Automation of Sight: From Photography to Computer Vision

author: Lev Manovich

year: 1997

Prologue

Nothing perhaps symbolizes mechanization as dramatically as the first assembly lines installed by Henry Ford in U.S. factories in 1913. It seemed that mechanical modernity was at its peak. Yet, in the same year the Spanish inventor Leonardo Torres y Quevedo had already advocated the industrial use of programmed machines. [1] He pointed out that although automatons existed before, they were never used to perform useful work:

The ancient automatons ...imitate the appearance and movement of living beings, but this has not much practical interest, and what is wanted is a class of apparatus which leaves out the merely visible gestures of man and attempts to accomplish the results which a living person obtains, thus replacing a man by a machine. [2]

With mechanization, work is performed by a human but his or her physical labor is augmented by a machine. Automation takes mechanization one step further: the machine is programmed to replace the functions of human organs of observation, effort, and decision.

Mass automation was made possible by the development of digital computers during World War II and thus became synonymous with computerization. The term "automation" was coined in 1947; and in 1949 Ford began the construction of the first automated factories.

Barely a decade later, automation of imaging and of vision was well underway. By the early 1960s, construction of static and moving two-dimensional and perspectival images, correction of artifacts in photographs, the identification of objects from their images, and many other visual tasks were already handled by computers. A number of new disciplines were emerging as well — computer image processing, computer vision, computer graphics, computer-aided design.

What these new disciplines had all in common is that they employed perspectival images. In other words, automation of imaging and vision was first of all directed at perspectival sight.

The reasons for this are two-fold. On the one hand, by the time digital computers became available, modern society was already heavily invested in lens-based methods of image gathering (photography, film, television) which all produced perspectival images. Therefore, it is not surprising that it would want first of all to automate various uses of such images in order to obtain a new return from its investment. On the other hand, the automation of perspectival sight has already begun well before this century with the development of perspective machines, descriptive and perspective geometry and, of course, photography. Computers certainly proved to be very fast perspectival machines, but they were hardly the first.

Perspective, Perspectival Machines, Photography

From the moment of adaptation of perspective, artists and draftsmen have attempted to aid the laborious manual process of creating perspectival images. [3] Between the sixteenth and the nineteenth century various "perspectival machines" were constructed. They were used to construct particularly challenging perspectival images, to illustrate the principles of perspective, to help students learn how to draw in perspective, to impress artists' clients, or to serve as intellectual toys. Already in the first decades of the sixteenth century, Dürer described a number of such machines. [4] One device is a net in the form of a rectangular grid, stretched between the artist and the subject. Another uses a string representing a line of sight. The string is fixed on one end, while the other end is moved successively to key points on the subject. The point where the string crosses the projection plane, defined by a wooden frame, is recorded by two crossed strings. For each position, a hinged board attached to the frame is moved and the point of intersection is marked on its surface. It is hard to claim that such a device, which gave rise to many variations, made the creation of perspectival images more efficient, however, the images it helped to produce had reassuring mechanical precision. Other major types of perspectival machines that appeared subsequently included the perspectograph, pantograph, physionotrace, and optigraph.

Why manually move the string imitating the ray of light from point to point? Along with perspectival machines a whole range of optical apparatuses was in use, particularly for depicting landscapes and conducting topographic surveys. They included versions of camera obscura from large tents to smaller, easily transportable boxes. After 1800, the artist's ammunition was strengthened by camera lucida, patented in 1806. [5] Camera lucida utilized a prism with two reflecting surfaces at 135˚. The draftsman carefully positioned his eye to see both the image and the drawing surface below and traced the outline of the image with a pencil.

Optical apparatuses came closer than previous perspectival devices to the automation of perspectival imaging. However, the images produced by camera obscura or camera lucida were only ephemeral and considerable effort was still required to fix these images. A draftsman had to meticulously trace the image to transform it into the permanent form of a drawing.

