision, structure from motion algorithms are more robust. Aloimonos, an advocate of the purposive approach, raises exactly this point, stating that "one can get more constraints on the motion parameters using many frames" (Aloimonos, 1990, p. 351). Importantly, using active vision in this fashion in no way entails the need for use of the purposive approach. In fact, many current instantiations of the recovery paradigm implicitly make use of active vision, and may benefit further by making this explicit. However, this is not the same level at which "purposivists" have sometimes proposed constraints be applied. For positing a multiple or even a many view constraint is entirely consistent with the approach as espoused by even the most ardent neo-Marrian reconstructionists.

6 Case Study: Structure From Motion

One of the criticisms of the recovery paradigm is that it has failed to take into account the real world; that is, current algorithms are not robust in the presence of noise and are hopelessly inefficient. Given the current state of the art, both charges are valid. An often cited example of these failures is the recovery of structure from motion (Aloimonos, 1990).

Elegant mathematical formulations of the problem have often glossed over, or ignored, the serious issues of noise and correspondence. Let us consider for example the problem of recovering a dense depth map from optical flow.[8] Theoretical results indicate that predictions of structure are very sensitive to small errors in the optical flow field (Aloimonos, 1990). To make matters worse, current formulations of optical flow are ill-posed and error prone. Most optical flow algorithms cannot produce flow fields with the necessary accuracy for meaningful structure to be computed.

[8]We are not advocating this as a fundamental goal of computer vision, but rather as an example of recent trends in the recovery paradigm that we find encouraging.

Another criticism of optical flow is that current algorithms do not take into account the real-time demands of an active perceiver, for instance a mobile robot. It has been suggested that because the computation of optical flow is ill-posed and requires regularization, algorithms for computing it are inherently iterative and ill-suited to real-time applications. The purposive paradigm's alternative to using optic flow is to find representations that are easier to compute; for example normal flow, or qualitative descriptions of the flow field.

Do the failures of optical flow mean that structure from motion is pointless? We think not. First, motion provides important structural information about the world, and it should be exploited. Second, recent work using robust and dynamic algorithms address the main criticisms lodged by advocates of the purposive paradigm. Therefore, we believe it is too soon to dismiss optic flow as hopeless. In fact, we are entering an exciting period in which robust approaches are being developed and the issues of incremental processing in a dynamic environment are being taken seriously. Next we consider these criticisms in light of current research.

6.1 Robust Optical Flow

First, we address the criticism that many current approaches to estimating optical flow are not robust. That is, the flow estimates are sensitive to noise, or *outliers*, that do not correspond to the assumptions of the approach. This can include sensor noise, specular reflections, shadows, and transparency; all commonly occurring situations.

One approach for dealing with the problems of outliers involves the use of robust statistical techniques (Hampel et. al., 1986). For example the work of Schunck (1989) uses a robust clustering of constraint lines to determine optical flow. Also recent work by Black (1991) reformulates the standard, gradient-based, least-squares optical flow equation of Horn and

Schunck (1981) using robust statistics. Correlation-based approaches to optical flow can also benefit from robust techniques which reduce the effects of outliers at motion discontinuities (Black and Anandan, 1990). Additionally, hierarchical, coarse-to-fine, approaches (Anandan, 1989) offer additional robustness by reducing the effects of high frequency noise at the coarse levels.

Optical flow algorithms have also been criticized for performing poorly at motion discontinuities. Most regularization schemes for recovering smooth flow fields have the detrimental effect of over-smoothing at motion boundaries. If the recovery of structure is our goal, then these boundaries are likely to be important. There are now many approaches which attempt to solve this problem. The most notable of which are the Markov random field (MRF) approaches (Black & Anandan, 1991; Konrad & Dubois, 1988; Murray & Buxton, 1987). These approaches represent motion discontinuities either explicitly using "line processes" (Geman & Geman, 1984), or implicitly using *weak continuity constraints* (Blake & Zisserman, 1987).

Recently, Black (1991) showed that a robust statistical formulation of the standard smoothness constraint provides an alternative way of viewing motion discontinuities. Singh (1991) also employs a statistical approach to avoid smoothing discontinuities, but it is not robust in a formal sense. Finally, in cases where the flow estimates are poor or uncertain, it is useful to have an estimate of the flow vector's certainty. Consequently, work on flow confidence measures (Anandan, 1984; Simoncelli, Adelson & Heeger, 1991) is an important contribution, for it may allow processes which use optic flow to ignore poor measurements and still produce meaningful results.

6.2 Temporal Persistence

The second main criticism of optical flow algorithms is that their computational expense prevents them from being used under real-world conditions. Many previous approaches have only considered the two-frame estimation problem and those which have considered longer sequences typically have done so in a batch fashion (Bolles, Baker, & Marimont, 1987; Heeger, 87). There have been recent advances on this front; in particular there are now a number of incremental approaches that compute optic flow dynamically and refine the flow estimates incrementally over an image sequence.

