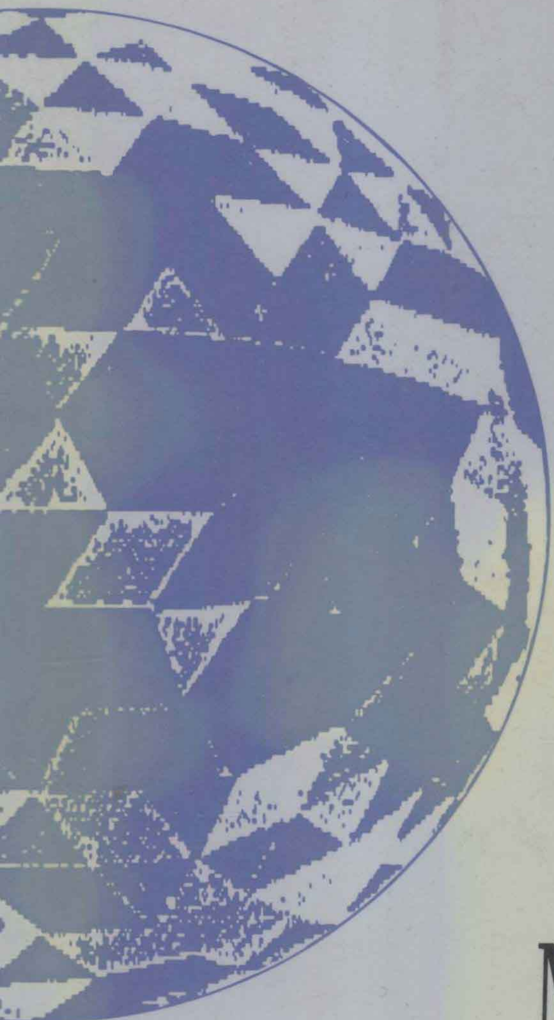
David Vernon



MACHINE VISION

Automated Visual Inspection and Robot Vision

Machine Vision

Automated Visual Inspection and Robot Vision

David Vernon
Department of Computer Science
Trinity College Dublin Ireland

Prentice Hall
New York London Toronto Sydney Tokyo Singapore

First published 1991 by
Prentice Hall International (UK) Ltd
66 Wood Lane End, Hemel Hempstead
Hertfordshire HP2 4RG
A division of
Simon & Schuster International Group

Typeset in 10 on 12 point Times
by MCS Typesetters, Salisbury, Wiltshire, England

Printed and bound in Great Britain
by Cambridge University Press

Library of Congress Cataloging-in-Publication Data
is available from the publisher

British Library Cataloguing in Publication Data

Vernon, David
Machine vision.
I. Title
006.3

ISBN 0-13-543398-3

1 2 3 4 5 95 94 93 92 91

Machine Vision

Everything that I can spy
Through the circle of my eye,
Everything that I can see
Has been woven out of me;
I have sown the stars, and threw
Clouds of morning and of eve
Up into the vacant blue;
Everything that I perceive,
Sun and sea and mountain high,
All are moulded by my eye:
Closing it, what shall I find?
– Darkness and a little wind.

James Stephens
The Hill of Vision

Preface

Machine vision is a multi-disciplinary subject, utilizing techniques drawn from optics, electronics, mechanical engineering, computer science, and artificial intelligence. This book is intended to be an in-depth introduction to Machine Vision which will allow the reader quickly to assimilate and comprehend the essentials of this evolving and fascinating topic. Significant emphasis will be placed on providing the reader with a solid grounding in the fundamental tools for image acquisition, processing, and analysis; a range of techniques, dealing with very simple two-dimensional systems, through more sophisticated robust two-dimensional approaches, to the current state of the art in three-dimensional robot vision, will be explained in some detail. Both application areas of automated visual inspection and robot vision are addressed. Recognizing that machine vision is just a component of a larger automation system, a brief introduction to robot programming will be provided, together with an explanation of the mechanisms by which robot vision modules interact with the programming language. It is important to recognize that the discipline of machine vision is presently undergoing a maturing process, with sophisticated techniques drawn from current research being exploited more and more in industrial systems. Without doubt, there is a long way to go, but the die is well cast. Acknowledging this trend, the last chapter of the book is devoted to the more research-orientated topics of three-dimensional image understanding and early visual processing (e.g. stereopsis and visual motion). It would indeed be foolhardy to attempt an exhaustive treatment of these areas; each deserves a volume on its own. However, if the essence of the philosophy of robot vision in its broadest sense is cogently imparted to the reader, then the exercise will have been successful and worth while.

