20

Progress in Computer Vision

Beginning around 1970 computer vision research grew into a highly developed subspecialty of AI, joining other specialized areas such as natural language processing, robotics, knowledge representation, and reasoning (to name just a few of them). In this chapter, I'll describe some of the important advances in computer vision during this period. Some of these were made in pursuit of specific applications in several fields such as aerial reconnaissance, cartography, robotics, medicine, document analysis, and surveillance.[1]

20.1 Beyond Line-Finding

In an earlier chapter, I described some filtering techniques for enhancing image quality and for extracting edges and lines in images. But much more can be done to extract properties of a scene using specific information about the conditions under which images are obtained and general information about the properties of objects likely to be in the scene.

20.1.1 *Shape from Shading*

In what has been called a "back-to-basics" movement, researchers began investigating how information about the physics and geometry of light reflection from surfaces could be used to reveal three-dimensional properties of a scene from a single two-dimensional image. A leader in this study was Berthold K. P. Horn (1943– ; Fig. 20.1). His MIT Ph.D. dissertation derived mathematical methods for determining the shape of an object from its shading.[2] Just as humans perceive an appropriately shaded image of a circle as a sphere, a computer vision system can be made to do so also. Making it do so, using information about the reflective properties of surfaces and the geometry of the imaging process, is what Horn did.

The basic idea of Horn's technique can be explained by referring to Fig. 20.2 in which an infinitesimal piece of surface receives illumination from a light source at an angle equal to i relative to the direction that points perpendicularly away from the surface piece. Suppose a light sensor (such as a TV camera), at an angle g relative to the direction of the light source and at an angle e relative to the direction of the surface, gathers the light reflected from the surface. The amount of light gathered from this surface patch depends on these three angles, the amount of illumination, and the reflectance properties of the surface. (Horn assumed what we would call a

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

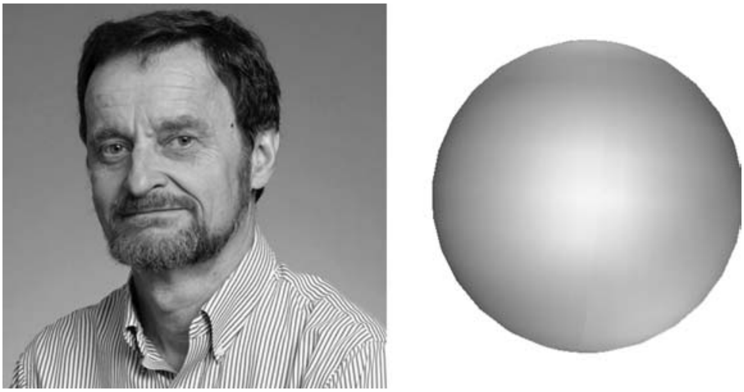

Figure 20.1. Berthold Horn (left) and a shaded circle (right). (Photograph courtesy of Berthold Horn.)

"matte" surface.) Because the amount of light gathered does vary in this manner, the image appears "shaded." Under certain circumstances, and with quite a bit of mathematical manipulation, the direction of the surface can be calculated if the other quantities are known. Then, by knowing the direction for many, many infinitesimal pieces of surface, the overall shape of the surface can be calculated (under the assumption that the surface is relatively smooth with no abrupt discontinuities).

Horn is now a professor of computer science and electrical engineering at MIT and continues to work on several topics related to computer vision. His thesis elicited a flurry of activity in the area of "shape from shading."[3] Several people extended the idea of shape from shading to attempt to calculate shape based on things other than shading, such as from multiple images (stereo), motion, texture, and contour. And, as we shall see in the next few pages, important work was done in extracting more than just the shape of objects.

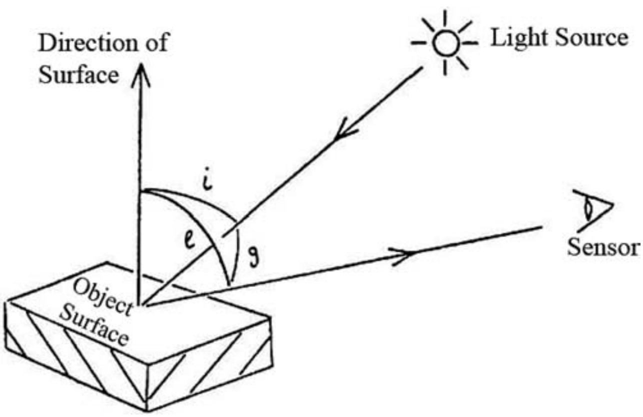

Figure 20.2. Light incident on and reflected by a small piece of a surface. (Illustration used with permission of Berthold Horn.)

