1 Applications and Tools

1.1 A Tool for Science and Technique

From the beginning of science, visual observation has played a major role. At that time, the only way to document the results of an experiment was by verbal description and manual drawings. The next major step was the invention of *photography* which enabled results to be documented objectively. Three prominent examples of scientific applications of photography are *astronomy*, *photogrammetry*, and *particle physics*. Astronomers were able to measure positions and magnitudes of stars and photogrammeters produced topographic maps from aerial images. Searching through countless images from hydrogen bubble chambers led to the discovery of many elementary particles in physics. These manual evaluation procedures, however, were time consuming. Some semi- or even fully automated optomechanical devices were designed. However, they were adapted to a single specific purpose. This is why quantitative evaluation of images did not find widespread application at that time. Generally, images were only used for documentation, qualitative description, and illustration of the phenomena observed.

Nowadays, we are in the middle of a second revolution sparked by the rapid progress in video and computer technology. Personal computers and workstations have become powerful enough to process image data. As a result, multimedia software and hardware is becoming standard for the handling of images, image sequences, and even 3-D visualization. The technology is now available to any scientist or engineer. In consequence, image processing has expanded and is further rapidly expanding from a few specialized applications into a standard scientific tool. Image processing techniques are now applied to virtually all the natural sciences and technical disciplines.

A simple example clearly demonstrates the power of visual information. Imagine you had the task of writing an article about a new technical system, for example, a new type of solar power plant. It would take an enormous effort to describe the system if you could not include images and technical drawings. The reader of your imageless article would also have a frustrating experience. He or she would spend a lot of time trying to figure out how the new solar power plant worked and might end up with only a poor picture of what it looked like.

B. Jähne, Digital Image Processing
ISBN 3-540-24035-7

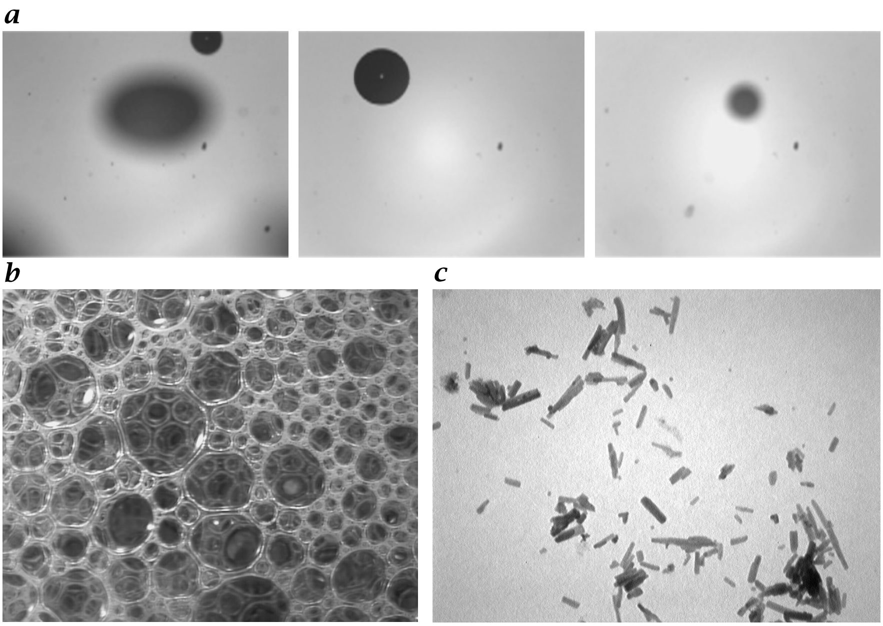

__Figure 1.1:__ Measurement of particles with imaging techniques: __a__ Bubbles submerged by breaking waves using a telecentric illumination and imaging system; from Geißler and Jähne [57]. __b__ Soap bubbles. __c__ Electron microscopy of color pigment particles (courtesy of Dr. Klee, Hoechst AG, Frankfurt).

Technical drawings and photographs of the solar power plant would be of enormous help for readers of your article. They would immediately have an idea of the plant and could study details in the images that were not described in the text, but which caught their attention. Pictures provide much more information, a fact which can be precisely summarized by the saying that "a picture is worth a thousand words".

Another observation is of interest. If the reader later heard of the new solar plant, he or she could easily recall what it looked like, the object "solar plant" being instantly associated with an image.

1.2 Examples of Applications

In this section, examples for scientific and technical applications of digital image processing are discussed. The examples demonstrate that image processing enables complex phenomena to be investigated, which could not be adequately accessed with conventional measuring techniques.

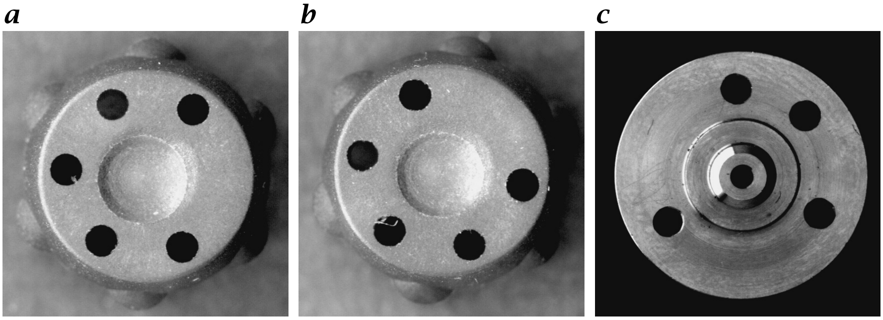

Figure 1.2: *Industrial parts that are checked by a visual inspection system for the correct position and diameter of holes (courtesy of Martin von Brocke, Robert Bosch GmbH).*

1.2.1 Counting and Gauging

A classic task for digital image processing is counting particles and measuring their size distribution. Figure 1.1 shows three examples with very different particles: gas bubbles submerged by breaking waves, soap bubbles, and pigment particles. The first challenge with tasks like this is to find an imaging and illumination setup that is well adapted to the measuring problem. The bubble images in Fig. 1.1a are visualized by a *telecentric* illumination and imaging system. With this setup, the principal rays are parallel to the optical axis. Therefore the size of the imaged bubbles does not depend on their distance. The sampling volume for concentration measurements is determined by estimating the degree of blurring in the bubbles.