The book is directed at final-year undergraduate and first-year graduate students in computer science and engineering, and at practising industrial engineers; the fundamental philosophy being to impart sufficient knowledge so that the reader will be competent to begin the implementation of a simple vision system and to enable him/her to study each issue independently in more depth. To that end, care is taken to provide adequate references to supporting texts, reports, and research papers. In this way the book may be viewed both as a self-contained introductory text and as a spring-board to more detailed and specific study.

Acknowledgements

Special thanks are due to Kenneth Dawson of the Computer Vision Group, Department of Computer Sciences, Trinity College, Dublin, for his work on the raw primal sketch, the extended Gaussian image, and the polyhedral models; and to Massimo Tistarelli and Prof. Giulio Sandini at the University of Genoa for their help with the examples of camera motion and stereopsis. Many people in the Trinity Computer Vision Group read draft versions of this book and I am grateful for their contributions. I would also like to record a note of thanks to Dr R. Dixon, The University, Manchester, for his many valuable comments on an earlier draft of this book. Several of the examples in this book were facilitated by research funded by the Commission of the European Communities under the European Strategic Programme for Research and Development in Information Technology: Project 419 – Image and Movement Understanding.

Contents

1

An introduction to computer vision

1.1 Computer vision: image processing or artificial intelligence?

What is computer vision and why would one be interested in studying it? It is perhaps easier to answer these two questions in reverse order. There are several reasons why one would be interested in computer vision, but the following two will serve to illustrate the many directions from which one can view the subject area:

1. All naturally occurring intelligent life-forms exhibit an ability to interact with and manipulate their environment in a coherent and stable manner. This interaction is facilitated by on-going intelligent interplay between perception and motion-control (i.e. action); visual perception is fundamentally important to most intelligent life.
2. Most manufacturers are concerned with the cosmetic integrity of their products; customers quite often equate quality of appearance with functional quality. So, to ensure the successful long-term marketing of an item, it is highly desirable that its appearance is checked visually before packaging and shipping. Likewise, it is desirable that the inspection process be automated and effected without human intervention.

These two motivations for the study of perception characterize two possible extremes of interest in the processing, analysis, and interpretation of visual imagery: from the philosophical and perhaps esoteric to the immediate and pragmatic. And the subject matter of everything between these two extremes presents one with wide and varied spectrums of commercial interest, difficulty and, indeed, success.

The answer to the first question (what is computer vision?) now becomes a little easier to identify. The world we live in and experience is filled with an endless variety of objects, animate and inanimate, and, to borrow a phrase from David Marr (of whom we shall hear more later in the book), it is by looking and seeing that we come to know what is where in this world. So, if vision is a means to an end – to know the world by looking – then computer vision is exactly the same except that the medium by which the knowledge is gained is now a computational instrument rather than the brain of some living creature. Without doubt, this is a very broad definition. But the subject matter of computer vision *is* this broad: topics such as image restoration, image enhancement, automated visual inspection, robot vision, computer-based image understanding of general three-dimensional scenes, and visual perception and cognition all fall under the umbrella of the term 'computer vision'.