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

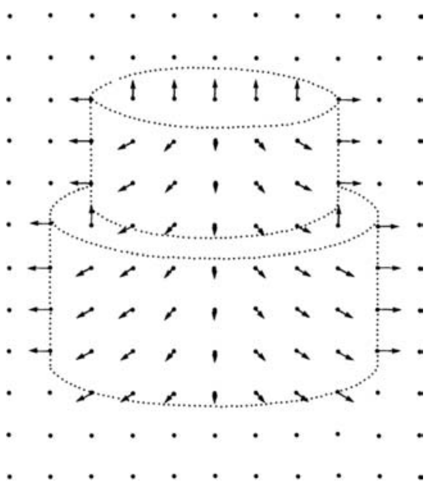

Figure 20.3. A $2\frac{1}{2}$-D sketch. (From David Marr and H. K. Nishihara, "Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes," *Proceedings of the Royal Society of London*, Series B, Biological Sciences, Vol. 200, No. 1140, p. 274, February 23, 1978.)

20.1.2 *The $2\frac{1}{2}$-D Sketch*

Even though a viewer sees only a two-dimensional image of a three-dimensional scene, David Marr (augmenting Horn's ideas) observed that, nevertheless, a viewer is able to infer (and thus perceive) from image shading and other depth cues *some* of the scene's three-dimensional attributes, such as surface shapes, shapes occluding other shapes, abrupt changes between smooth surfaces, and other depth information. Marr called the representation of these attributes a "$2\frac{1}{2}$-D sketch" (because it was not *fully* three dimensional). According to Marr's theory of vision (described in his book[4]), the next step of visual processing, after producing the primal sketch (see p. 133) of blobs and edges, is to produce this $2\frac{1}{2}$-D sketch. An example sketch is shown in Fig. 20.3 in which arrows pointing perpendicularly away from surfaces are superimposed on the primal sketch of an image from which they are inferred.

Finally, according to Marr, the information in the $2\frac{1}{2}$-D sketch, along with stored information about object shapes, would be used to locate specific objects in the image and thus produce a 3-D model of the scene. I'll describe what he had to say about that process shortly.

20.1.3 *Intrinsic Images*

Two researchers at SRI, Jay Martin Tenenbaum (1943– ; Fig. 20.4) and Harry Barrow (recently relocated from Edinburgh), developed some image-processing techniques quite similar to those used in producing the $2\frac{1}{2}$-D sketch.[5] They noted that the intensity value at each pixel of an image resulted from a tangled combination of several factors, including properties of the ambient illumination and reflective and geometric properties of objects in the scene. They thought that these factors could be untangled to recover important three-dimensional information about the scene.

Barrow and Tenenbaum proposed that each of these factors (all of which influenced intensity) could be represented by imaginary images that they called "intrinsic images." These images were to consist of a grid of "pixels" overlaying a projection

Figure 20.4. Jay Martin Tenenbaum (left) and Harry Barrow (right). (Photographs courtesy of J. Martin Tenenbaum and of Harry Barrow.)

of the scene and in registration with the intensity image. One intrinsic image, for example, was an illumination image. It consisted of pixels whose values were the amounts of illumination falling on the pixels of the projected scene. These values, of course, were not known, but Barrow and Tenenbaum proposed that they could be estimated from the intensity image and from the other intrinsic images.

As examples, I show a set of such intrinsic images in Fig. 20.5. The actual image of intensity values is shown at the top. The known value of a pixel in that image depends on the unknown values of pixels in the intrinsic images below. In fact, the values of the pixels in all of the images, intrinsic and actual, are interdependent. The arrows in the figure reflect that fact. (There should also be some arrows going up.) Based on the values of pixels in some of the images, the values of others can be computed by using known physical relationships, constraints among the images, and other reasonable assumptions. These values, in turn, allow the computation of others. In essence, these computations "propagate" pixel values throughout the set of intrinsic images (much like how levels in the Blackboard architecture affect other levels). As Barrow and Tenenbaum later summarized their method, "We envisaged this recovery process as a set of interacting parallel local computations, more like solving a system of simultaneous equations by relaxation than like a feed forward sequence of stages."[6] Barrow and Tenenbaum also used some of their ideas about intrinsic images to work on the problem of interpreting line drawings as three-dimensional surfaces.[7]

Barrow and Tenenbaum intended their work to be useful not only in computer vision but also as a potential model of "precognitive" vision processes in humans. However, in a 1993 "retrospective" about their work they wrote[8]