It is much more difficult to measure the shape of the soap bubbles shown in Fig. 1.1b, because they are transparent. Therefore, deeper lying bubbles superimpose the image of the bubbles in the front layer. Moreover, the bubbles show deviations from a circular shape so that suitable parameters must be found to describe their shape.

A third application is the measurement of the size distribution of color pigment particles. This significantly influences the quality and properties of paint. Thus, the measurement of the distribution is an important quality control task. The image in Fig. 1.1c taken with a transmission electron microscope shows the challenge of this image processing task. The particles tend to cluster. Consequently, these clusters have to be identified, and — if possible — to be separated in order not to bias the determination of the size distribution.

Almost any product we use nowadays has been checked for defects by an automatic *visual inspection* system. One class of tasks includes the checking of correct sizes and positions. Some example images are shown in Fig. 1.2. Here the position, diameter, and roundness of the

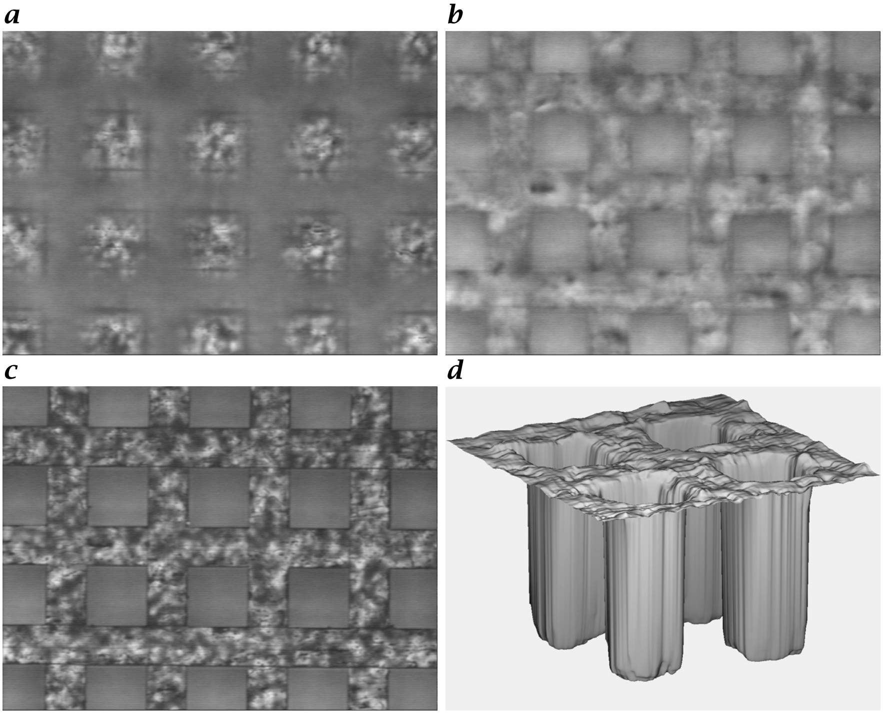

*Figure 1.3: Focus series of a press form of PMMA with narrow rectangular holes imaged with a confocal technique using statistically distributed intensity patterns. The images are focused on the following depths measured from the bottom of the holes: **a** 16 µm, **b** 480 µm, and **c** 620 µm (surface of form). **d** 3-D reconstruction. From Scheuermann et al. [178].*

holes is checked. Figure 1.2c illustrates that it is not easy to illuminate metallic parts. The edge of the hole on the left is partly bright and thus it is more difficult to detect and to measure the holes correctly.

1.2.2 Exploring 3-D Space

In images, 3-D scenes are projected on a 2-D image plane. Thus the depth information is lost and special imaging techniques are required to retrieve the topography of surfaces or volumetric images. In recent years, a large variety of range imaging and volumetric imaging techniques have been developed. Therefore image processing techniques are also applied to *depth maps* and *volumetric images*.

Figure 1.3 shows the reconstruction of a press form for microstructures that has been imaged by a special type of confocal microscopy [178]. The form is made out of PMMA, a semi-transparent plastic ma-

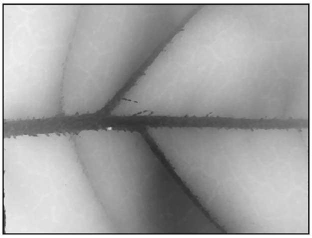

Figure 1.4: *Depth map of a plant leaf measured by optical coherency tomography (courtesy of Jochen Restle, Robert Bosch GmbH).*

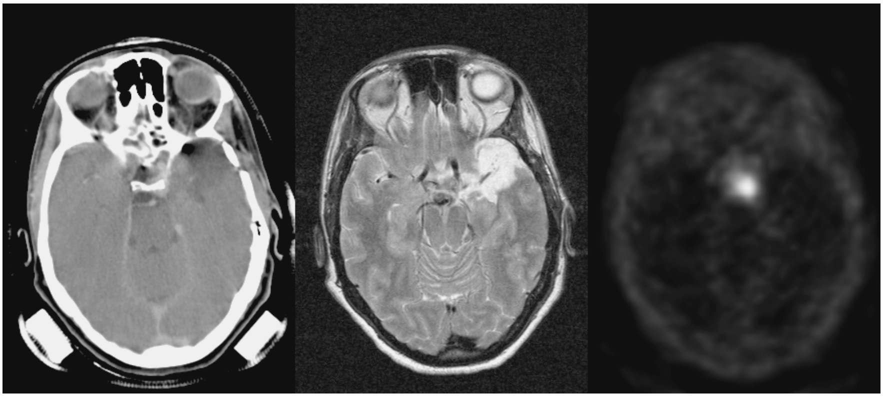

Figure 1.5: *Horizontal scans at the eye level across a human head with a tumor. The scans are taken with x-rays (left), T2 weighted* magnetic resonance tomography *(middle), and* positron emission tomography *(right; images courtesy of Michael Bock, DKFZ Heidelberg).*

terial with a smooth surface, so that it is almost invisible in standard microscopy. The form has narrow, 500 μm deep rectangular holes.