With photography, this time-consuming process was finally eliminated. The process of **imaging physical reality**, the creation of perspectival representations of real objects was now automated. Not surprisingly, photography was immediately employed in a variety of fields, from aerial photographic surveillance to criminal detection. Whenever the real had to be captured, identified, classified, stored, photography was put to work.

Photography automated one use of perspectival representation — but not others. According to Bruno Latour, the greatest advantage of perspective over other kinds of representations is that it establishes a "four-lane freeway" between physical reality and its representation. [6] We can combine real and imagined objects in a single geometric model and go back and forth between reality and the model. By the twentieth century, the creation of a geometric model of both existing and imagined reality still remained a time-consuming manual process, requiring the techniques of perspectival and analytical geometry, pencil, ruler, and eraser. Similarly, if one wanted to visualize the model in perspective, hours of drafting were required. And to view the model from another angle, one had to start all over again. The automation of geometrical modeling and display had to wait for the arrival of digital computers.

3-D Computer Graphics: Automation of Perspectival Imaging

Digital computers were developed towards the end of World War II. The automation of the process of constructing perspectival images of both existent and non-existent objects and scenes followed quite soon. [7] By the early 1960s Boeing designers already relied on 3-D computer graphics for the simulation of landings on the runway and of pilot movement in the cockpit. [8]

By automating perspectival imaging, digital computers completed the process which began in the Renaissance. This automation became possible because perspectival drawing has always been a step-by-step procedure, an algorithm involving a series of steps required to project coordinates of points in 3-D space onto a plane. Before computers the steps of the algorithm were executed by human draftsmen and artists. With a computer, these steps can be executed automatically and, therefore, much more efficiently.

The details of the actual perspective-generating algorithm which could be executed by a computer were published in 1963 by Lawrence G. Roberts, then a graduate student at MIT. [9] The perspective-generating algorithm constructs perspectival images in a manner quite similar to traditional perspectival techniques. In fact, Roberts had to refer to German textbooks on perspectival geometry from the early 1800s to get the mathematics of perspective. [10] The algorithm reduces reality to solid objects, and the objects are further reduced to planes defined by straight lines. The coordinates of the endpoint of each line are stored in a computer. Also stored are the parameters of a virtual camera — the coordinates of a point of view, the direction of sight, and the position of a projection plane. Given this information, the algorithm generates a perspectival image of an object, point by point.

The subsequent development of computer graphics can be seen as the struggle to automate other operations involved in producing perspectival stills and moving images. The computerization of perspectival construction made possible the automatic generation of a perspectival image of a geometric model as seen from an arbitrary point of view — a picture of a virtual world recorded by a virtual camera. But, just like with the early perspectival machines described by Dürer, early computer graphics systems did not really save much time over traditional methods. To produce a film of a simulated landing, Boeing had to supplement computer technology with manual labor. As in traditional animation, twenty-four plots were required for each second of film. These plots were computer-generated and consisted of simple lines. Each plot was then hand-colored by an artist. Finished plots were filmed, again manually, on an animation stand. [11] Gradually, throughout the 1970s and the 1980s, the coloring stage was automated as well. Many algorithms were developed to add the full set of depth cues to a synthetic image — hidden line and hidden surface removal, shading, texture, atmospheric perspective, shadows, reflections, and so on. [12]

At the same time, to achieve interactive perspectival display, special hardware was built. Each step in the process of 3-D image synthesis was delegated to a special electronic circuit: a clipping divider, a matrix multiplier, a vector generator. Later on, such circuits became specialized computer chips, connected together to achieve real-time, high resolution, photorealistic 3-D graphics. Silicon Graphics Inc., one of the major manufacturers of computer graphics hardware, labeled such a system "geometry engine".

The term appropriately symbolizes the second stage of the automation of perspectival imaging. At the first stage, the photographic camera, with perspective physically built into its lens, automated the process of creating perspectival images of existing objects. Now, with the perspectival algorithm and other necessary geometric operations embedded in silicon, it becomes possible to display and interactively manipulate models of non-existent objects as well.