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

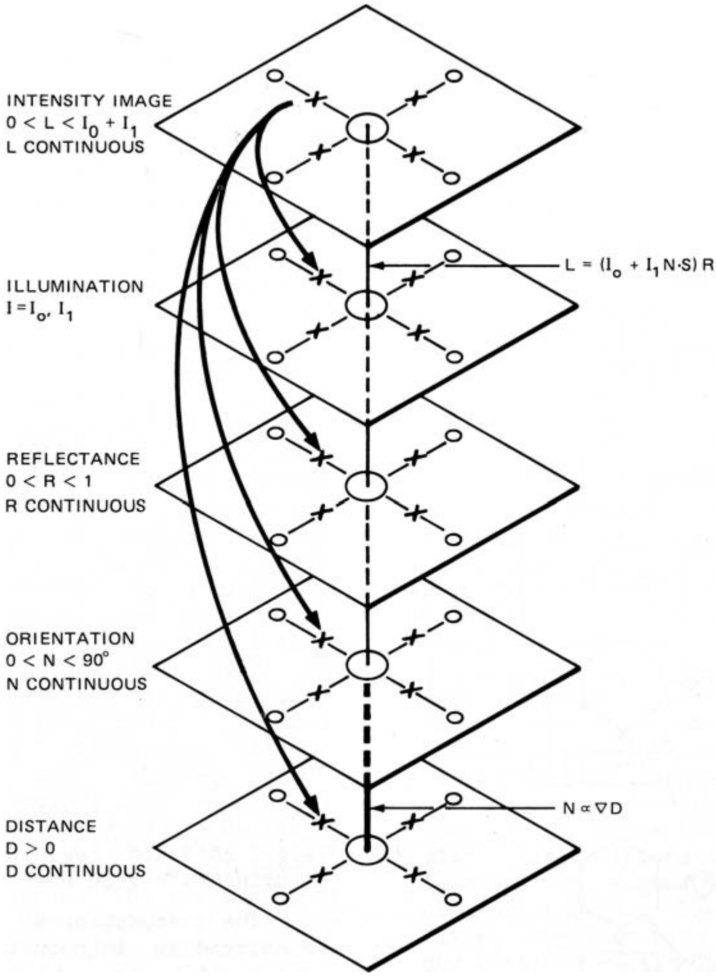

Figure 20.5. Intrinsic images. (Used with permission of Harry Barrow and Jay M. Tenenbaum.)

Despite the maturity of computational vision and the rapid developments in neural systems, we still have a long way to go before we can come close to our goal of understanding visual perception. To do so we will need to draw upon what we have learned in many fields, including neuroscience, neural networks, experimental psychology and computational vision.

20.2 Finding Objects in Scenes

20.2.1 *Reasoning about Scenes*

Even before the development of shape-from-shading and other methods for recovering depth information from scenes, a number of researchers had worked on methods for finding objects in scenes. I described many of these techniques in Section 9.3.

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

During the early 1970s, Thomas Garvey completed a Stanford Ph.D. thesis on a system for locating objects, such as desks, chairs, and wastebaskets, in images of office scenes.[9] As Garvey wrote in his summary,

> The system uses information about the appearances of objects, about their interrelationships, and about available sensors to produce a plan for locating specified objects in images of room scenes.

In related work, Barrow and Tenenbaum developed a system, called MSYS, for reasoning about scenes "in which knowledge sources compete and cooperate until a consistent explanation of the scene emerges by consensus."[10] MSYS analyzed images of office scenes and attempted to find the most likely interpretation for the regions in an image (desk top, back of chair, floor, doorway, and so on) given a number of candidate interpretations and their probabilities. Knowing relationships between regions (such as "chair backs are usually adjacent to chair seats"), MSYS tried to find the most likely overall set of region interpretations.

An example of a scene considered by MSYS is shown in Fig. 20.6. Some of the regions in the scene have been detected and labeled with possible interpretations.

As Barrow and Tenenbaum wrote, MSYS's reasoning might proceed as follows:

> Regions PIC, WBSKT, and CBACK cannot be WALL or DOOR, because their brightnesses are much less than that along the top edge of the image vertically above them, which violates [knowledge about the brightness of walls and doors]. Consequently, region PIC must be the PICTURE, WBSKT must be WASTEBASKET, and CBACK must be CHAIRBACK.
>
> Region LWALL and RWALL must then be WALL, since they are adjacent to region PIC, and DOOR cannot be adjacent to PICTURE.
>
> Region DR cannot be WALL because all regions labeled WALL are required to have the same brightness. Therefore, region DR must be DOOR.