In order to make the transparent material visible, a statistically distributed pattern is projected through the microscope optics onto the focal plane. This pattern only appears sharp on parts that lie in the focal plane. The pattern gets more blurred with increasing distance from the focal plane. In the focus series shown in Fig. 1.3, it can be seen that first the patterns of the material in the bottom of the holes become sharp (Fig. 1.3a), then after moving the object away from the optics, the final image focuses at the surface of the form (Fig. 1.3c). The depth of the surface can be reconstructed by searching for the position of maximum contrast for each pixel in the focus series (Fig. 1.3d).

Figure 1.4 shows the depth map of a plant leaf that has been imaged with another modern optical 3-D measuring technique known as *white-light interferometry* or *coherency radar*. It is an interferometric technique that uses light with a coherency length of only a few wavelengths.

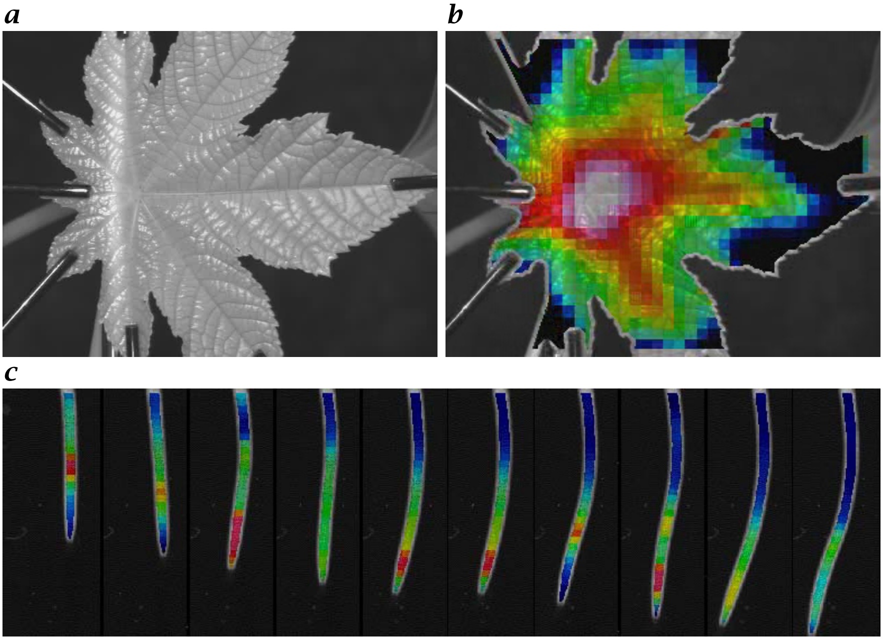

*Figure 1.6: Growth studies in botany: **a** Rizinus plant leaf; **b** map of growth rate; **c** Growth of corn roots (courtesy of Uli Schurr and Stefan Terjung, Institute of Botany, University of Heidelberg).*

Thus interference patterns occur only with very short path differences in the interferometer. This effect can be utilized to measure distances with accuracy in the order of a wavelength of light used.

Medical research is the driving force for the development of modern volumetric imaging techniques that allow us to look into the interior of 3-D objects. Figure 1.5 shows a scan through a human head. Whereas *x-rays* (*computer tomography*, *CT*) predominantly delineate the bone structures, the T2-weighted *magnetic resonance tomography* (*MRT*) shows the soft tissues, the eyes and scar tissue with high signal intensity. With *positron emission tomography* (*PET*) a high signal is observed at the tumour location because here the administered positron emitter is accumulating.

1.2.3 Exploring Dynamic Processes

The exploration of dynamic processes is possible by analyzing *image sequences*. The enormous potential of this technique is illustrated with a number of examples in this section.

In botany, a central topic is the study of the growth of plants and the mechanisms controlling growth processes. Figure 1.6a shows a Riz-

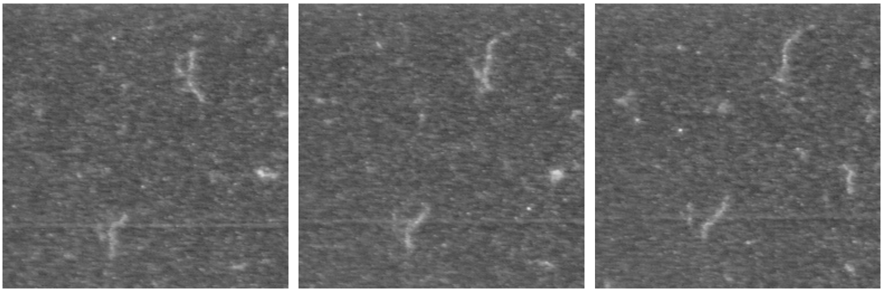

Figure 1.7: *Motility assay for motion analysis of motor proteins (courtesy of Dietmar Uttenweiler, Institute of Physiology, University of Heidelberg).*

inus plant leaf from which a map of the growth rate (percent increase of area per unit time) has been determined by a time-lapse image sequence where about every minute an image was taken. This new technique for growth rate measurements is sensitive enough for area-resolved measurements of the diurnal cycle.

Figure 1.6c shows an image sequence (from left to right) of a growing corn root. The gray scale in the image indicates the growth rate, which is largest close to the tip of the root.

In science, images are often taken at the limit of the technically possible. Thus they are often plagued by high noise levels. Figure 1.7 shows fluorescence-labeled motor proteins that are moving on a plate covered with myosin molecules in a so-called *motility assay*. Such an assay is used to study the molecular mechanisms of muscle cells. Despite the high noise level, the motion of the filaments is apparent. However, automatic motion determination with such noisy image sequences is a demanding task that requires sophisticated image sequence analysis techniques.

