Preface

What is new in the third edition?

Image processing and computer vision has been, and continues to be, subject to much research and development. The research develops into books and so the books need updating. We have always been interested to note that our book contains stock image processing and computer vision techniques which are yet to be found in other regular textbooks (OK, some is to be found in specialist books, though these rarely include much tutorial material). This has been true of the previous editions and certainly occurs here.

In this third edition, the completely new material is on new methods for low- and high-level feature extraction and description and on moving object detection, tracking, and description. We have also extended the book to use color and more modern techniques for object extraction and description especially those capitalizing on wavelets and on scale space. We have of course corrected the previous production errors and included more tutorial material where appropriate. We continue to update the references, especially to those containing modern survey material and performance comparison. As such, this book—IOHO—remains the most AU:1 up-to-date text in feature extraction and image processing in computer vision.

Why did we write this book?

We always expected to be asked: "why on earth write a new book on computer vision?", and we have been. A fair question is "there are already many good books on computer vision out in the bookshops, as you will find referenced later, so why add to them?" Part of the answer is that any textbook is a snapshot of material that exists prior to it. Computer vision, the art of processing images stored within a computer, has seen a considerable amount of research by highly qualified people and the volume of research would appear even to have increased in recent years. That means a lot of new techniques have been developed, and many of the more recent approaches are yet to migrate to textbooks. It is not just the new research: part of the speedy advance in computer vision technique has left some areas covered only in scanty detail. By the nature of research, one cannot publish material on technique that is seen more to fill historical gaps, rather than to advance knowledge. This is again where a new text can contribute.

Finally, the technology itself continues to advance. This means that there is new hardware, new programming languages, and new programming environments. In particular for computer vision, the advance of technology means that computing power and memory are now relatively cheap. It is certainly considerably cheaper than when computer vision was starting as a research field. One of

the authors here notes that the laptop in which his portion of the book was written on has considerably more memory, is faster, and has bigger disk space and better graphics than the computer that served the entire university of his student days. And he is not that old! One of the more advantageous recent changes brought by progress has been the development of mathematical programming systems. These allow us to concentrate on mathematical technique itself rather than on implementation detail. There are several sophisticated flavors of which Matlab, one of the chosen vehicles here, is (arguably) the most popular. We have been using these techniques in research and in teaching, and we would argue that they have been of considerable benefit there. In research, they help us to develop technique faster and to evaluate its final implementation. For teaching, the power of a modern laptop and a mathematical system combines to show students, in lectures and in study, not only how techniques are implemented but also how and why they work with an explicit relation to conventional teaching material.

We wrote this book for these reasons. There is a host of material we could have included but chose to omit; the taxonomy and structure we use to expose the subject are of our own construction. Our apologies to other academics if it was your own, or your favorite, technique that we chose to omit. By virtue of the enormous breadth of the subject of image processing and computer vision, we restricted the focus to feature extraction and image processing in computer vision for this has been the focus of not only our research but also where the attention of established textbooks, with some exceptions, can be rather scanty. It is, however, one of the prime targets of applied computer vision, so would benefit from better attention. We have aimed to clarify some of its origins and development, while also exposing implementation using mathematical systems. As such, we have written this text with our original aims in mind and maintained the approach through the later editions.

The book and its support

Each chapter of this book presents a particular package of information concerning feature extraction in image processing and computer vision. Each package is developed from its origins and later referenced to more recent material. Naturally, there is often theoretical development prior to implementation. We have provided working implementations of most of the major techniques we describe, and applied them to process a selection of imagery. Though the focus of our work has been more in analyzing medical imagery or in biometrics (the science of recognizing people by behavioral or physiological characteristic, like face recognition), the techniques are general and can migrate to other application domains.

You will find a host of further supporting information at the book's web site http://www.ecs.soton.ac.uk/~msn/book/. First, you will find the **worksheets** (the Matlab and Mathcad implementations that support the text) so that you can study

the techniques described herein. The **demonstration** site too is there. The web site will be kept up-to-date as much as possible, for it also contains links to other material such as web sites devoted to techniques and applications as well as to available software and online literature. Finally, any errata will be reported there. It is our regret and our responsibility that these will exist, and our inducement for their reporting concerns a pint of beer. If you find an error that we don't know about (not typos like spelling, grammar, and layout) then use the "mailto" on the web site and we shall send you a pint of good English **beer**, free!