20.2.2 *Using Templates and Models*

Much of the early work on object recognition was based on using object "templates" that could be matched against images. Martin A. Fischler and Robert A. Elschlager elaborated this idea by using "stretchable templates" that permitted more powerful matching techniques. They used these to find objects such as faces or particular terrain features in photographs containing such objects.[11] The process depended on having a general representation for the object being sought and then a process for matching that representation against the photograph. Their representations were based on breaking an object down into a number of primitive parts and "specifying an allowable range of spatial relations which these 'primitive parts' must satisfy for the object to be present." For the object to be present in a picture, "it is required that [the] primitives occur (or at least that some significant subset of them occurs), and also that they occur within a certain spatial relationship one to the other . . . " As Fischler and Elschlager pointed out, it is usually the case that determining whether or not some of the parts occur depends on whether or not the whole object occurs,

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

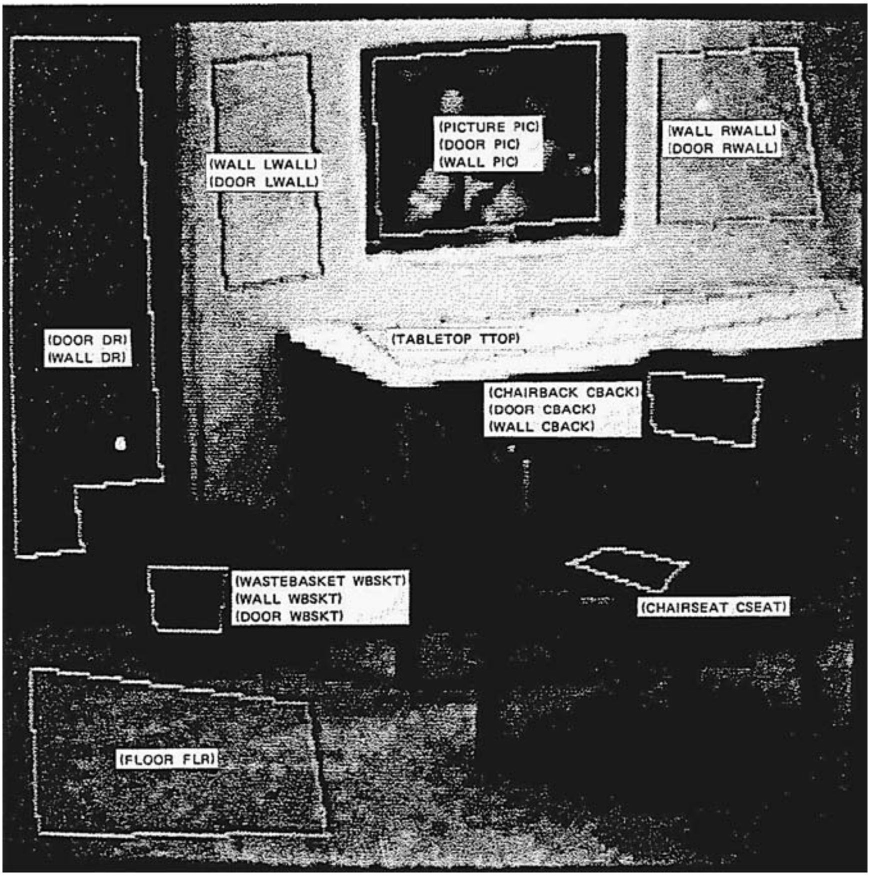

Figure 20.6. An MSYS scene with some regions detected and labeled. (Illustration used with permission of SRI International.)

and vice versa. The main contribution of their paper was the development of a dynamic-programming-style method for dealing with this circularity.

Earlier I had described David Marr's work on processes for producing a primal sketch and a $2\frac{1}{2}$-D sketch. These were the first two stages in Marr's theory of vision. He argued that these stages could uncover important shape information without specific knowledge of the shapes of objects likely to be in a scene. He had written:[12]

> Most early visual processes extract information about the visible surfaces directly, without particular regard to whether they happen to be part of a horse, or a man, or a tree. . . . As for the question of what additional knowledge should be brought to bear, general knowledge must be enough – general knowledge embedded in the early visual processes as general constraints, together with the geometrical consequences of the fact that the surfaces co-exist in three-dimensional space.

Specific knowledge about shapes, he argued, should be utilized in a third stage. It is this stage that uses three-dimensional models of objects. He proposed using a