The next example is taken from oceanography. The small-scale processes that take place in the vicinity of the ocean surface are very difficult to measure because of undulation of the surface by waves. Moreover, point measurements make it impossible to infer the 2-D structure of the waves at the water surface. Figure 1.8 shows a space-time image of short wind waves. The vertical coordinate is a spatial coordinate in the wind direction and the horizontal coordinate the time. By a special illumination technique based on the *shape from shading* paradigm (Section 8.5.3), the along-wind slope of the waves has been made visible. In such a spatiotemporal image, motion is directly visible by the inclination of lines of constant gray scale. A horizontal line marks a static object. The larger the angle to the horizontal axis, the faster the object is moving. The image sequence gives a direct insight into the complex nonlinear dynamics of wind waves. A fast moving large wave modulates the motion of shorter waves. Sometimes the short waves move with

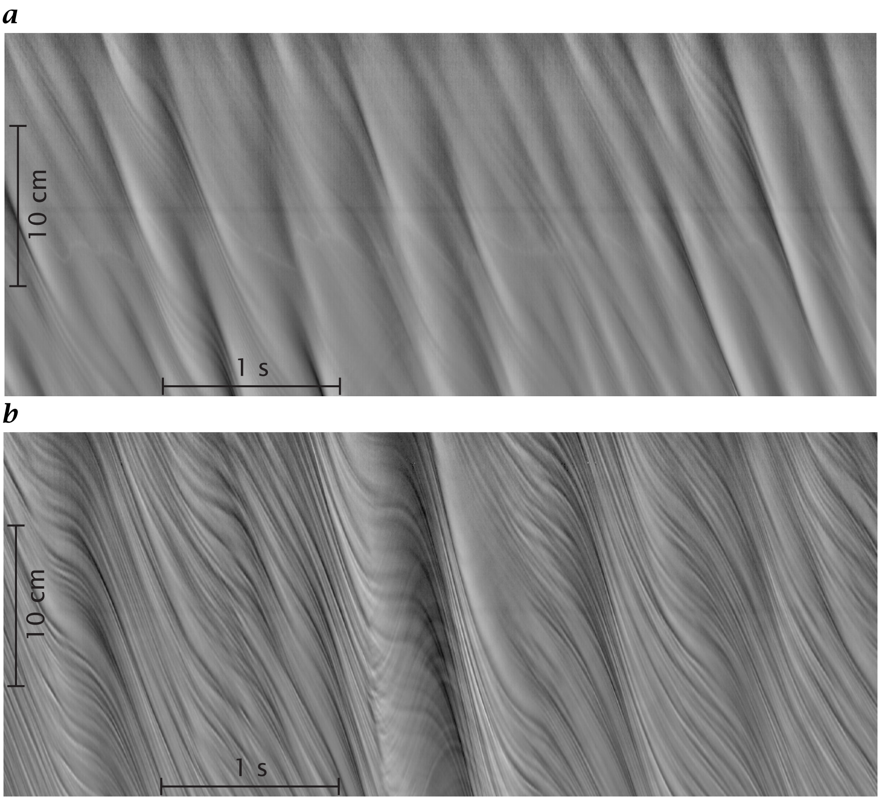

Figure 1.8: *A space-time image of short wind waves at a wind speed of* ***a*** *2.5 and* ***b*** *7.5 m/s. The vertical coordinate is the spatial coordinate in wind direction, the horizontal coordinate the time.*

the same speed (bound waves), but mostly they are significantly slower showing large modulations in the phase speed and amplitude.

The last example of image sequences is on a much larger spatial and temporal scale. Figure 1.9 shows the annual cycle of the tropospheric column density of NO_2. NO_2 is one of the most important trace gases for the atmospheric ozone chemistry. The main sources for tropospheric NO_2 are industry and traffic, forest and bush fires (biomass burning), microbiological soil emissions, and lighting. Satellite imaging allows for the first time the study of the regional distribution of NO_2 and the identification of the sources and their annual cycles.

The data have been computed from spectroscopic images obtained from the GOME instrument of the ERS2 satellite. At each pixel of the images a complete spectrum with 4000 channels in the ultraviolet and visible range has been taken. The total atmospheric column density of the NO_2 concentration can be determined by the characteristic absorp-

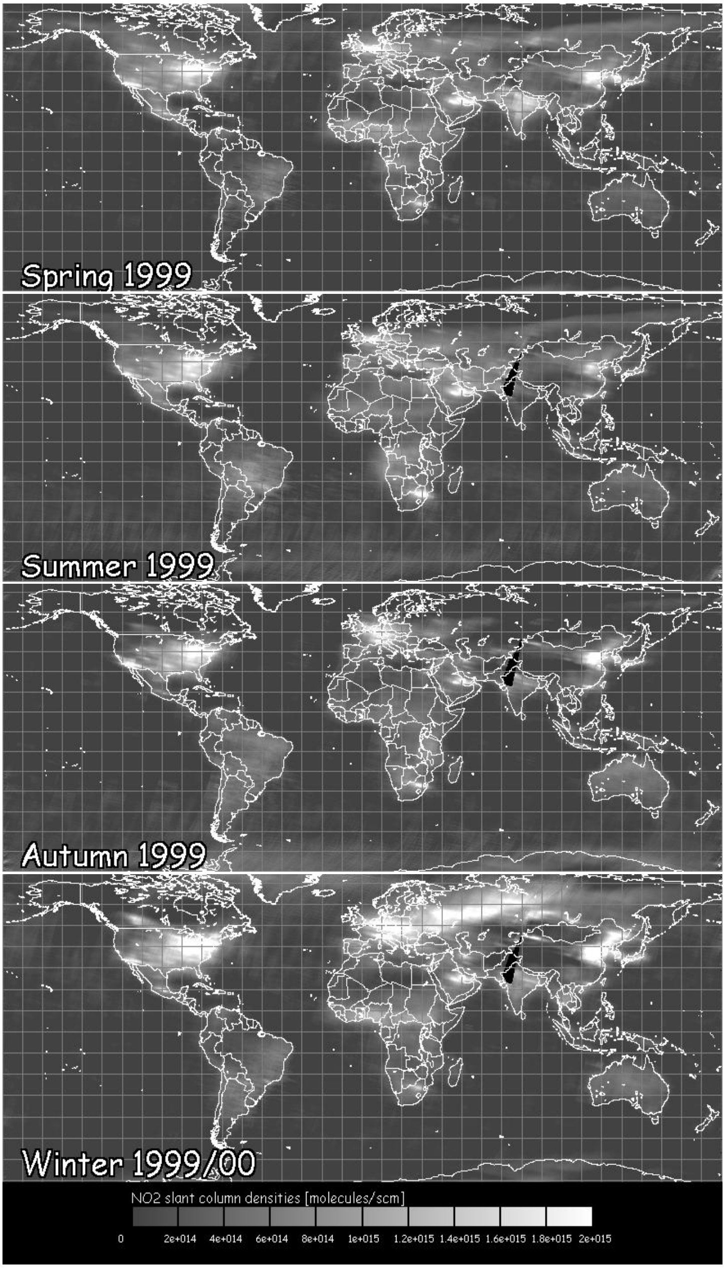

Figure 1.9: *Maps of tropospheric NO_2 column densities showing four three-month averages from 1999 (courtesy of Mark Wenig, Institute for Environmental Physics, University of Heidelberg).*