Computer Vision: Automation of Sight

In his papers, published between 1963 and 1965, Roberts formalized the mathematics necessary for generating and modifying perspective views of geometric models on the computer. This, writes William J. Mitchell, was "an event as momentous, in its way, as Brunelleschi's perspective demonstration." [13] However, Roberts developed techniques of 3-D computer graphics having in mind not the automation of perspectival imaging but another, much more daring goal — "to have the machine recognize and understand photographs of three-dimensional objects." [14] Thus, the two fields were born simultaneously: 3-D computer graphics and computer vision, **automation of imaging and of sight**.

The field of computer vision can be seen as the culmination of at least two centuries-long histories. The first is the history of mechanical devices designed to aid human perception, such as Renaissance perspectival machines. This history reaches its final stage with computer vision, which aims to replace human sight altogether. The second is the history of automata, whose construction was especially popular in the seventeenth and eighteenth centuries. Yet, despite the similarity in appearance, there is a fundamental difference between Enlightenment automata which imitated human's or animal's bodily functions and the modern-day robots equipped with computer vision systems, artificial legs, arms, etc. As noted by Leonardo Torres, old automata, while successfully copying the appearance and movement of living beings, had no economic value. Indeed, such voice synthesis machines as Wolgang von Kempelen's 1778 device which directly imitated the functioning of the oral cavity, or Abbé Mical's **Têtes Parlantes** (1783) operated by a technician hiding offstage and pressing a key on a keyboard were used only for entertainment. [15] When in 1913 Torres called for automata that would "accomplish the results which a living person obtains, thus replacing a man by a machine" he was expressing a fundamentally new idea of using automata for productive labor. A few years later, the brother of the Czech writer Karel Capek coined the word **robot** from the Czech word **robota**, which means "forced labor". [16] Capek's play **R.U.R.** (1921) and Fritz Lang's **Metropolis** (1927) clearly demonstrate this new association of automata with physical industrial labor.

Therefore, it would be erroneous to conclude that, with computer vision, twentieth-century technology simply added the sense of sight to eighteenth-century mechanical statues. But even to see computer vision as the continuation of Torres', Capek's, or Lang's ideas about industrial automation which replaces manual labor would not be fully accurate. The idea of computer vision became possible and the economic means to realize this idea became available only with the shift from industrial to post-industrial society after World War II. The attention turned from the automation of the body to the automation of the mind, from physical to mental labor. This new concern with the automation of mental functions such as vision, hearing, reasoning, problem-solving is exemplified by the very names of the two new fields that emerged during the 1950s and 1960s — artificial intelligence and cognitive psychology. The latter gradually replacing behaviorism, the dominant psychology of the "Fordism" era. The emergence of the field of computer vision is a part of this cognitive revolution, a revolution which was financed by the military escalation of the Cold War. [17] This connection is solidified in the very term "artificial intelligence" which may refer simultaneously to two meanings of "intelligence": reason, the ability to learn or understand, and information concerning an enemy or a possible enemy or an area. Artificial intelligence: artificial reason to analyze collected information, collected intelligence.

In the 1950s, faced with the enormous task of gathering and analyzing written, photographic, and radar information about the enemy, the CIA and the NSA (National Security Agency) began to fund the first artificial intelligence projects. One of the earliest projects was a Program for Mechanical Translation, initiated in the early 1950s in the attempt to automate the monitoring of Soviet communications and media. [18] The work on mechanical translation was probably the major cause of many subsequent developments in modern linguistics, its move towards formalization; it can be discerned in Noam Chomsky's early theory which, by postulating the existence of language universals in the domain of grammar, implied that translation between arbitrary human languages could be automated. The same work on mechanical translation was also one of the catalysts in the development of the field of pattern recognition, the precursor to computer vision. Pattern recognition is concerned with automatically detecting and identifying predetermined patterns in the flow of information. A typical example is character recognition, the first stage in the process of automating translation. Pattern recognition was also used in the U.S.

for the monitoring of Soviet radio and telephone communication. Instead of listening to every transmission, an operator would be alerted if computer picked up certain words in the conversation.