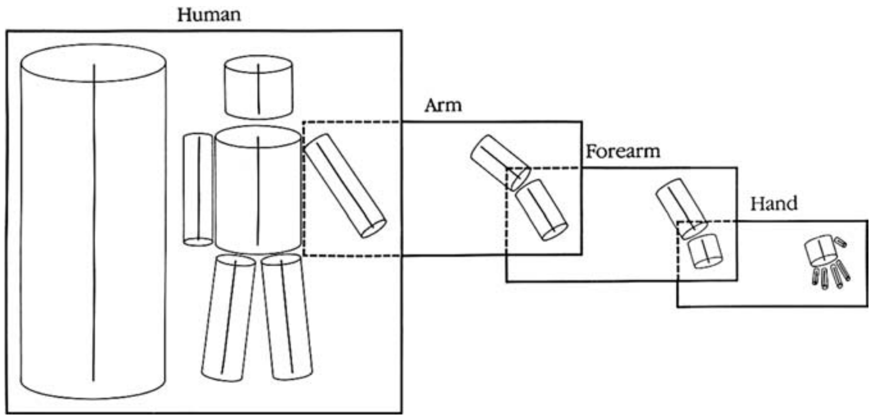

Figure 20.7. An example of one of Marr's 3-D model hierarchies. (From David Marr, *Vision*, San Francisco: W. H. Freeman and Co., p. 306, 1982.)

hierarchy of models in which a gross model is decomposed into subparts and these into subsubparts and so on. For example, the shape of a human might be modeled as in Fig. 20.7. Each box corresponds to a 3-D model and its submodel. On the left side of the box is an axis-oriented model; on the right side is how that model is represented as submodels. (Directions of the axes can be adjusted to fit matching parts of the image.)

In this third stage, comparing models of this sort with shape information and other 3-D information contained in the $2\frac{1}{2}$-D sketch helps to identify and locate objects in a scene. For Marr, vision was "the *process* of discovering from images what is present in the world and where it is."[13]

Marr was not the first to suggest the use of cylinders as models of parts of objects. In a 1971 IEEE conference paper, Thomas O. Binford (1936–) introduced the idea of "generalized cylinders" (sometimes called "generalized cones").[14] A later paper defined them as follows: "A generalized cone is defined by a planar cross section, a space curve spine, and a sweeping rule. It represents the volume swept out by the cross section [not necessarily a circular one] as it is translated along [an axis called a spine], held at some constant angle to the spine, and transformed according to the sweeping rule."[15]

Binford had several Stanford Ph.D. students who used models to help identify objects in scenes. Of these I might mention Gerald J. Agin,[16] Ramakant Nevatia,[17] and Rodney A. Brooks (1954–),[18] all of whom contributed to what came to be called "model-based vision." (Brooks later became a professor at MIT, where he worked on other topics. His subsequent work will be discussed later.)

Brooks's ACRONYM system[19] used generalized cones to model several different kinds of objects. ACRONYM used these models to help identify and locate objects in images. Some examples of the kinds of generalized cones that can be used as building blocks of models and model objects are shown in Fig. 20.8.

Other views regarding what vision is all about competed with those of Marr and others who were attempting to use vision to reconstruct entire scenes. Some,

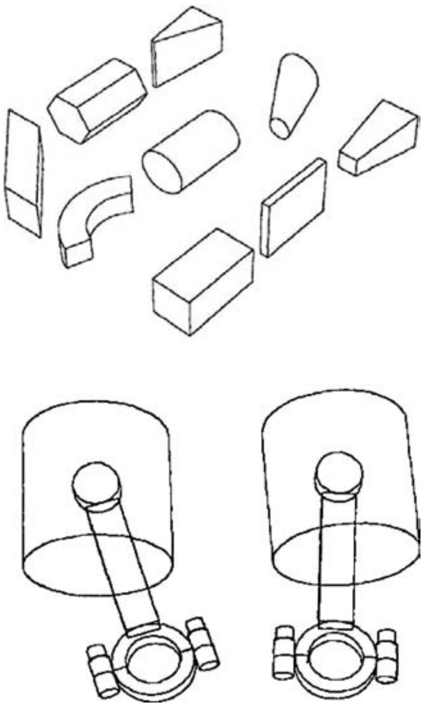

Figure 20.8. Primitive generalized cones and piston models constructed from generalized cones. (From Rodney A. Brooks, "Symbolic Reasoning among 3-D Models and 2-D Images," *Artificial Intelligence*, Vol. 17, Nos. 1–3, pp. 285–348, 1981.)

especially those involved in robotics, claimed that the purpose of vision was to perceive just what was required to guide action. Many of the vision routines in Shakey were embedded in its action programs. Professor Yiannis Aloimonos at the University of Maryland is one of the researchers advocating this "purposive" or "interactive" approach. He claims that the goal of vision is action. When vision is "considered in conjunction with action, it becomes easier." He goes on to explain that "the descriptions of space-time that the system needs to derive are not general purpose, but are purposive. This means that these descriptions are good for restricted sets of tasks, such as tasks related to navigation, manipulation and recognition."[20] In the neuroscience community, to which Marr wanted to make a contribution, there were Patricia S. Churchland, V. S. Ramachandran, and Terrence J. Sejnowski, who later wrote "What is vision for? Is a perfect internal recreation of the three-dimensional world really necessary? Biological and computational answers to these questions lead to a conception of vision quite different from pure vision [as advocated by Marr]. Interactive vision . . . includes vision with other sensory systems as partners in helping to guide actions."[21]

