Visual reconstruction for researchers

Visual reconstruction by Andrew Blake and Andrew Zisserman. Published by MIT Press, USA. 1987. 225 pp. £22.50.

Computer vision poses a formidable information processing problem. Conceptually, visual information processing can be divided into three stages – low, intermediate and high. The low and intermediate levels are critical in the visual processing hierarchy; it is here that information about the structure of images and surfaces is extracted and refined into a coherent representation which can drive subsequent high-level object analysis tasks.

Blake and Zisserman's excellent monograph concerns itself with a class of inverse or reconstruction problem central to low and intermediate level visual processing. The authors' theme is visual reconstruction which they characterize as reducing visual data to stable descriptions. Such descriptions must make explicit both the continuous regions and the disconti-nuity boundaries implicit in raw, noisy data. This amounts to a piecewise continuous reconstruction of the data. This book presents a detailed investigation into this problem using the powerful class of weak continuity constraints.

The book is remarkably well written. Using readily accessible explanations and vivid examples, the authors present novel research including theoretical and empirical results gathered over the past five years. Formal analysis is postponed till well into the book and much of the supporting mathematics is relegated to the appendices. The result is a nicely balanced and self-contained presentation which is at the same time concise and clear.

The text begins by introducing the notion of piecewise continuous modelling and makes the important point that cooperative processing is the most attractive way to compute these models. Next, a selection of examples in image discontinuity detection, surface reconstruction and curve description provides a perspective on the broad applicability of the algorithms developed later in the book. The authors review prior reconstruction work, mentioning pioneering research at MIT, USA and other institutions, and contrasting their approach with more traditional filtering tech-niques. Piecewise continuous reconstruction is an evolving area of investigation which continues to attract the attention of computer vision researchers internationally.

The second part of the book examines two instances of weak continuity constraints in one and two dimensions. These constraints take the form of energy functionals that penalize the lack of continuity or smoothness in the reconstructed solution. The basis for the constraints is the popular first-order (string and membrane) and second-order (rod and plate) strain energies and their viewpoint invariant extensions. The authors' contribution is to augment the basic energies with an additional term that permits discontinuities at a cost. The resulting weak continuity constraints formalize reconstruction as an optimization problem. Because discontinuities incur a fixed penalty, the energy function is highly non-convex, exhibiting many local suboptima. The authors present a detailed mathematical analysis of weak continuity models, including a cogent elucidation of the roles played by the model's parameters in determining the scale, sensitivity and noise resistance of the reconstruction process.

The third part of the book investigates the discrete form of the reconstruction problem and its numerical solution. As is by now standard practice for variational problems in vision, discretization is achieved using finite-element techniques. The authors propose a graduated non-convexity (GNC) descent algorithm which successfully avoids local suboptima. At the time of writing, the authors apparently had not realized that GNC is a discrete version of the continuation method, an established numerical technique for solving difficult nonlinear problems. The continuation method prescribes a homotopy (embedding) of a nonconvex optimization problem into a parameterized family of problems. The original nonconvex problem lies at one extreme of the continuation parameter's range while a (more) convex approximation to this problem lies at the other extreme. The continuation method first solves the simple convex problem, then 'bootstraps' itself towards a good solution of the difficult, nonconvex problem by solving a sequence of increasingly nonconvex approximations using the previous solution(s) as an initial guess for the next optimization.

Other researchers have successfully employed continuation methods to solve severely non-convex problems such as stereo matching. Their results corroborate Blake and Zisserman's observation that this deterministic approach is a powerful and efficient alternative to stochastic optimization methods such as simulated annealing. The GNC algorithm is apparently effective for the problem at hand but it has certain limitations. It adjusts the continuation parameter in a heuristic sequence of discrete steps rather than continuously as do 'true' continuation methods. The authors' analysis of the conditions under which GNC produces good solutions is not entirely satisfactory for most practical situations that do not involve dense data. To the book's credit, it presents many piecewise-continuous surface reconstruction examples including a variety of intensity and depth data sets, both natural and synthetic; however, most of the examples involve dense data. Whether the GNC technique can be effective with sparse data, such as that provided by feature-based stereo matching, remains to be demonstrated convincingly. The difficulties that the weak membrane evidently encounters with sparse

data may be attributable in part to the absence of a line process component (present in competing piecewise reconstruction techniques) which favours figural continuity in the surface discontinuities.