As a "logistics of perception" came to dominate modern warfare and surveillance and as the space race began, image processing became another major new field of research. [19] Image processing comprises techniques to improve images for human or computer interpretation. In 1964, the space program for the first time used image processing to correct distortions in the pictures of the Moon introduced by an onboard television camera of Ranger 7. [20] In 1961, the National Photographic Interpretation Center (NPIC) was created to produce photoanalysis for the rest of the U.S. intelligence community and, as Manual De Landa points out, by the end of the next decade computers "were routinely used to correct for distortions made by satellite's imaging sensors and by atmospheric effects, sharpen out-of-focus images, bring multicolored single images out of several pictures taken in different spectral bands, extract particular features while diminishing or eliminating their backgrounds altogether…" De Landa also notes that computer analysis of photographic imagery became the only way to deal with the pure volume of intelligence being gathered: "It became apparent during the 1970s that there is no hope of keeping up with the millions of images that poured into NPIC … by simply looking at them the way they had been looked at in World War II. The computers therefore also had to be taught to compare new imagery of a given scene with old imagery, ignoring what had not changed and calling the interpreter's attention to what had." [21]

The techniques of image processing, which can automatically increase an image's contrast, remove the effects of blur, extract edges, record differences between two images, and so on, greatly eased the job of human photoanalysts. And the combining of image processing with pattern recognition made it possible in some cases to delegate the analysis of photographs to a computer. For instance, the technique of pattern matching used to recognize printed characters can also be used to recognize objects in a satellite photograph. In both cases, the image is treated as consisting of two-dimensional forms. The contours of the forms are extracted from the image are then compared to templates stored in computer memory. If a contour found in the image matches a particular template, the computer signals that a corresponding object is present in a photograph.

A general-purpose computer vision program has to be able to recognize not just two-dimensional but three-dimensional objects in a scene taken from an arbitrary angle. [22] Only then it can be used to recognize an enemy's tank, to guide an automatic missile towards its target, or to control a robotic arm on the factory floor. The problem with using simple template matching is that "a two-dimensional representation of a two-dimensional object is substantially like the object, but a two-dimensional representation of a three-dimensional object introduces a perspective projection that makes the representation ambiguous with respect to the object." [23] While pattern recognition was working for images of two-dimensional objects, such as letters or chromosomes, a different approach was required to "see" in 3-D.

Roberts' 1965 paper "Machine Recognition of Three-dimensional Solids" is considered to be the first attempt at solving the general task of automatically recognizing three-dimensional objects. [24] His program was designed to understand the artificial world composed solely of polyhedral blocks — a reduction of reality to geometry that would have pleased Cézanne. Using image processing techniques, a photograph of a scene was first converted into a line drawing. Next, the techniques of 3-D computer graphics were used:

Roberts' program had access to three-dimensional models of objects: a cube, a rectangular solid, a wedge, and a hexagonal prism. They were represented by coordinates (x, y, z) of their vertices. A program recognized these objects in a line drawing of the scene. A candidate model was selected on the basis of simple features such as a number of vertices. Then the selected model was rotated, scaled, projected, and matched with the input line drawing. If the match was good, the object was recognized, as were its position and size. Roberts' program could handle even a composite object made of multiple primitive shapes; it subtracted parts of a line drawing from the drawing as they were recognized, and the remaining portions were analyzed further. [25]