There is a certain amount of mathematics in this book. The target audience is the third- or fourth-year students of BSc/BEng/MEng in electrical or electronic engineering, software engineering, and computer science, or in mathematics or physics, and this is the level of mathematical analysis here. Computer vision can be thought of as a branch of applied mathematics, though this does not really apply to some areas within its remit and certainly applies to the material herein. The mathematics essentially concerns mainly calculus and geometry, though some of it is rather more detailed than the constraints of a conventional lecture course might allow. Certainly, not all the material here is covered in detail in undergraduate courses at Southampton.

Chapter 1 starts with an overview of computer vision hardware, software, and established material, with reference to the most sophisticated vision system yet "developed": the **human vision** system. Though the precise details of the nature of processing that allows us to see are yet to be determined, there is a considerable range of **hardware** and **software** that allow us to give a computer system the capability to acquire, process, and reason with imagery, the function of "sight." The first chapter also provides a comprehensive **bibliography** of material you can find on the subject including not only textbooks but also available software and other material. As this will no doubt be subject to change, it might well be worth consulting the web site for more up-to-date information. The preference for journal references is those which are likely to be found in local university libraries or on the Web, *IEEE Transactions* in particular. These are often subscribed to as they are relatively of low cost and are often of very high quality.

Chapter 2 concerns the basics of signal processing theory for use in computer vision. It introduces the **Fourier transform** that allows you to look at a signal in a new way, in terms of its frequency content. It also allows us to work out the minimum size of a picture to conserve information, to analyze the content in terms of frequency, and even helps to speed up some of the later vision algorithms. Unfortunately, it does involve a few equations, but it is a new way of looking at data and at signals and proves to be a rewarding topic of study in its own right. It extends to wavelets, which are a popular analysis tool in image processing.

In Chapter 3, we ~~then~~ start to look at **basic** image processing techniques, where image points are mapped into a new value first by considering a single point in an original image and then by considering groups of points. Not only do we see common operations to make a picture's appearance better, especially for

human vision, but also we see how to reduce the effects of different types of commonly encountered image noise. We shall see some of the modern ways to remove noise and thus clean images, and we shall also look at techniques which process an image using notions of shape rather than mapping processes.

Chapter 4 concerns **low-level features** which are the techniques that describe the content of an image, at the level of a whole image rather than in distinct regions of it. One of the most important processes we shall meet is called **edge detection**. Essentially, this reduces an image to a form of a caricaturist's sketch, though without a caricaturist's exaggerations. The major techniques are presented in detail, together with descriptions of their implementation. Other image properties we can derive include measures of **curvature**, which developed into modern methods of **feature extraction**, and measures of **movement**. These are also covered in this chapter.

These edges, the curvature, or the motion need to be grouped in some way so that we can find shapes in an image and are dealt in Chapter 5. Using basic thresholding rarely suffices for shape extraction. One of the newer approaches is to group low-level features to find an object—in a way this is object extraction without shape. Another approach to **shape extraction** concerns analyzing the **match** of low-level information to a known template of a target shape. As this can be computationally very cumbersome, we then progress to a technique that improves computational performance, while maintaining an optimal performance. The technique is known as the **Hough transform** and it has long been a popular target for researchers in computer vision who have sought to clarify its basis, improve its speed, and to increase its accuracy and robustness. Essentially, by the Hough transform, we estimate the parameters that govern a shape's appearance, where the shapes range from **lines** to **ellipse**s and even to **unknown shapes**.

In Chapter 6, some applications of shape extraction require ~~too~~ determine rather more than the parameters that control appearance, and require to be able to **deform** or **flex** to match the image template. For this reason, the chapter on shape extraction by matching is followed by one on **flexible shape** analysis. This is a topic that has shown considerable progress of late, especially with the introduction of **snakes** (**active contours**). The newer material is the formulation by level set methods and brings new power to shape extraction techniques. These seek to match a shape to an image by analyzing local properties. Further, we shall see how we can describe a shape by its **skeleton** though with practical difficulty which can be alleviated by **symmetry** (though this can be slow), and also how global constraints concerning the **statistics** of a shape's appearance can be used to guide final extraction.

Up to this point, we have not considered techniques that can be used to describe the shape found in an image. In Chapter 7, we shall find that the two major approaches concern techniques that describe a shape's perimeter and those that describe its area. Some of the **perimeter description** techniques, the Fourier descriptors, are even couched using Fourier transform theory that allows analysis of their frequency content. One of the major approaches to **area description**, statistical moments, also has a form of access to frequency components, though it is

AU:2

of a very different nature to the Fourier analysis. One advantage is that insight into descriptive ability can be achieved by **reconstruction** which should get back to the original shape.