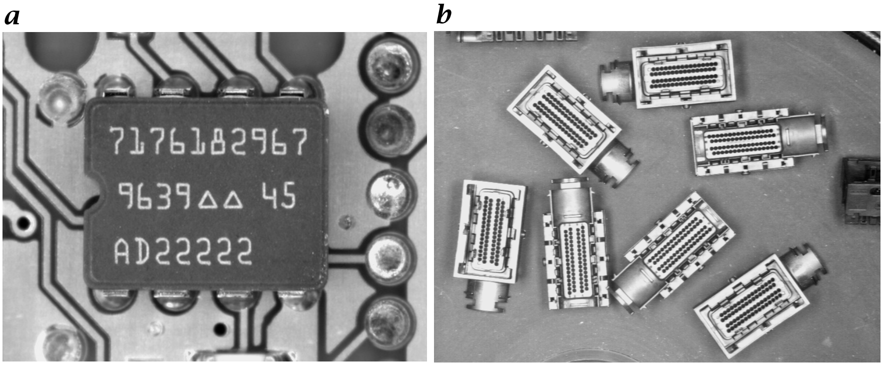

Figure 1.10: *Industrial inspection tasks:* ***a*** *Optical character recognition.* ***b*** *Connectors (courtesy of Martin von Brocke, Robert Bosch GmbH).*

tion spectrum that is, however, superimposed by the absorption spectra of other trace gases. Therefore, a complex nonlinear regression analysis is required. Furthermore, the stratospheric column density must be subtracted by suitable image processing algorithms.

The resulting maps of tropospheric NO_2 column densities in Fig. 1.9 show a lot of interesting detail. Most emissions are related to industrialized countries. They show a clear annual cycle in the Northern hemisphere with a maximum in the winter.

1.2.4 Classification

Another important task is the classification of objects observed in images. The classical example of classification is the recognition of characters (*optical character recognition* or short *OCR*). Figure 1.10a shows a typical industrial OCR application, the recognition of a label on an integrated circuit. Object classification includes also the recognition of different possible positioning of objects for correct handling by a robot. In Fig. 1.10b, connectors are placed in random orientation on a conveyor belt. For proper pick up and handling, whether the front or rear side of the connector is seen must also be detected.

The classification of defects is another important application. Figure 1.11 shows a number of typical errors in the inspection of integrated circuits: an incorrectly centered surface mounted resistor (Fig. 1.11a), and broken or missing bond connections (Fig. 1.11b–f).

The application of classification is not restricted to industrial tasks. Figure 1.12 shows some of the most distant galaxies ever imaged by the Hubble telescope. The galaxies have to be separated into different classes according to their shape and color and have to be distinguished from other objects, e. g., stars.

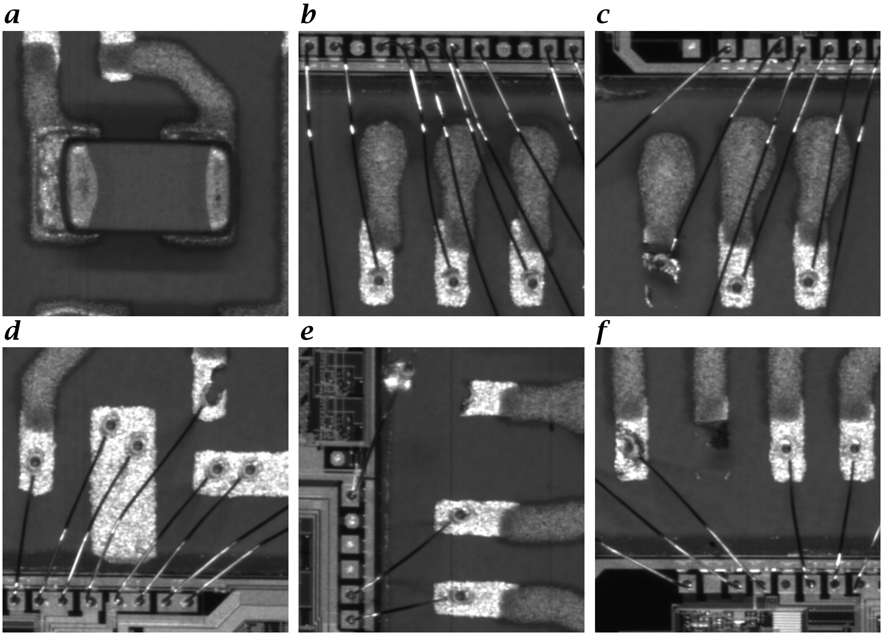

Figure 1.11: *Errors in soldering and bonding of integrated circuits. Courtesy of Florian Raisch, Robert Bosch GmbH).*

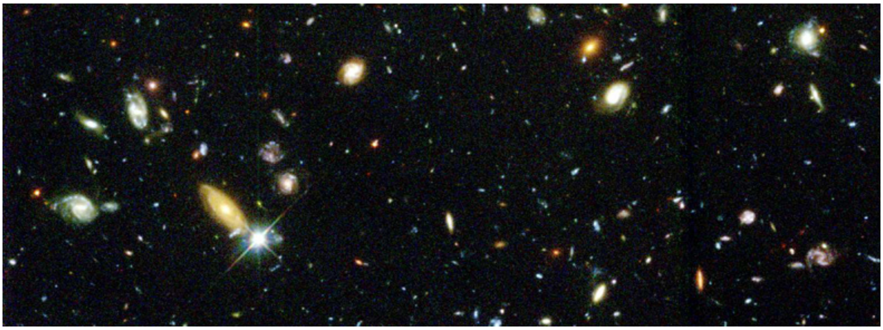

Figure 1.12: *Hubble deep space image: classification of distant galaxies (http://hubblesite.org/).*