Was this enough to completely automate human vision? This depends upon how we define vision. The chapter on computer vision in **The Handbook of Artificial Intelligence** (1982) opens with the following definition: "Vision is the information-processing task of understanding a scene from its projected images." [26] But what does "understanding a scene" mean? With computer vision research financed by the military-industrial complex, the definition of understanding becomes highly pragmatic. In the best tradition of the pragmatism of James and Pierce, cognition is equated with action. The computer can be said to "understand" a scene if it can act on it — move objects, assemble details, destroy targets. Thus, in the field of computer vision "understanding a scene" implies two goals. First, it means the identification of various objects represented in an image. Second, it means reconstruction of three-dimensional space from the image. A robot, for instance, need not only to recognize particular objects, but it has to construct a representation of the surrounding environment to plan its movements. Similarly, a missile not only has to identify a target but also to determine the position of this target in three-dimensional space. It can be seen that Roberts' program simultaneously fulfilled both goals. His program exemplified the approach taken by most computer vision researchers in the following two decades. A typical program first reconstructs the three-dimensional scene from the input image and then matches the reconstructed objects to the models stored in memory. If the match is good, the program can be said to recognize the object, while simultaneously recording its position. A computer vision program thus acts like a blind person who "sees" objects (i.e., identifies them) by reading their shapes through touch. As for a blind person, understanding the world and the recognition of shapes are locked together; they cannot be accomplished independently of one another.

In summary, early computer vision was limited to recognition of two-dimensional forms. Later, researchers began to tackle the task of recognizing 3-D objects which involves reconstruction of shapes from their perspectival representations (a photograph or a video image). From this point on, the subsequent history of computer vision research can be seen as a struggle against perspective inherent to the photographic optics.

The Retreat of Perspective

With the emergence of the field of computer vision, perspectival sight reaches its apotheosis and at the same time begins its retreat. At first computer vision researchers believed that they could invert the perspective and reconstruct the represented scene from a single perspectival image. Eventually, they realized that it is often easier to bypass perspectival images altogether and use other means as a source of three-dimensional information.

Latour points out that with the invention of perspective it became possible to represent absent things and plan our movement through space by working on representations. To quote him again, "one cannot smell or hear or touch Sakhalin island, but you can look at the map and determine at which bearing you will see the land when you send the next fleet." [27] Roberts' program extended these abilities even further. Now the computer could acquire full knowledge of the three-dimensional world from a single perspectival image! And because the program determined the exact position and orientation of objects in a scene, it became possible to see the reconstructed scene from another viewpoint. It also became possible to predict how the scene would look from an arbitrary viewpoint. [28] Finally, it also became possible to guide automatically the movement of a robot through the scene.

Roberts' program worked using only a single photograph — but only because it was presented with a highly artificial scene and because it "knew" what it could expect to see. Roberts limited the world which his program could recognize to simple polyhedral blocks. The shapes of possible blocks were stored in a computer. Others simplified the task even further by painting all objects in a scene the same color.

However, given an arbitrary scene, composed of arbitrary surfaces of arbitrary color and lighted in an arbitrary way, it is very difficult to reconstruct the scene correctly from a single perspectival image. The image is "underdetermined". First, numerous spatial layouts can give rise to the same two-dimensional image. Second, "the appearance of an object is influenced by its surface material, the atmospheric conditions, the angle of the light source, the ambient light, the camera angle and characteristics, and so on," and all of these different factors are collapsed together in the image. [29] Third, perspective, like any other type of projection, does not preserve many geometric properties of a scene. Parallel lines turn into convergent lines; all angles change; equal lines appear unequal. All this makes it very difficult for a computer to determine which lines belong to a single object.

Thus, perspective, which until now served as a model of visual automation, becomes the drawback which needs to be overcome. Perspective, this first step towards the rationalization of sight (Ivins) has eventually become a limit to its total rationalization — the development of computer vision.

The realization of the ambiguities inherent in a perspectival image itself came after years of vision research. In trying to compensate for these ambiguities, laboratories began to scrutinize the formal structure of a perspectival image with a degree of attention unprecedented in the history of perspective. For instance, in 1968 Adolpho Guzman classified the types of junctions that appear in line representations after he realized that a junction type can be used to deduce whether regions of either side of a junction line were part of the same object. [30] In 1986 David Lowe presented a method to calculate the probability that a particular regularity in an image (for instance, parallel lines) reflects the physical layout of the scene rather than being an accident due to a particular viewpoint. [31] All other sources of depth information such as shading, shadows or texture gradients were also systematically studied and described mathematically.