Chapter 8 describes **texture** analysis and also serves as a vehicle for introductory material on **pattern classification**. Texture describes patterns with no known analytical description and has been the target of considerable research in computer vision and image processing. It is used here more as a vehicle for material that precedes it, such as the Fourier transform and area descriptions though references are provided for access to other generic material. There is also introductory material on how to classify these patterns against known data, with a selection of the distance measures that can be used within that, and this is a window on a much larger area, to which appropriate pointers are given.

Finally, Chapter 9 concerns detecting and analyzing **moving objects**. Moving objects are detected by separating the foreground from the background, known as **background subtraction**. Having separated the moving components, one approach is then to follow or **track** the object as it moves within a sequence of image frames. The moving object can be described and recognized from the tracking information or by collecting together the sequence of frames to derive moving object descriptions.

The **appendices** include materials that are germane to the text, such as **camera models** and **coordinate geometry**, the method of **least squares**, a topic known as **principal components analysis**, and methods of **color description**. These are aimed to be short introductions and are appendices since they are germane to much of the material throughout but not needed directly to cover it. Other related material is referenced throughout the text, especially online material.

In this way, the text covers all major areas of feature extraction and image processing in computer vision. There is considerably more material in the subject than is presented here; for example, there is an enormous volume of material in 3D computer vision and in 2D signal processing, which is only alluded to here. Topics that are specifically not included are 3D processing, watermarking, and image coding. To include all these topics would lead to a monstrous book that no one could afford or even pick up. So we admit we give a snapshot, and we hope more that it is considered to open another window on a fascinating and rewarding subject.

In gratitude

Prof. Xiaoguang Jia, Prof. Yan Qiu Chen, Dr. Adrian Evans, Dr. Colin Davies, Dr. Mark Jones, Dr. David Cunado, Dr. Jason Nash, Dr. Ping Huang, Dr. Liang Ng, Dr. David Benn, Dr. Douglas Bradshaw, Dr. David Hurley, Dr. John Manslow, Dr. Mike Grant, Bob Roddis, Dr. Andrew Tatem, Dr. Karl Sharman, Dr. Jamie Shutler, Dr. Jun Chen, Dr. Andy Tatem, Dr. Chew-Yean Yam, Dr. James Hayfron-Acquah, Dr. Yalin Zheng, Dr. Jeff Foster, Dr. Peter Myerscough, Dr. David Wagg, Dr. Ahmad Al-Mazeed, Dr. Jang-Hee Yoo, Dr. Nick Spencer, Dr. Stuart Mowbray, Dr. Stuart Prismall, Dr. Peter Gething, Dr. Mike Jewell, Dr. David Wagg, Dr. Alex Bazin, Hidayah Rahmalan, Dr. Xin Liu, Dr. Imed Bouchrika, Dr. Banafshe Arbab-Zavar, Dr. Dan Thorpe, Dr. Cem Direkoglu, Dr. Sina Samangooei, Dr. John Bustard, Alastair Cummings, Mina Ibrahim, Muayed Al-Huseiny, Gunawan Ariyanto, Sung-Uk Jung, Richard Lowe, Dan Reid, George Cushen, Nick Udell, Ben Waller, Anas Abuzaina, Mus'ab Sahrim, Ari Rheum, Thamer Alathari, Tim Matthews and John Evans (for the great hippo photo), and to Jamie Hutton, Ben Dowling, and Sina again (for the Java demonstrations site). There has been much input from Mark's postdocs too, omitting those already mentioned, they include Dr. Hugh Lewis, Dr. Richard Evans, Dr. Lee Middleton, Dr. Galina Veres, Dr. Baofeng Guo, and Dr. Michaela Goffredo. We are also very grateful to other past Southampton students on BEng and MEng Electronic Engineering, MEng Information Engineering, BEng and MEng Computer Engineering, MEng Software Engineering, and BSc Computer Science who have pointed out our earlier mistakes (and enjoyed the beer), have noted areas for clarification, and in some cases volunteered some of the material herein. Beyond Southampton, we remain grateful to the reviewers of the three editions, to those who have written in and made many helpful suggestions, and to Prof. Daniel Cremers, Dr. Timor Kadir, Prof. Tim Cootes, Prof. Larry Davis, Dr. Pedro Felzenszwalb, Prof. Luc van Gool, and Prof. Aaron Bobick, for observations on and improvements to the text and/or for permission to use images. To all of you, our very grateful thanks.

Final message

We ourselves have already benefited much by writing this book. As we already know, previous students have also benefited and contributed to it as well. It remains our hope that it does inspire people to join in this fascinating and rewarding subject that has proved to be such a source of pleasure and inspiration to its many workers.

Mark S. Nixon
Electronics and Computer Science,
University of Southampton

Alberto S. Aguado
Black Rock Studio
December 2011