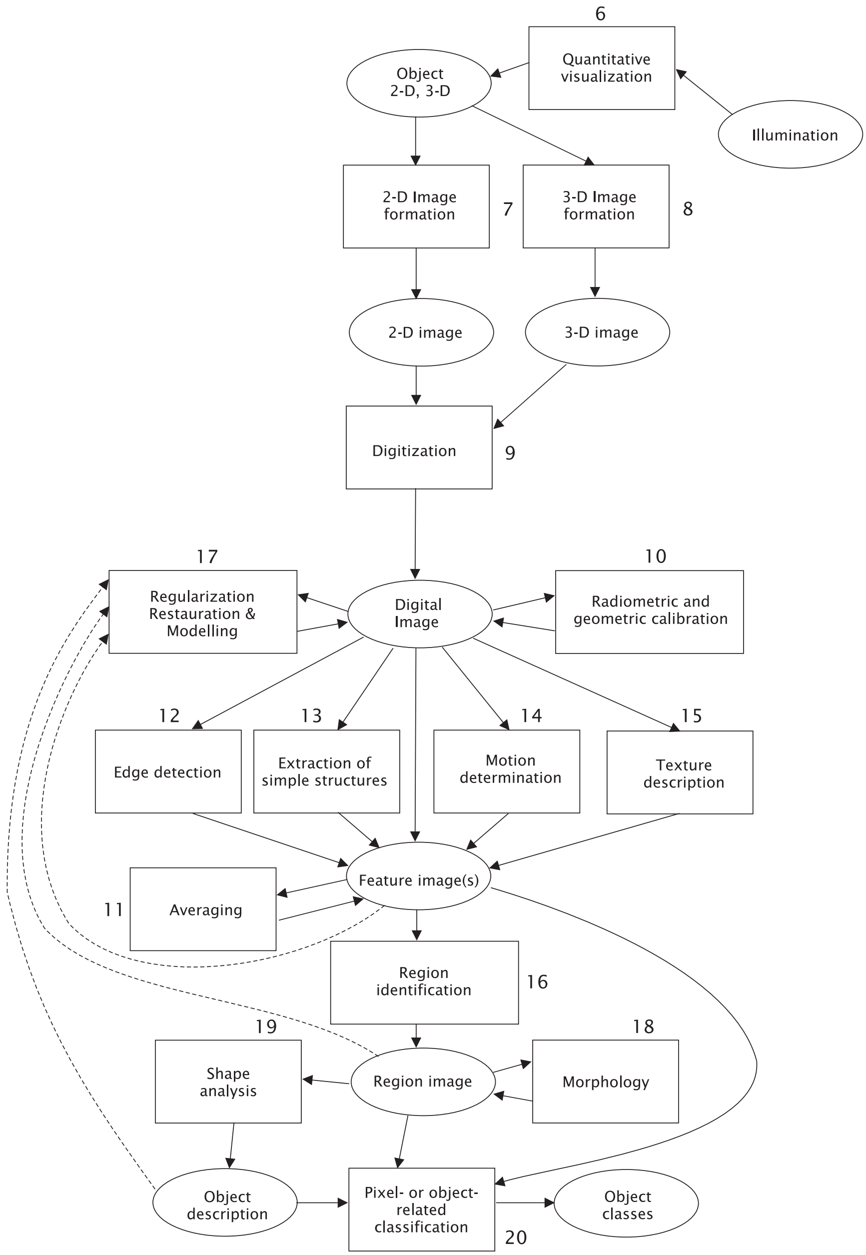

Figure 1.13: *A hierarchy of digital image processing tasks from image formation to image comprehension. The numbers by the boxes indicate the corresponding chapters of this book.*

1.3 Hierarchy of Image Processing Operations

Image processing is not a one-step process. We are able to distinguish between several steps which must be performed one after the other until we can extract the data of interest from the observed scene. In this way a *hierarchical processing* scheme is built up as sketched in Fig. 1.13. The figure gives an overview of the different phases of image processing, together with a summary outline of this book.

Image processing begins with the capture of an image with a suitable, not necessarily optical, acquisition system. In a technical or scientific application, we may choose to select an appropriate imaging system. Furthermore, we can set up the illumination system, choose the best wavelength range, and select other options to capture the object feature of interest in the best way in an image (Chapter 6). 2-D and 3-D image formation are discussed in Chapters 7 and 8, respectively. Once the image is sensed, it must be brought into a form that can be treated with digital computers. This process is called *digitization* and is discussed in Chapter 9.

The first steps of digital processing may include a number of different operations and are known as *image preprocessing*. If the sensor has nonlinear characteristics, these need to be corrected. Likewise, brightness and contrast of the image may require improvement. Commonly, too, coordinate transformations are needed to restore geometrical distortions introduced during image formation. Radiometric and geometric corrections are elementary pixel processing operations that are discussed in Chapter 10.

A whole chain of processing steps is necessary to analyze and identify objects. First, adequate filtering procedures must be applied in order to distinguish the objects of interest from other objects and the background. Essentially, from an image (or several images), one or more *feature images* are extracted. The basic tools for this task are averaging (Chapter 11), edge detection (Chapter 12), the analysis of simple neighborhoods (Chapter 13) and complex patterns known in image processing as *texture* (Chapter 15). An important feature of an object is also its *motion*. Techniques to detect and determine motion are discussed in Chapter 14.

Then the object has to be separated from the background. This means that regions of constant features and discontinuities must be identified by *segmentation* (Chapter 16). This can be an easy task if an object is well distinguished from the background by some local features. This is, however, not often the case. Then more sophisticated segmentation techniques are required (Chapter 17). These techniques use various optimization strategies to minimize the deviation between the image data and a given model function incorporating the knowledge about the objects in the image.

The same mathematical approach can be used for other image processing tasks. Known disturbances in the image, for instance caused by a defocused optics, motion blur, errors in the sensor, or errors in the transmission of image signals, can be corrected (*image restoration*). Images can be reconstructed from indirect imaging techniques such as *tomography* that deliver no direct image (*image reconstruction*).

