BOOK REVIEW

Real-time Computer Vision

Christopher M. Brown and Demetri Terzopoulos, Eds., 232 + xvi pages. ISBN: 0-521-47278-4. Cambridge University Press (1995) $49.95.

Reviewed by K. Venkatesh Prasad
Ricoh California Research Center
2882, Sand Hill Road, Suite 115
Menlo Park, California 94025-7022

Most of us human beings with unimpaired vision have been endowed with such sophisticated visual skills that vision is, and always has been, an effortless, "real-time" process. So effortless is the visual process for us that we tend to take it and the machinery behind it—the eyes and the brain—for granted, until we begin to investigate either the machine implementability, biology, or the psychology of the process. Those of us attempting to endow computers with vision will readily admit how humbling the process can be and why real-time vision is really a revolutionary advancement in the field of computer vision.

The book *Real-time Computer Vision* (RTCV), edited by Brown and Terzopoulos, is an excellent and timely testament to the advancements in the field. The chapters in the book are based on selected lectures presented in 1993 during a computer vision program sponsored by the Isaac Newton Institute of Mathematical Sciences, Cambridge University. With a judicious choice of substance and style, the editors have assembled a book that is neutral with respect to the raging ideological debates in vision theory—reconstructionist versus purposive versus behaviorist vision. All these differing schools of thought have inspired the creation of algorithms that work in real-time, and the book contains a chapter based on each ideology. Algorithms are the focus of the book with only occasional reference to ideology or even the underlying enabling technologies. Advanced students and researchers, particularly those interested in visually guided motion, will find this book valuable. The nine chapters in the book span a broad range of topics that include shape recovery, model-based vehicle tracking, active exploration, tracking heads and eyes, controlling robot behavior, visual monitoring, and controlling distributed robots.

The book begins with an introduction to some of the terminology and concepts central to RTCV. With regard to terminology, one preliminary question is "what is meant by RTCV?" The quantitative definitions offered are frequently arbitrary. Depending on where one is in the world, one could hear real-time defined by a different frequency measure (in Hz): 30 Hz in the 110-V/60-Hz electric power world and 25 Hz in the 220-V/50-Hz electric power world, implying that results of a visual process can be obtained every 30th (or 25th) of a second. A reasonable working definition, provided by the editors, is that RTCV "refers to the computer analysis of dynamic scenes at rates sufficiently high to effect practical visually-guided control or decision-making in everyday situations. This means relevant visual information must be extracted several times a second, say from 10 to 30 Hz." To prepare the reader for the book's content, the editors thoughtfully enumerate and briefly review concepts of particular importance to the foundation of RTCV: the principles of estimation (based on the Kalman filter), affine geometry, and dynamic modeling and control. The editors urge the mastering of these and other traditional topics in computer vision such as clustering, feature finding, and recovery of three-dimensional scene parameters.

Eight chapters follow the introductory chapter and are grouped into three parts: Visual Tracking, Model-Based Vision and Exploration, and Visual Control. The three chapters in the visual tracking part deal with automatic tracking algorithms that are not critically domain or application dependent. For instance, the first chapter deals with the tracking of general curves in real-time, and the algorithm developed is shown to be applicable to curves originating from human hands, circuit boards, and flowers. The value of a statistical basis for the automatic control of scales of operation is established in this chapter; the Kalman filter is used to form a principled procedure to estimate the spatial and temporal scales of operation. The three chapters in the model-based vision and exploration part contain algorithms that are somewhat application dependent but deal with visual tasks and scenes that are more complex than those in the first part. The visual tasks include locating, recognizing, and tracking multiple vehicles in cluttered traffic scenes; or detecting motion in a scene followed by a redirecting of a foveal region of attention to the motion region; or estimating a complete geometric and topologic model of a surface from a sequence of views taken together. The final part on visual control contains three chapters related to visually guided motor control. The first of these chapters addresses the hand-eye coordination problem and presents a technique for the improvement of robotic manipulation based on visual feedback. The second chapter in this part deals with the cooperative behavior between robotic agents using visual information. The final chapter presents an experimental proof for a conjecture that complex tasks in a known environment can be solved by a "division of labor" approach, i.e., by "using a hybrid system comprising symbolic and nonsymbolic components ...".

From an editorial standpoint, the book appears carefully assembled. One might run into an occasional typographical error (for example, the words "with outwith" appear on page 209 where one might expect just the word "without"), but the flow of the substance is coherent and unambiguous. A detailed subject index and a comprehensive bibliography might have been useful additions. The advanced reader is likely to find the book self contained and richly rewarding, while the beginning reader with an undergraduate computer science, electrical engineering, mathematical, or physical science background might find the book to be a source of inspiration. The book is well suited to a second-level graduate course in computer vision. Engineers involved with exploratory research and development involving real-time and embedded systems or hardware implementation of algorithms might find plenty of food for thought in this book. In any case, this book would make a good addition to any technical library.

The contents of the book, in addition to a preface and an introduction are as follows: Part I: Visual Tracking—A framework for spatio-temporal control in the tracking of visual contours; Tracking moving heads; Tracking and measuring drivers eyes; Part II: Model-Based Vision and Exploration—Model-based vision for traffic scenes using the ground-plane constraint; Active exploration of dynamic and static scenes; Robust shape recovery from occluding contours using a linear smoother; Part III: Visual Control—Visual robot guidance from uncalibrated stereo; Control of visually guided behaviors; Hybrid problems need hybrid solutions—tracking and controlling toy cars.

K. Venkatesh Prasad received his BE degree in electronics and communication engineering from the University of Madras in 1980. He obtained an MSEE degree from the Indian Institute of Technology, Madras, in 1984, based on research in the field of electromagnetics and antennas. In 1987, he obtained an MS degree with an emphasis in computer engineering from Washington State University, Pullman. In 1990, he received a PhD degree in electrical and computer engineering from Rutgers University, New Brunswick, New Jersey. From 1990 to 1992, he was a postdoctoral associate at Rutgers University, a visiting scientist at the California Institute of Technology (Caltech), Pasadena, and a Caltech affiliate at NASA's Jet Propulsion Laboratory, Pasadena. Since December 1992 he has been a senior research scientist at Ricoh Corporation's California Research Center working in the area of visual information processing applied to visual speech, digital documents, and video processing.