The book concludes by offering some interesting opinions about the nature and role of early visual processing. We find ourselves in agreement with most but not all of these opinions. One of their puzzling positions (first put forth in chapter 1) is that explicit surface reconstruction is of restricted value: 'the main purpose of producing a depth map . . . is to make discontinuities in the visible surface'. Discontinuities are indeed crucial but our own work in this area indicates that it is difficult to describe discontinuities meaningfully without simultaneously describing the continuous surfaces that they bound. Among the authors' arguments against explicit surface reconstruction is the ambiguity of the reconstruction in the absence of shading or texture information. Fortunately, however, it is possible to model this shape ambiguity and to use it to guide subsequent processing. A more compelling argument in favour of explicit surface reconstruction is that it facilitates the integration of multiple sources of partial and noisy visual information, an issue which is presently attracting much attention. Furthermore, explicit surface reconstruction appears consistent with the psychophysics of surface perception, is necessary for surface shape analysis and is clearly useful for geometric modelling and other engineering applications.

Another of Blake and Zisserman's opinions is that mechanical modelling is preferable to probabilistic modelling, in part because it provides continuous models. Although the mechanical approach is very powerful and perhaps more 'natural' for specifying surface model parameters, the probabilistic approach can also accommodate continuous models because of the equivalence between the Gibbs distribution and Markov random fields. Moreover, the probabilistic approach is superior for modelling the degradation (noise) introduced by vision sensors, for evaluating and reducing the uncertainty in the estimates and for estimating the values of hidden parameters. The issue of whether to use deterministic or stochastic approaches to compute solutions is therefore relatively independent of model formulation and more relevant to the search for efficient algorithms. For applications such as visual reconstruction, specialized deterministic algorithms such as GNC apparently are more efficient than the more generally applicable simulated annealing technique. The development of probabilistic and deterministic models remains a fertile area of research. This book champions one side of an exciting interplay between the two approaches which is yielding interesting new results.

This book is a clear and timely exposition of an effective approach to piecewise continuous reconstruction. The book is essential reading for researchers in low and intermediate level computer vision, offering ample food for thought and inspiration for future work. For readers outside the vision community, it provides an elegant exemplar of the computational approach to vision.

Richard Szeliski and
Demetri Terzopoulos,
Schlumberger Palo Alto Research,
Palo Alto,
California, USA

Robotic object recognition

Robotic object recognition using vision and touch by Peter K Allen. Published by Kluwer, Boston, USA. 1987. 172pp. $40.

This book describes a research project being undertaken at the University of Pennsylvania, USA which has the aim of recognizing kitchen utensils. These items present quite a challenge for a robot recognition system because of their range of shapes and degree of complexity. If this project succeeds it will help to hasten the advent of the household robot. Object recognition was chosen as a starting point because it is an essential first step in many grasping, manipulation, inspection and assembly operations.

A combination of passive stereo vision and active exploratory touch is used to gather information about the objects. The author claims that neither stereo vision nor touch by itself can perform the recognition task. However, by combining the two sensing modes, the resulting system can perform better than the two sensing modes operating in isolation.

This book provides more of a progress report than the description of a technique which can be immediately applied. This is not however a criticism of the book. Object recognition tasks seem deceptively simple to humans but the technological problems involved in giving a robot the same capabilities are immense.

This book will be of interest to graduate students studying robot sensing – computer vision and/or touch sensing. People in industry or universities engaged in research into intelligent robotics will also find the book useful. It will be a valuable addition to my bookshelf.

R Andrew Russell
Monash University
Victoria, Australia

Machine vision for end users

Applying machine vision by Nello Zuech. Published by Wiley-Interscience, NY, USA. 1988. 265 pp. £30.45.

In the preface to this book the author states that: