Reviewed in LEONARDO, Journal of the International Society for the Arts, Sciences and Technology.
Vol 17, Number 4, 1984

The Interpretation of Visual Motion. Shimon Ullman. MIT Press, Cambridge, Mass. 1979. 229 pp., illus. Cloth. ISBN: 0-262-21007-X.

It's a bird! It's a plane! It's Superman! Unambiguously and provably so, according to one of Shimon Ullman's theories concerning the perception of objects in motion. This theorem, named the "structure from motion theorem", is worth quoting. It epitomizes the style and method of the book as a whole. "Given three distinct orthographic views of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined" (p. 148). In other words, a human observer, or, more to the point, a computerized robot's eye, given three separate views or frames of a movie sequence of the flying Superman, can be absolutely sure that the figure is of a certain shape and flying in a certain direction, even if Superman is flying at night with just four lights attached to his body giving the visual cues to the idealized observer. For most of us who are blessed with the gift of sight, all this seems superfluous and rather tedious, since we already instinctively know it to be true.

Ullman, however, has the well-defined aim of studying the *tasks* accomplished by a visual system in order to investigate the computations it performs. It almost seems that the author was searching for an ideal visual system and chose the human visual system because it is efficient, but not quite so efficient as certain mathematical algorithms (p. 155). This approach is fully justified, since the research is the product of MIT's Artificial Intelligence Laboratory. The thrust of the work is to analyze those aspects of human vision of motion that can be useful in teaching a machine how to see. This aim is not clearly stated, but the discussion is colored by the computational approach to vision research. The results of the meticulous experiments are expressed mathematically and used to refute older, more intuitive results by Gibson (p. 140) and others. One cannot help but think there are many other aspects of the subject than those treated in this book. For example, what effect does body acceleration have on our experience of motion? When a 3-dimensional object is seen in motion, is it being seen with one eye or with two? What can we learn from such effects as the waterfall illusion or the Pulfrich pendulum illusion about the way we perceive motion? More basically, what is the physiological and neurological basis of the perception of motion? All these matters are untreated in this very specialized and poorly illustrated book. Is it fair for a layman to criticize a work in this way? I believe so: scientific truth is relative. To Ullman a baby looks at a milk bottle from different viewpoints, "attaching them to the new superordinate 'node'". But does a baby have to learn all the formulas in the Appendix entitled "Structure from Perspective Projections", as a robot must? There is a constant temptation for people working on artificial intelligence to carry the notion, "this is the best way to design a robot vision", all the way to "human vision is almost like a robot's vision".

Most of the experiments discussed involve two pictures of dots or lines that were presented rapidly one after another to an observer, who saw apparent motion if there was displacement of the stimuli. Ullman shows that very few points or visual cues are sufficient to define an object's motion in two- or three-dimensional space. The correspondence between the objects in the two frames is analyzed using a "minimal mapping theory"—essentially a mathematical approach to matching the two views topologically. The last two chapters of the book are entitled "The Interpretation of Structure from Motion" and "The Perception of Motion from Structure". These relate to the perception of the shape of an unknown object known from the moving two-dimensional views the observer sees, and the perception of the motion of a known object, such as a cube.

Although quite abst act, *The Interpretation of Visual Motion* is not abstract enough, in the sense that it only deals with the "real" visual world. There is no mention of the contortions of the pseudoscopically viewed scene. The mathematical treatment is limited to generally rigid objects moving in our familiar perceptual world following the rules of perspective. But what about the perception of motion in artificial reality? In computer graphics, objects not only can be transformed magically from a cube to an angel to a field of stars, but the very space in which these objects move can be given new properties—for example, a spherical coordinate space where the perceived motion increases the more the object approaches the center. All this makes one long for a nonscientific, quiet vision, the world of a Millet, or at least a child running after an enticing moving target, a well-thrown ball.

Reviewed by **Vladimir Tamari**, 1-27-11 Shimmachi, Setagaya-ku, Tokyo 154, Japan.