For example, Singh (1991) uses a Kalman filter (Gelb, 1974) to estimate optical flow incrementally. There are a number of other analogous approaches for estimating depth from motion (Heel, 1991; Matthies, Szeliski, & Kanade, 1989). While there are problems with the Kalman filter approach, it brings us closer to the objective of dynamic optical flow. An alternative *incremental minimization* approach (Black & Anandan, 1991) uses a robust formulation similar to those already discussed and solves the difficult minimization problem incrementally over a sequence of images. This approach is unique in that it explicitly incorporates a *temporal persistence* constraint in the formulation of the optical flow equation. Temporal persistence provides a powerful additional constraint, *at the level of the physical world*, on the interpretation of visual motion and results in increased robustness.

In a different direction, various researchers have been looking at the real-time computation of optical flow. Notable is recent work on on implementing optical flow equations using analog devices (Koch, Luo, & Mead, 1988). This work incorporates many of the ideas presented for preserving motion discontinuities. The work of Nishihara (1984) also deserves mention. His simple technique for correlation using zero-crossings can be implemented efficiently in hardware.

7 Final Thoughts

We conclude by reiterating that we are not denying a role for the purposive approach in the study of vision, but that we believe it is better suited for understanding and mimicking the overall visual behavior of frogs rather than humans. We acknowledge that there are some aspects to human visual behavior, particularly those associated with "automatic" or "unconscious" processing, that may warrant a purposive analysis. For instance, partial information from optic flow may provide adequate constraints for navigation or wayfinding; likewise, partial information about surfaces may suffice for grasping an object. Indeed, to date much of the actual research done within the purposive framework has focused on precisely these types of problems (Aloimonos, 1990). However, it is our position that it is crucial that the modules computing such information be general enough that this same information may be utilized in the reconstruction of the scene (and indeed this routinely occurs).[9] Therefore, if the purposive approach does have a role in understanding general purpose vision, it seems likely to be at the level of well-defined and narrowly constrained tasks, but without obviating the need for recovery and reconstruction. Moreover, the sometimes unstated goal of much of computer vision, to develop complex visual processing systems capable of producing symbolic input to AI programs, is alive and well. This is true not only because of the present day successes of the reconstructive approach in computer vision (some of which we have discussed here), but because we believe such an approach holds out the best hope for ultimately understanding and duplicating the adaptive nature of human vision.

[9]We retain some skepticism as to whether object recognition may be construed as purposive, since unlike frogs, human recognition performance is not tied to any salient environmental conditions.

References

[1] Aloimonos, J. (1990). "Purposive and qualitative active vision," *Proceedings of the 10th International Conference on Pattern Recognition,* June 16–21, Atlantic City, New Jersey, Vol. 1, 346–360, Los Alamitos, California: IEEE Computer Society Press.

[2] Aloimonos, J. (Y), Weiss, I., & Bandyopadhyay, A. (1987). "Active vision," *Proceedings of the First International Conference on Computer Vision,* June 8–11, London, England, 35–54, Los Alamitos, California: IEEE Computer Society Press.

[3] Aloimonos, Y., & Rosenfeld, A. (1991). "A response to "Ignorance, myopia, and naivete in computer vision systems" by R. C. Jain and T. O. Binford," *CVGIP: Image Understanding,* Vol. 53, No. 1, 120–124.

[4] Anandan, P. (1989). "A computational framework and an algorithm for the measurement of visual motion," *Int. Journal of Computer Vision*, 2, 283–310.

[5] Anandan, P. (1984). "Computing dense displacement fields with confidence measures in scenes containing occlusion," *SPIE Vol. 521 Intelligent Robots and Computer Vision*, 184–194.

[6] Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). "What some concepts might not be," *Cognition,* 12, 263–308.

[7] Banaji, M. R. (1991). "The physical and mental bases of human thought and the impending death of dualism," To appear in *IEEE Expert,* No. 6.

[8] Biederman, I. (1987). "Recognition–by–components: A theory of human image understanding," *Psychological Review,* 94, 115–147.

[9] Black, M. J. (1991). "A robust gradient method for determining optical flow," submitted, *Computer Vision and Pattern Recognition*, CVPR-92.

[10] Black, M. J., & Anandan, P. (1991). "Robust dynamic motion estimation over time," *Proc. Computer Vision and Pattern Recognition*, CVPR-91, Maui, Hawaii, June, 296–302.

[11] Blake, A., & Zisserman, A. (1987). *Visual Reconstruction,* Cambridge, MA: The MIT Press.

[12] Bolles, R. C., Baker, H. H., & Marimont, D. H. (1987). "Epipolar-plane image analysis: An approach to determining structure from motion," *Internation Journal of Computer Vision*, Vol. 1, No. 1, 7–57.

[13] Braunstein, M. L., Hoffman, D. D., Shaprio, L. R., Andersen, G. J., & Bennett, B. M. (1987). "Minimum points and views for the recovery of three–dimensional structure," *Journal of Experimental Psychology: Human Perception and Performance,* 13 (3), 335–343.

[14] Carey, S, (1985). *Conceptual Change in Childhood.* Cambridge, MA: The MIT Press.

[47] Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, MA: The MIT Press.

[48] White, R. (1989). "Visual thinking in the ice age," *Scientific American,* 261 (1), 92–99.