In any case, models still play an important role in computer vision. (However, one prominent vision researcher told me that the "residue of model-based vision is close to zero,"[22] and another told me that "most current robotic systems use vision hacks" instead of general-purpose, science-based scene-analysis methods.[23])

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

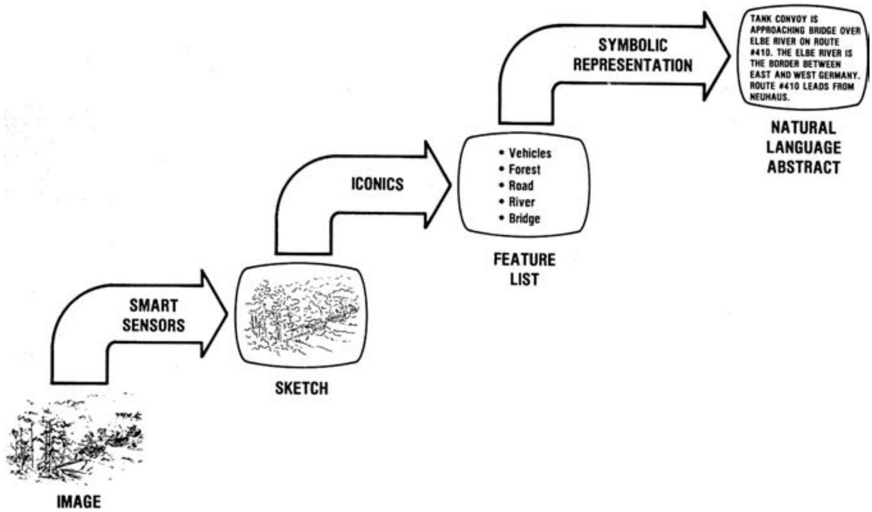

Figure 20.9. An illustration of IU goals. (Illustration used with permission of SAIC.)

20.3 DARPA's Image Understanding Program

Much of the computer vision work in the United States was being funded by DARPA, and there were concerns among vision researchers (as always) about continuing support. Tenenbaum recalls attending a DARPA meeting in 1974 where the future of computer vision research was being discussed. The program officer monitoring DARPA-supported vision work, Air Force Major David L. Carlstrom, was at the meeting and was interested in pulling together the various efforts in the field. Because DARPA had been supporting work in this area for some years, Carlstrom needed a new name that would indicate that DARPA was starting something new. Tenenbaum told me that he recommended to Carlstrom that the new initiative be called "the image understanding program."[24] (Recall that there was already an ongoing DARPA-supported effort in speech understanding, so the phrase sounded "DARPA-friendly.")

In 1976, DARPA launched its Image Understanding (IU) program. It grew to be a major effort composed of the leading research laboratories doing work in this area as well as "teams" pairing a university with a company. The individual labs participating were those at MIT, Stanford, University of Rochester, SRI, and Honeywell. The university/industry teams were USC–Hughes Research Laboratories, University of Maryland–Westinghouse, Inc., Purdue University–Honeywell, Inc., and CMU–Control Data Corporation.

Regular workshops were held to report progress. The proceedings of one held in April 1977 stated the goals of the program: "The Image Understanding Program is planned to be a five year research effort to develop the technology required for automatic and semiautomatic interpretation and analysis of military photographs and related images."[25] DARPA's ultimate goal for the IU program was well captured by the illustration on the cover of that proceedings, shown in Fig. 20.9.

As the diagram implies, military commanders would like computer vision systems to be able to analyze a photograph and to produce a written description of its important components and their relationships.

Some of the computer vision research that I have already described, such as work on the $2\frac{1}{2}$-D sketch, intrinsic images, generalized cylinders, and ACRONYM, was supported by the IU Program. But there was always some tension between DARPA's goals and those of people doing computer vision research. DARPA wanted the program to produce "field-able" systems. J. C. R. Licklider emphasized this point at a preliminary IU workshop in March 1975:[26]

> At the end of the five year period the technology developed must be in a state in which it can be utilized by the DoD components to solve their specific problems without requiring a significant research effort to figure our how to apply the technology to the specific problems. For this reason, the program must result in a demonstration at the end of the five year period that an important DoD problem has been solved.

