BOOK REVIEW

Computer and Robot Vision

Robert M. Haralick and Linda G. Shapiro. Volume 1, 672 + xvi pages, ISBN 0-201-10877-1, $70.95. Volume 2, 630 + x pages, ISBN 0-201-56943-4, $66.75. Addison Wesley, Reading, Massachusetts (1993).

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

For nearly two decades—the sixties and seventies—a major proportion of problems addressed in the field of computer vision dealt with clean images in toy worlds. A well-kept "secret" was computation time. The eighties saw steady progress toward solving real-world vision problems, beginning with noisy images and computing, in real time, the three-dimensional shape and range of real-world objects. Much of this progress, however, was still reported from the privileged echelons of academia or corporate research laboratories. In the present decade we seem well set to launch into an era of real-world real-time computer vision with everyday life applications.

It is with this backdrop that Haralick and Shapiro, both distinguished researchers in the field of computer vision, have coauthored this two-volume book. The book is written with an emphasis on practical issues and addresses those aspects of computer vision that are required in fields such as robotics, automated inspection, cartography, face recognition, medical image diagnostics, and in-orbit spacecraft maintenance.

The field of computer vision is rather diverse, with a growing gamut of applications and algorithmic solutions on the one hand and a constantly evolving set of theoretical underpinnings and high-speed computational strategies on the other hand. The applications of computer vision are rapidly increasing, in part, because of the cost-effective availability of enabling technologies: cameras, frame grabbers, computer memory, and display monitors, to name a few; and, in part, because the field has reached a level of maturity at which algorithms that work on "real images" are now readily available. Some applications, which a decade ago might not have been considered practical, have now become a reality, some still in advanced laboratories, but many others in just everyday life applications. These include computer vision "driven" highway vehicles, real-time photo interpretation, postal packet inspection, telerobotic in-orbit spacecraft inspection, and medical image diagnostics. The theoretical foundations of computer vision are in constant evolution because of the feedback from biological vision-related discoveries in the fields of neurobiology and psychophysics. When the applications demand a computer vision-based solution and suitable algorithms are available to handle the task, what eventually makes computer vision "work" is that it empowers machines to capture an image of a scene and compute a symbolic description of the objects in the scene in real time. The key words here are "compute" and "real time". The rapid strides in parallel and distributed computation have made possible the existence of computer vision algorithms that work in real time. Therefore, one might consider the field of computer vision to be driven by progress in three complex intercoupled domains: vision applications, theoretical paradigms of vision, and high-speed computation.

The multifaceted nature of the field makes writing a textbook quite a challenge. Since the central theme of the book is on the practical issues of computer vision, Haralick and Shapiro have judiciously chosen to make only a passing mention, or no mention at all, of topics dealing with the theoretical or high-speed computational underpinnings of computer vision. Topics not covered in either of the two volumes include optical, biological, and psychophysical aspects of vision, such as the optics of image formation, the neurobiology of vision, visual perception, and parallel architectures in vision.

The contents of the two volumes are ordered by the three-level theme of vision, popularized by the late David Marr. The first volume deals with issues of low-level vision (based on pixel-level analysis) and intermediate-level vision (based on edge or region analysis). The second volume describes topics in high-level vision (knowledge-based scene analysis).

Each chapter contains an exhaustive bibliography that places the contents of the chapter in the proper context of related activity. There are sufficient illustrations and figures with gray-scale images to make the assimilation of the material efficient, especially for readers with limited prior exposure to the field. The print quality of some of the figures, particularly the images of histograms or other plots, appears slightly underexposed or lacking in contrast, but this does not hinder the overall interpretation of the concepts. There are ample examples and homework exercises contained in each chapter, making them well suited for both self-study and formal classroom study. The book assumes a maturity in mathematics at the advanced undergraduate or beginning graduate student level. For those battling the jargon-driven world of today's high technology, the authors have very thoughtfully provided an exhaustive glossary of computer vision terms and a rather complete index of terms. Overall, the volumes appear to be well proofread and edited.

There are a couple of desirable attributes of a two-volume book that are conspicuous by their absence from this book: one is a comprehensive bibliography, or at least an index of cited authors, at the end of each volume, and the second is a complete two-volume table of contents in both books.

This two-volume book is suitable for use as a one-year textbook. If you are looking for a comprehensive, cohesive, and self-contained collection of material on the broad and growing topic of computer vision, then these two volumes might be your one-stop choice.

To allay any doubts you may have, a listing of chapter titles follows. The first volume includes Computer Vision: Overview; Binary Machine Vision: Thresholding and Segmentation; Binary Machine Vision: Region Analysis; Statistical Pattern Recognition; Mathematical Morphology; Neighborhood Operators; Conditioning and Labeling; The Facet Model; Image Segmentation; and Arc Extraction and Segmentation. In addition there are three appendices to this volume. The first appendix contains a review of the use of an ellipse as a single-shape approximation to a simple, compact two-dimensional region. The second appendix provides a review of the basics of linear algebra, orthogonal projections, and least-squares approximations. The third appendix is quite unusual in that it discusses an issue rarely broached in computer vision textbooks: What protocol should be adhered to in carrying out experiments? The authors write: "The experimental protocol states the quantity to be measured, the accuracy to which it is to be measured, the population of scenes/images on which the algorithm is to be applied. Then the protocol must give the experimental design and the data analysis plan."

The chapter titles of the second volume include Illumination; Perspective Projective Geometry; Analytic Photogrammetry; Motion and Surface Structure from Time Varying Image Sequences; Image Matching; The Consistent Labeling Problem; Object Models and Matching; Knowledge-Based Vision; Accuracy; and a glossary of computer vision terms.

Haralick and Shapiro's two-volume book provides more than adequate material for a full-year university course. The book is self-contained and, pedagogically speaking, well structured with ample examples and exercises, both of the pencil and paper variety and of the longer kind involving computer programming. Professionals in the field of robotics, industrial manufacturing, medicine, and the space sciences, desirous of getting acquainted or updated with the developments in the field, will find the book a rich resource. As we launch into the era of real-world computer vision these books will be a valuable set to hold on to.

K. Venkatesh Prasad received his BE degree in electronics and communication engineering from the University of Madras, India, 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 research scientist at Ricoh Corporation's California Research Center working in the area of visual information processing applied to visual speech and document analysis.