Despite these advances, a single perspectival image remained too ambiguous a source of information for practical computer vision systems. An alternative has been to use more than one image at a time. Computer stereo systems employ two cameras which, like human eyes, are positioned a distance apart. If the common feature can be identified in both images, then the position of an object can be simply determined through geometric calculations. Other systems use a series of continuous images recorded by a video camera.

But why struggle with the ambiguity of perspectival images at all? Instead of inferring three-dimensional structure from a two-dimensional representation, it is possible to measure depth directly by employing various remote sensing technologies. In addition to video cameras, modern vision systems also utilize a whole range of different range finders such as lasers or ultrasound. [32] Range finders are devices which can directly produce a three-dimensional map of an object. The same basic principle employed in radar is used: the time required for an electro-magnetic wave to reach an object and reflect back is proportional to the distance to the object. But while radar reduces an object to a single point and in fact is blind to close-by objects, a range finder operates at small distances. By systematically scanning the surface of an object, it directly produces a "depth map," a record of an object's shape which can be then matched to geometric models stored in computer memory thus bypassing the perspectival image altogether.

Epilogue

The Renaissance's adaptation of perspective represented the first step in the automation of sight. While other cultures used sophisticated methods of space mapping, the importance of perspective lies not in its representational superiority but in its algorithmic character. This algorithmic character enabled the gradual development of visual languages of perspective and descriptive geometry and, in parallel, of perspectival machines and technologies, from a simple net described by Dürer to photography and radar. And when digital computers made possible mass automation in general, automation of perspectival vision and imaging followed soon.

The use of computers allowed to extend perspective, utilizing to the extreme its inherent qualities such as the algorithmic character and the reciprocal relationship it establishes between reality and representation. The perspective algorithm, a foundation of both computer graphics and computer vision, is used to generate perspectival views given a geometric model and to deduce the model given a perspectival view. Yet, while giving rise to new technologies of "geometric vision," perspective also becomes a limit to the final automation of sight — recognition of objects by a computer. Finally, it is displaced from its privileged role, becoming just one among other techniques of space mapping and visualization.

References:

[1] Charles Eames and Ray Eames, **A Computer Perspective: Background to the Computer Age** (Cambridge: Harvard University Press, 1990), 65-67.

[2] Qtd. in ibid., 67.

[3] For a survey of perspectival instruments, see Martin Kemp, **The Science of Art** (New Haven: Yale University Press, 1990), 167-220.

[4] Ibid., 171-172.

[5] Ibid., 200.

[6] See Bruno Latour, "Visualization and Cognition: Thinking with Eyes and Hands," **Knowledge and Society: Studies in the Sociology of Culture Past and Present** 6 (1986): 1-40.

[7] For a comprehensive account of 3-D computer graphics techniques, see J. William Mitchell, **The Reconfigured Eye: Visual Truth in the Post-Photographic Era** (Cambridge, The MIT Press, 1992), 117-162.

[8] Jasia Reichardt, **The Computer in Art** (London and New York: Studio Vista and Van Nostrand Reinhold Company, 1971), 15.

[9] L.G. Roberts, **Machine Perception of Three-Dimensional Solids**, MIT Lincoln Laboratory TR 315, 1963; L.G. Roberts, **Homogeneous Matrix Representations and Manipulation of N-Dimensional Constructs**, MIT Lincoln Laboratory MS 1405, 1965.

[10] "Retrospectives II: The Early Years in Computer Graphics at MIT, Lincoln Lab, and Harvard," in **SIGGRAPH '89 Panel Proceedings** (New York: The Association for Computing Machinery, 1989), 72.