Although for centuries man has been interested in solving the puzzle of how man comes to 'see', the first computational experiments in developing artificial machine vision systems were conducted in the late 1950s and, over the last twenty-five to thirty years computer-based vision systems of widely varying degrees of complexity have been used in many diverse areas such as office automation, medicine, remote sensing by satellite, and in both the industrial world and the military world. The applications have been many and varied, encompassing character recognition, blood cell analysis, automatic screening of chest X-rays, registration of nuclear medicine lung scans, computer-aided tomography (CAT), chromosome classification, land-use identification, traffic monitoring, automatic generation of cartographic projections, parts inspection for quality assurance industrial, part identification, automatic guidance of seam welders, and visual feedback for automatic assembly and repair. Military applications have included the tracking of moving objects, automatic navigation based on passive sensing, and target acquisition and range-finding.

As we have seen, computer vision is concerned with the physical structure of a three-dimensional world by the automatic analysis of images of that world. However, it is necessary to qualify the use of the word *image*. First, the image is a two-dimensional one and, hence, we inevitably lose information in the projection process, i.e. in passing from a three-dimensional world to a two-dimensional image. Quite often, it is the recovery of this lost information which forms the central problem in computer vision. Second, the images are digital images: they are discrete representations (i.e. they have distinct values at regularly sampled points) and they are quantized representations (i.e. each value is an integer value).

Computer vision includes many techniques which are useful in their own right, e.g. image processing (which is concerned with the transformation, encoding, and transmission of images) and pattern recognition (frequently the application of statistical decision theory to general patterns, of which visual patterns are but one instance). More significantly, however, computer vision includes techniques for the useful description of shape and of volume, for geometric modelling, and for so-called cognitive processing. Thus, though computer vision is certainly concerned with the processing of images, these images are only the raw material of a much broader science which, ultimately, endeavours to emulate the perceptual capabilities of man and, perhaps, to shed some light upon the manner by which he accomplishes his amazingly adaptive and robust interaction with his environment.

1.2 *Industrial machine vision vs. image understanding*

Computer vision, then, is an extremely broad discipline (or set of disciplines) and in order to get to grips with it, we need to identify some way of classifying different approaches. To begin with, we note that humans live and work within a general three-dimensional world, pursuing many goals and objectives in an unconstrained and constantly changing environment in which there are many varied and, often, ill-defined objects. Industrial automation, on the other hand, is given to performing single repeated tasks involving relatively few objectives, all of which are known and defined, in manufacturing environments which are normally constrained and engineered to simplify those tasks. Industrial systems do not yet work with *general* three-dimensional environments (although the environments they do work in are often much less structured than one would suppose) and vision systems for manufacturing still exploit many assumptions, which would not generally apply to unconstrained worlds with many objects and many goals, in order to facilitate processing and analysis. There is a considerable dichotomy between the two approaches – a situation which must change and is changing – it is for this reason that the final chapter is concerned with advanced techniques and their migration to the industrial environment. Let us look a little closer at each of these classes of computer vision.

Approaches associated with general environments are frequently referred to by the terms 'image understanding' or 'scene analysis'. The latter term is now quite dated as it typically refers to approaches and systems developed during the 1970s. Vision systems specifically intended for the industrial environment are often referred to generically as 'industrial machine vision systems'.

Image understanding vision systems are normally concerned with three-dimensional scenes, which are partially constrained, but viewed from one (and often several) unconstrained viewpoint. The illumination conditions may be known, e.g. the position of the room light might be assumed, but usually one will have to contend with shadows and occlusion, i.e. partially hidden objects. As such, the data or scene representation is truly a two-dimensional image representation of a three-dimensional scene, with high spatial resolutions (i.e. it is extremely detailed) and high grey-scale resolutions (i.e. it exhibits a large variation in grey-tone). Occasionally, colour information is incorporated but not nearly as often as it should be. Range data is sometimes explicitly available from active range-sensing devices, but a central theme of image understanding is the automatic extraction of both range data and local orientation information from several two-dimensional images using e.g., stereopsis, motion, shading, occlusion, texture gradients, or focusing. One of the significant aspects of image understanding is that it utilizes several redundant information representations (e.g. based on the object edges or boundaries, the disparity between objects in two stereo images, and the shading of the object's surface); and it also incorporates different levels of representation in order to organize the information being made explicit in the representation in an increasingly powerful and meaningful manner. For example, an image understanding system would endeavour to model the scene with some form of parameterized three-dimensional object models built from several low-level processes based on distinct visual cues. At present, image-understanding systems utilize both explicit knowledge (or models) and software-embedded knowledge for reasoning, that is, for controlling image analysis.

