1 Introduction

1.1 Overview

This is where we start, by looking at the human visual system to investigate what is meant by vision, then on to how a computer can be made to sense pictorial data and then how we can process an image. The overview of this chapter is shown in Table 1.1; you will find a similar overview at the start of each chapter. There are no references (citations) in the overview, citations are made in the text and are collected at the end of each chapter.

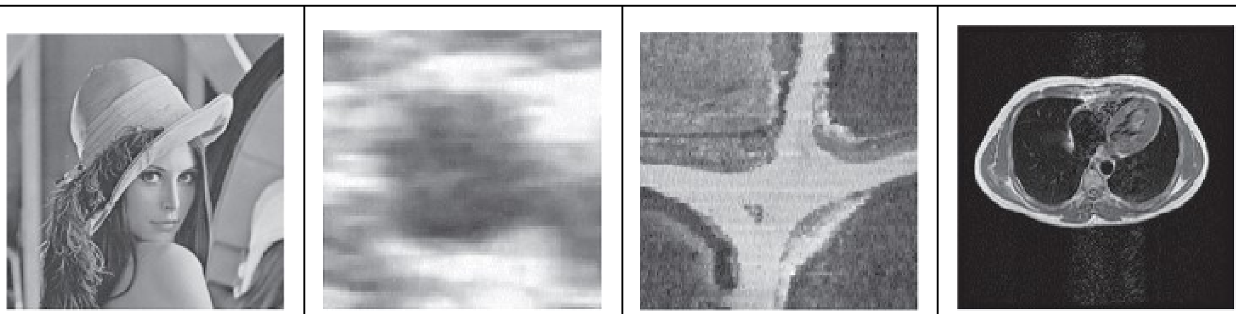
Table 1.1 Overview of Chapter 1

Main topic	Sub topics	Main points
Human vision system	How the *eye* works, how visual *information* is *processed* and how it can *fail*.	Sight, lens, retina, image, colour, monochrome, processing, brain, visual illusions.
Computer vision systems	How electronic *images* are formed, how *video* is fed into a *computer* and how we can *process* the information using a computer.	Picture elements, pixels, video standard, camera technologies, pixel technology, performance effects, specialist cameras, video conversion, computer languages, processing packages. Demonstrations of working techniques.
Mathematical systems	How we can process images using *mathematical packages*; introduction to the *Matlab* and *Mathcad* systems.	Ease, consistency, support, visualization of results, availability, introductory use, example worksheets.
Literature	Other *textbooks* and other places to find *information* on image processing, computer vision and feature extraction.	Magazines, textbooks, websites and this book's website.

1.2 Human and computer vision

A computer vision system processes images acquired from an electronic camera, which is like the human vision system where the brain processes images derived from the eyes. Computer vision is a rich and rewarding topic for study and research for electronic engineers, computer scientists and many others. Increasingly, it has a commercial future. There are now many vision systems in routine industrial use: cameras inspect mechanical parts to check size, food is inspected for quality, and images used in astronomy benefit from computer vision techniques. Forensic studies and biometrics (ways to recognize people) using computer vision include automatic face recognition and recognizing people by the ‘texture’ of their irises. These studies are paralleled by biologists and psychologists who continue to study how our human vision system works, and how we see and recognize objects (and people).

A selection of (computer) images is given in Figure 1.1; these images comprise a set of points or *picture elements* (usually concatenated to *pixels*) stored as an *array of numbers* in a *computer*. To recognize faces, based on an image such as Figure 1.1(a), we need to be able to analyse constituent shapes, such as the shape of the nose, the eyes and the eyebrows, to make some measurements to describe, and then recognize, a face. (Figure 1.1a is perhaps one of the most famous images in image processing. It is called the Lena image, and is derived from a picture of Lena Sjööblom in *Playboy* in 1972.) Figure 1.1(b) is an ultrasound image of the carotid artery (which is near the side of the neck and supplies blood to the brain and the face), taken as a cross-section through it. The top region of the image is near the skin; the bottom is inside the neck. The image arises from combinations of the reflections of the ultrasound radiation by tissue. This image comes from a study that aimed to produce three-dimensional (3D) models of arteries, to aid vascular surgery. Note that the image is very *noisy*, and this obscures the shape of the (elliptical) artery. Remotely sensed images are often analysed by their *texture* content. The perceived texture is different between the road junction and the different types of foliage seen in Figure 1.1(c). Finally, Figure 1.1(d) is a magnetic resonance imaging (MRI) image of a cross-section near the middle of a human body. The chest is at the top of the image, the lungs and blood vessels are the dark areas, and the internal organs and the fat appear grey. Nowadays, MRI images are in routine medical use, owing to their ability to provide high-quality images.