[11] This mixture of automated and pre-industrial labor is characteristic of the early uses of computers for the production of images. In 1955 the psychologist Attneave was the first to construct an image which was to become one of the icons of the age of digital visuality — random squares pattern. A pattern consisted of a grid made from small squares colored black or white. A computer-generated table of random numbers has been used to determine the colors of the square — odd number for one color, even number for another. Using this procedure, two research assistants manually filled in 19,600 squares of the pattern. Paul Vitz and Arnold B. Glimcher, **Modern Art and Modern Science** (New York: Praeger Publishers, 1984), 234. Later, many artists, such as Harold Cohen, used computers to generate line drawings which they then colored by hand, transferred to canvas to serve as a foundation for painting, etc.

[12] For further discussion of the problem of realism in 3-D computer graphics, see Lev Manovich, "Real" Wars: Esthetics and Professionalism in Computer Animation," **Design Issues** 6, no. 1 (Fall 1991): 18-25; Lev Manovich, "Assembling Reality: Myths of Computer Graphics," **Afterimage** 20, no. 2 (September 1992): 12-14.

[13] Mitchell, **The Reconfigured Eye**, 118.

[14] "Retrospectives II: The Early Years in Computer Graphics at MIT, Lincoln Lab, and Harvard," 57.

[15] Remko Scha, "Virtual Voices," **MediaMatic** 7, no. 1 (1992): 33. Scha describes two fundamental approaches taken by the developers of voice imitating machines: the genetic method which imitates the physiological processes that generate speech sounds in the human body and the gennematic method which is based on the analysis and reconstruction of speech sounds themselves without considering the way in which the human body produces them. While the field of computer vision, and other fields of artificial intelligence, first clearly followed gennematic method, in the 1980s, with the growing popularity of neural networks, there was a shift towards the genetic method — direct imitation of the physiology of the visual system. In a number of laboratories, scientists begin to build artificial eyes which move, focus, and analyze information exactly like human eyes.

[16] Eames and Eames, **A Computer Perspective**, 100.

[17] Manuel De Landa, "Policing the Spectrum," in **War in the Age of Intelligent Machines** (New York: Zone Books, 1991), 194-203.

[18] Ibid., 214.

[19] The first paper on image processing was published in 1955. L.S.G. Kovasznay, and H.M. Joseph, "Image Processing," **Proceedings of IRE** 43 (1955): 560-570.

[20] Rafael C. Gonzalez, and Paul Wintz, **Digital Image Processing** (Reading, Mass.: Addison-Wesley Publishing Company, 1977), 2.

[21] Qtd. in De Landa, "Policing the Spectrum," 200.

[22] Within the field of computer vision, a scene is defined as a collection of three-dimensional objects depicted in an input picture. David McArthur, "Computer Vision and Perceptual Psychology," **Psychological Bulletin** 92, no. 2 (1982): 284.

[23] Paul R. Cohen and Edward A. Feigenbaum, eds., **The Handbook of Artificial Intelligence** (Los Altos, CA: William Kaufmann, Inc., 1982), 3: 139.

[24] L.G. Roberts, "Machine perception of three-dimensional solids," **Optical and Electro Optical Information Processing**, ed. J.T. Tippett (Cambridge: The MIT Press, 1965).

[25] Cohen and Feigenbaum, **The Handbook of Artificial Intelligence**, 3: 129.

[26] Ibid., 127.

[27] Latour, "Visualisation and Cognition," 8.

[28] Cohen and Feigenbaum, The Handbook of Artificial Intelligence, 3: 141.

[29] Ibid., 128.

[30] Ibid., 131.

[31] David Lowe, **Three-Dimensional Object Recognition from Single Two-Dimensional Images**, Robotics Report 62 (New York: Courant Institute of Mathematical Sciences, New York University, 1986).

[32] Cohen and Feigenbaum, **The Handbook of Artificial Intelligence**, 3: 254-259.