Air Force Major Larry Druffel at DARPA assumed leadership of the IU program in 1978. In November 1978, he advised "The prudent approach is to consolidate those techniques which are sufficiently mature to transfer to DoD agencies."[27] By 1979, the program's goals had expanded to include cartography and mapping. A "memorandum of understanding" (MOU) between DARPA and the Defense Mapping Agency (DMA) was concluded to support automatic mapping efforts through the development of a DARPA/DMA "testbed." In November 1979, Druffel wrote[28]

> Plans are progressing for a demonstration system to evaluate the maturity of IU technology by automating mapping, charting, and geodesy functions. While focussing on specific cartographic photointerpretation functions, the system should offer the entire image exploitation community an opportunity to assess the future application of Image Understanding methodologies to their specific problem.

The "five-year" program did not end in 1981. It continued under the DARPA leadership of Navy Commander Ron Ohlander, Air Force Lt. Col. Robert L. Simpson Jr., and others until approximately 2001. In 1985 Simpson summarized some of its accomplishments:[29]

> Originally conceived as a five year program in 1975 by Lt. Col. David Carlstrom, the first several years of IU established the strong base of low-level vision techniques and knowledge-based subsystems that began to differentiate computer vision from what is usually called "image processing." In the late 1970s and early 1980s, under the direction of Lt. Col. Larry Druffel, the program saw the development of model-based vision systems such as ACRONYM and demonstration of IU techniques in more meaningful concept demonstrations such as the DARPA/DMA image understanding testbed. These demonstrations and their potential for future military use warranted the continuation of the IU program beyond its initial five year lifespan. Under Cmd. Ron Ohlander, IU technology continued to mature to the point that the DARPA Strategic Computing Program could justify a major application, the autonomous land vehicle.

As Ohlander said, the IU program was extended beyond its projected five-year lifetime. It is said that even as early as 1984, DARPA had spent over $4 million on this effort.[30] One potential application was computer vision for robot-controlled military vehicles – a component of DARPA's "Strategic Computing" program. I'll describe that application and others in more detail in later chapters.

As a growing subspecialty of artificial intelligence, papers on computer vision began to appear in new journals devoted to the subject, including *Computer Vision and Image Understanding* and *IEEE Transactions on Pattern Analysis and Machine Intelligence*. The field's textbooks around this time included *Pattern Classification and Scene Analysis*[31] and two books titled *Computer Vision*.[32]

Notes

1. For an extensive list of computer vision applications see the CVonline Web site at http://homepages.inf.ed.ac.uk/rbf/CVonline/applic.htm. [258]
2. Berthold K. P. Horn, "Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View," MIT Department of Electrical Engineering Ph.D. thesis, MIT Artificial Intelligence Laboratory Technical Report 232, November 1970; available online at http://people.csail.mit.edu/bkph/AIM/AITR-232-OCR-OPT.pdf. In his thesis, Horn credits Thomas Rindfleisch's Rindfleisch, Thomas 1966 work on using image brightness in studies of lunar topography. [258]
3. For a modern discussion of the problem, see, for example, Emmanuel Prados and Olivier Faugeras, "Shape from Shading," in N. Paragios, Y. Chen, and O. Faugeras (eds.), *Handbook of Mathematical Models in Computer Vision*, pp. 375–388, New York: Springer-Verlag, 2006; available online at http://perception.inrialpes.fr/Publications/2006/PF06a/chapter-prados-faugeras.pdf. [259]
4. David Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco: W. H. Freeman and Co., 1982. [260]
5. Harry G. Barrow and Jay Martin Tenenbaum, "Recovering Intrinsic Scene Characteristics from Images," in A. Hanson and E. Riseman (eds.), *Computer Vision Systems*, pp. 3–26, New York: Academic Press, 1978. Available online at http://web.mit.edu/cocosci/Papers/Barrow-Tenenbaum78.pdf and at http://www.ai.sri.com/pubs/files/737.pdf. [260]
6. Harry G. Barrow and Jay Martin Tenenbaum, "Retrospective on 'Interpreting Line Drawings as Three-Dimensional Surfaces,'" *Artificial Intelligence*, Vol. 59, Nos. 1–2, pp. 71–80, 1993. [261]
7. Harry G. Barrow and Jay Martin Tenenbaum, "Interpreting Line Drawings as Three-Dimensional Surfaces," *Artificial Intelligence*, Vol. 17, pp. 75–116, 1981. Available online at http://web.mit.edu/cocosci/Papers/Barrow-Tenenbaum81.pdf. [261]
8. Harry G. Barrow and Jay Martin Tenenbaum, "Retrospective on 'Interpreting Line Drawings as Three-Dimensional Surfaces,'" *Artificial Intelligence*, Vol. 59, Nos. 1–2, pp. 71–80, 1993. [261]
9. Thomas D. Garvey, "Perceptual Strategies for Purposive Vision," Stanford University Ph.D. thesis, published as SRI International AI Center Technical Note 117, September 1976. Abstract available online at http://www.ai.sri.com/pub_list/759. [263]
10. Harry G. Barrow and J. Martin Tenenbaum, "MSYS: A System for Reasoning about Scenes," SRI International AI Center Technical Note 121, April 1976. Available online at http://www.ai.sri.com/pubs/files/757.pdf. [263]
11. Martin A. Fischler and Robert A. Elschlager, "The Representation and Matching of Pictorial Structures," *IEEE Transactions on Computers*, Vol. C-22, No. 1, pp. 67–92, January 1973. [263]
12. David Marr, *op. cit.*, pp. 272–4. [264]
13. David Marr, *op. cit.*, pp. 23–60. [265]