(a) Face from a camera	(b) Artery from ultrasound	(c) Ground by remote sensing	(d) Body by magnetic resonance

Figure 1.1 Real images from different sources

There are many different image sources. In medical studies, MRI is good for imaging soft tissue, but does not reveal the bone structure (the spine cannot be seen in Figure 1.1d); this can be achieved by using computed tomography (CT), which is better at imaging bone, as opposed to soft tissue. Remotely sensed images can be derived from infrared (thermal) sensors or synthetic-aperture radar, rather than by cameras, as in Figure 1.1(c). Spatial information can be provided by two-dimensional arrays of sensors, including sonar arrays. There are perhaps more varieties of sources of spatial data in medical studies than in any other area. But computer vision techniques are used to analyse any form of data, not just the images from cameras.

Synthesized images are good for *evaluating* techniques and finding out how they work, and some of the bounds on *performance*. Two synthetic images are shown in Figure 1.2. Figure 1.2(a) is an image of circles that were specified *mathematically*. The image is an ideal case: the circles are perfectly defined and the brightness levels have been specified to be constant. This type of synthetic image is good for evaluating techniques which find the borders of the shape (its edges) and the shape itself, and even for making a description of the shape. Figure 1.2(b) is a synthetic image made up of sections of real image data. The borders between the regions of image data are exact, again specified by a program. The image data comes from a well-known texture database, the Brodatz album of textures. This was scanned and stored as a computer image. This image can be used to analyse how well computer vision algorithms can identify regions of differing texture.

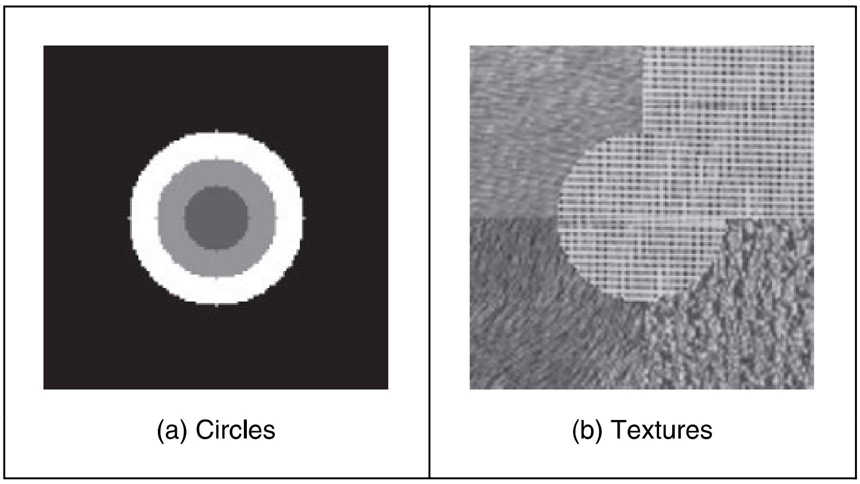


Figure 1.2 Examples of synthesized images

This chapter will show you how basic computer vision systems work, in the context of the human vision system. It covers the main elements of human vision, showing you how your eyes work (and how they can be deceived). For computer vision, this chapter covers the hardware and the software used for image analysis, giving an introduction to Mathcad and Matlab, the software tools used throughout this text to implement computer vision algorithms. Finally, a selection of pointers to other material is provided, especially those for more detail on the topics covered in this chapter.

1.3 The human vision system

Human vision is a sophisticated system that senses and acts on *visual stimuli*. It has evolved for millions of years, primarily for defence or survival. Intuitively, computer and human vision appear to have the same function. The purpose of both systems is to interpret *spatial* data, data that is indexed by more than one dimension. Even though computer and human vision are functionally similar, you cannot expect a computer vision system to replicate exactly the function of the human eye. This is partly because we do not understand fully how the vision system of the eye and brain works, as we shall see in this section. Accordingly, we cannot design a system to replicate its function exactly. In fact, some of the properties of the human eye are useful when developing computer vision techniques, whereas others are actually undesirable in a computer vision system. But we shall see computer vision techniques which can, to some extent, replicate, and in some cases even improve upon, the human vision system.