Now that we know the geometrical shape of the object, we can use morphological operators to analyze and modify the shape of objects (Chapter 18) or extract further information such as the mean gray value, the area, perimeter, and other parameters for the form of the object (Chapter 19). These parameters can be used to classify objects (*classification*, Chapter 20). Character recognition in printed and handwritten text is an example of this task.

While it appears logical to divide a complex task such as image processing into a succession of simple subtasks, it is not obvious that this strategy works at all. Why? Let us discuss a simple example. We want to find an object that differs in its gray value only slightly from the background in a noisy image. In this case, we cannot simply take the gray value to differentiate the object from the background.

Averaging of neighboring image points can reduce the noise level. At the edge of the object, however, background and object points are averaged, resulting in false mean values. If we knew the edge, averaging could be stopped at the edge. But we can determine the edges only after averaging because only then are the gray values of the object sufficiently different from the background.

We may hope to escape this circular argument by an iterative approach. We just apply the averaging and make a first estimate of the edges of the object. We then take this first estimate to refine the averaging at the edges, recalculate the edges and so on. It remains to be studied in detail, however, whether this iteration converges at all, and if it does, whether the limit is correct.

In any case, the discussed example suggests that more difficult image processing tasks require feedback. Advanced processing steps give parameters back to preceding processing steps. Then the processing is not linear along a chain but may iteratively loop back several times. Figure 1.13 shows some possible feedbacks. The feedback may include non-image processing steps.

If an image processing task cannot be solved with a given image, we may decide to change the illumination, zoom closer to an object of interest or to observe it under a more suitable view angle. This type of approach is known as *active vision*. In the framework of an intelligent system exploring its environment by its senses we may also speak of an *action-perception cycle*.

1.4 Image Processing and Computer Graphics

For some time now, image processing and *computer graphics* have been treated as two different areas. Knowledge in both areas has increased considerably and more complex problems can now be treated. Computer graphics is striving to achieve *photorealistic* computer-generated images of three-dimensional scenes, while image processing is trying to reconstruct one from an image actually taken with a camera. In this sense, *image processing* performs the inverse procedure to that of computer graphics. In computer graphics we start with knowledge of the shape and features of an object — at the bottom of Fig. 1.13 — and work upwards until we get a two-dimensional image. To handle image processing or computer graphics, we basically have to work from the same knowledge. We need to know the interaction between illumination and objects, how a three-dimensional scene is projected onto an image plane, etc.

There are still quite a few differences between an image processing and a graphics workstation. But we can envisage that, when the similarities and interrelations between computer graphics and image processing are better understood and the proper hardware is developed, we will see some kind of general-purpose workstation in the future which can handle computer graphics as well as image processing tasks. The advent of multimedia, i. e., the integration of text, images, sound, and movies, will further accelerate the unification of computer graphics and image processing. The term "*visual computing*" has been coined in this context [66].

1.5 Cross-disciplinary Nature of Image Processing

By its very nature, the science of image processing is cross-disciplinary in several aspects. First, image processing incorporates concepts from various sciences. Before we can process an image, we need to know how the digital signal is related to the features of the imaged objects. This includes various physical processes from the interaction of radiation with matter to the geometry and radiometry of imaging. An imaging sensor converts the incident irradiance in one or the other way into an electric signal. Next, this signal is converted into digital numbers and processed by a digital computer to extract the relevant data. In this chain of processes (see also Fig. 1.13) many areas from *physics*, *computer science* and *mathematics* are involved including among others, optics, solid state physics, chip design, computer architecture, algebra, analysis, statistics, algorithm theory, graph theory, system theory, and numerical mathematics. From an engineering point of view, contributions from *optical engineering*, *electrical engineering*, *photonics*, and *software engineering* are required.

Image processing has a partial overlap with other disciplines. Image processing tasks can partly be regarded as a measuring problem, which is part of the science of *metrology*. Likewise, *pattern recognition* tasks are incorporated in image processing in a similar way as in *speech processing*. Other disciplines with similar connections to image processing are the areas of *neural networks*, *artificial intelligence*, and *visual perception*. Common to these areas is their strong link to biological sciences.

When we speak of *computer vision*, we mean a computer system that performs the same task as a biological vision system to "discover from images what is present in the world, and where it is" [132]. In contrast, the term *machine vision* is used for a system that performs a vision task such as checking the sizes and completeness of parts in a manufacturing environment. For many years, a vision system has been regarded just as a passive observer. As with biological vision systems, a computer vision system can also actively explore its surroundings by, e. g., moving around and adjusting its angle of view. This, we call *active vision*.

There are numerous special disciplines that for historical reasons developed partly independently of the main stream in the past. One of the most prominent disciplines is *photogrammetry* (measurements from photographs; main applications: mapmaking and surveying). Other areas are *remote sensing* using aerial and satellite images, *astronomy*, and *medical imaging*.

The second important aspect of the cross-disciplinary nature of image processing is its widespread application. There is almost no field in natural sciences or technical disciplines where image processing is not applied. As we have seen from the examples in Section 1.2, it has gained crucial importance in several application areas. The strong links to so many application areas provide a fertile ground for further rapid progress in image processing because of the constant inflow of techniques and ideas from an ever-increasing host of application areas.

A final cautionary note: a cross-disciplinary approach is not just a nice extension. It is a necessity. Lack of knowledge in either the application area or image processing tools inevitably leads at least to suboptimal solutions and sometimes even to a complete failure.

1.6 Human and Computer Vision

We cannot think of image processing without considering the *human visual system*. This seems to be a trivial statement, but it has far-reaching consequences. We observe and evaluate the images that we process with our visual system. Without taking this elementary fact into consideration, we may be much misled in the interpretation of images.
The first simple questions we should ask are:

- What intensity differences can we distinguish?