Most industrial machine vision systems contrast sharply with the above approach. The scenes in an industrial environment are usually assumed to be two-dimensional, comprising known isolated rigid parts, frequently with a contrasting visual backdrop. Lighting is almost always a critical factor and must be very carefully organized. Typically, the ambient room lighting will be totally inadequate, and even confusing, so that each inspection station will require its own set of dedicated lights, each designed for the task in hand. The images which industrial machine vision systems use are frequently two-dimensional binary images (pure black and white, with no intermediate grey-levels) of essentially two-dimensional scenes. There is normally just one simple internal object representation or model; the analysis strategy being to extract salient features (e.g. area, circularity, or some other measure of shape) and to make some decision, typically using feature-based discrimination. This process frequently uses software-embedded (hard-coded) knowledge of the scene.

There are two complementary areas of industrial machine vision: robot vision and automated visual inspection. Both of these use essentially the same techniques and approaches, although the visual inspection tasks are, in general, not as difficult as those involved in visual perception for robotic parts manipulation, identification, and assembly. This is because the inspection environment is usually easier to control and the accept/reject decisions required for inspection are often easier to determine than the location and identification information needed for assembly. The significant problems associated with robotic part handling, too, has meant that advanced three-dimensional robot vision has not received the attention it merits.

1.3 Sensory feedback for manufacturing systems: why vision?

The answer to this question must necessarily be double-barrelled:

1. We need feedback because the manufacturing system is not perfect and free of errors: we wish to ensure that we are informed when errors begin to creep into the process, so that we can take corrective action and ensure that quality and productivity are maintained.
2. We use vision because it is by far the most versatile sense available and conveys extremely rich information when compared with, e.g., sonar or infra-red sensing. Furthermore, unlike tactile sensing, it senses the environment in

a remote manner rather than having to be in contact with the objects being analysed.

Systems which are equipped with (useful) visual capabilities are inherently adaptive and can deal with uncertainty in the environment, or at least that is what one would hope for. The upshot of this is that, by incorporating vision in the manufacturing process, not only can we identify when things go wrong (e.g. in visual inspection) but the uncertain and variable nature of the manufacturing environment can be catered for.

Unfortunately, vision, while versatile, is also the most complex of the senses, due mainly to the fact that most information in visual images is implicitly coded and requires extensive processing and analysis to make it explicit. *Visual sensing is difficult*: in humans, ten of the estimated one hundred cortical areas in the brain are devoted to vision and much work remains to be done before we can claim to have even a modest grasp of visual sensing.

Given that one acknowledges that vision is (potentially) a very powerful sense, let us look at some of the motivations for using visual processes in the industrial workplace.

☐ *Safety and reliability*

Considerations of safety and reliability usually arise in environments which are hazardous for humans (e.g. in close proximity to a milling bit) or because manufactured parts are of critical importance and 100 per cent inspection is required (e.g. defects in a brake lining might conceivably cause loss of life).

Machine vision also facilitates consistency in inspection standards; such systems don't suffer from the 'Monday-morning' syndrome and their performance can be (and should be) quantitatively assessed.

☐ *Product quality*

High-volume production using humans seldom facilitates inspection of all parts but automated visual inspection techniques may make it feasible; this depends on the complexity of the task and the effective throughput that is required by the manufacturing system. The latter consideration is particularly important if the vision system is to be incorporated in an on-line manner, i.e. inspecting each part as it is manufactured.

☐ *Flexible automation*

In environments where quality assurance is performed by a machine, it is feasible to integrate the inspection task into the complete production or manufacturing cycle, and allow it to provide feedback to facilitate on-line control. This provides for the adaptive requirements of AMT (advanced manufacturing technology) systems and facilitates the overall control by computer, such as is found (or, more realistically, as will be found in the future) in advanced computer integrated manufacturing (CIM) environments.