https://doi.org/10.1017/CBO9780511819346.024 Published online by Cambridge University Press

14. Thomas O. Binford, "Visual Perception by Computer," *Proceedings of the IEEE Conference on Systems and Control*, Miami FL, December 1971. [265]
15. Rodney A. Brooks, "Symbolic Reasoning among 3-D Models and 2-D Images," *Artificial Intelligence*, Vol. 17, Nos. 1–3, pp. 285–348, 1981. [265]
16. Gerald J. Agin, "Representation and Description of Curved Objects," Stanford University Ph.D. thesis, published as Stanford Artificial Intelligence Project Memo AIM-173, October 1972. See also Gerald J. Agin and Thomas O. Binford, "Computer Descriptions of Curved Objects," *Proceedings of the Third International Joint Conference on Artificial Intelligence*, pp. 629–640, August 1973; later published as Gerald J. Agin and Thomas O. Binford, "Computer Descriptions of Curved Objects," *lEEE Transactions on Computers*, Vol. 25, No. 4, April 1976. [265]
17. Ramakant Nevatia, "Structured Descriptions of Complex Curved Objects for Recognition and Visual Memory," Stanford University Department of Electrical Engineering Ph.D. thesis, published as Stanford Artificial Intelligence Laboratory Memo AIM-250, October 1974. [265]
18. Rodney A. Brooks, "Symbolic Reasoning among 3-D Models and 2-D Images," Stanford University Computer Science Department Ph.D. thesis, 1981, published as Stanford CS Department Report STAN-CS-81-861. Also published as Rodney A. Brooks, *op. cit.*. [265]
19. The system was first reported in Rodney A. Brooks, Russell Greiner, and Thomas O. Binford, "The ACRONYM Model-Based Vision System," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, pp. 105–113, Tokyo, 1979. A later revised version was reported in Brooks's *Artificial Intelligence* paper just cited. [265]
20. From his Web page at http://www.cfar.umd.edu/˜yiannis/. [266]
21. Patricia S. Churchland, V. S. Ramachandran, and Terrence J. Sejnowski, "A Critique of Pure Vision," in Christof Koch and Joel L. Davis (eds.), *Large-Scale Neuronal Theories of the Brain*, pp. 23–65, Cambridge, MA: MIT Press, 1994. Available online at http://philosophy.ucsd.edu/faculty/pschurchland/papers/kochdavis94critiqueofpurevision.pdf. [266]
22. Martin A. Fischler, private communication, August 1, 2007. [266]
23. Jay Martin Tenenbaum, private communication, July 31, 2007. [266]
24. Private communication, July 31, 2007. [267]
25. Lee S. Bauman (ed.), *Proceedings: Image Understanding Workshop*, Science Applications, Inc., Report No. SAI-78-549-WA, April 1977. [267]
26. Quoted in the Foreword of the *Proceedings: Image Understanding Workshop*, published by Science Applications, Inc., May 1978. [268]
27. Quoted in the Foreword of the *Proceedings: Image Understanding Workshop*, published by Science Applications, Inc., November 1978. [268]
28. Quoted in the Foreword of the *Proceedings: Image Understanding Workshop*, published by Science Applications, Inc., November 1979. [268]
29. Quoted in the Foreword of the *Proceedings: Image Understanding Workshop*, published by Science Applications International Corporation,December 1985. [268]
30. Alex Roland with Philip Shiman, *Strategic Computing: DARPA and the Quest for Machine Intelligence*, p. 220, Cambridge, MA: MIT Press, 2002. [268]
31. Richard O. Duda and Peter E. Hart, *Pattern Classification and Scene Analysis*, New York: John Wiley and Sons, Inc., 1973. [269]
32. Michael Brady, *Computer Vision*, Amsterdam: North-Holland Publishing Co., 1981, and Dana H. Ballard and C. M. Brown, *Computer Vision*, New York: Prentice Hall, Inc., 1982. [269]