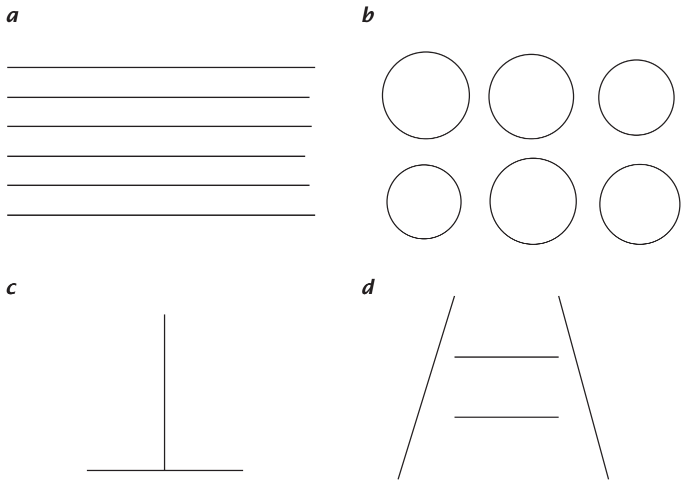

Figure 1.14: *Test images for distance and area estimation:* ***a*** *parallel lines with up to 5 % difference in length;* ***b*** *circles with up to 10 % difference in radius;* ***c*** *the vertical line appears longer, though it has the same length as the horizontal line;* ***d*** *deception by perspective: the upper line (in the background) appears longer than the lower line (in the foreground), though both are equally long.*

- What is the spatial resolution of our eye?
- How accurately can we estimate and compare distances and areas?
- How do we sense colors?
- By which features can we detect and distinguish objects?

It is obvious that a deeper knowledge would be of immense help for computer vision. Here is not the place to give an overview of the human visual system. The intention is rather to make us aware of the elementary relations between human and computer vision. We will discuss diverse properties of the human visual system in the appropriate chapters. Here, we will make only some introductory remarks. A detailed comparison of human and computer vision can be found in Levine [121]. An excellent up-to-date reference to human vision is also the monograph by Wandell [210].

The reader can perform some experiments by himself. Figure 1.14 shows several test images concerning the question of estimation of distance and area. He will have no problem in seeing even small changes in the length of the parallel lines in Fig. 1.14a. A similar area comparison with circles is considerably more difficult (Fig. 1.14b). The other examples show how the estimate is biased by the context of the image. Such phenomena are known as *optical illusions*. Two examples of estimates for length are shown in Fig. 1.14c, d. These examples show

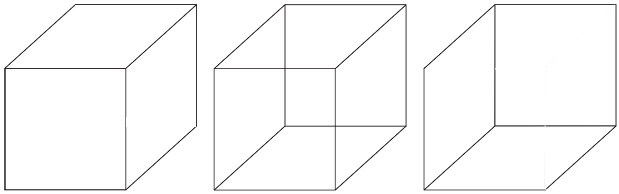

Figure 1.15: *Recognition of three-dimensional objects: three different representations of a cube with identical edges in the image plane.*

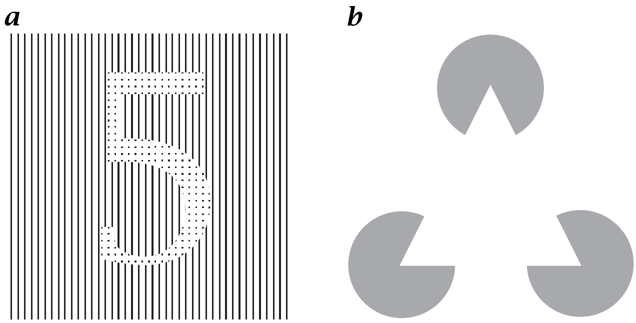

Figure 1.16: a *Recognition of boundaries between textures;* ***b*** *"interpolation" of object boundaries.*

that the human visual system interprets the context in its estimate of length. Consequently, we should be very careful in our visual estimates of lengths and areas in images.

The second topic is that of the recognition of objects in images. Although Fig. 1.15 contains only a few lines and is a planar image not containing any direct information on depth, we immediately recognize a cube in the right and left image and its orientation in space. The only clues from which we can draw this conclusion are the hidden lines and our knowledge about the shape of a cube. The image in the middle, which also shows the hidden lines, is ambivalent. With some training, we can switch between the two possible orientations in space.

Figure 1.16 shows a remarkable feature of the human visual system. With ease we see sharp boundaries between the different textures in Fig. 1.16a and immediately recognize the figure 5. In Fig. 1.16b we identify a white equilateral triangle, although parts of the bounding lines do not exist.

From these few observations, we can conclude that the human visual system is extremely powerful in recognizing objects, but is less well suited for accurate measurements of gray values, distances, and areas.

In comparison, the power of computer vision systems is marginal and should make us feel humble. A digital image processing system can only perform elementary or well-defined fixed image processing tasks such as real-time quality control in industrial production. A computer vision system has also succeeded in steering a car at high speed on a highway, even with changing lanes. However, we are still worlds away from a universal digital image processing system which is capable of “understanding” images as human beings do and of reacting intelligently and flexibly in real time.

Another connection between human and computer vision is worth noting. Important developments in computer vision have been made through progress in understanding the human visual system. We will encounter several examples in this book: the *pyramid* as an efficient data structure for image processing (Chapter 5), the concept of local orientation (Chapter 13), and motion determination by filter techniques (Chapter 14).

1.7 Components of an Image Processing System

This section briefly outlines the capabilities of modern image processing systems. A general purpose image acquisition and processing system typically consists of four essential components:

1. An image acquisition system. In the simplest case, this could be a CCD camera, a flatbed scanner, or a video recorder.
2. A device known as a frame grabber to convert the electrical signal (normally an analog video signal) of the image acquisition system into a digital image that can be stored.
3. A personal computer or a workstation that provides the processing power.
4. Image processing software that provides the tools to manipulate and analyze the images.

1.7.1 Image Sensors

Digital processing requires images to be obtained in the form of electrical signals. These signals can be digitized into sequences of numbers which then can be processed by a computer. There are many ways to convert images into digital numbers. Here, we will focus on video technology, as it is the most common and affordable approach.

The milestone in image sensing technology was the invention of semiconductor photodetector arrays. There are many types of such sensors, the most common being the *charge coupled device* or *CCD*. Such a sensor consists of a large number of photosensitive elements. During the accumulation phase, each element collects electrical charges, which are generated by absorbed photons. Thus the collected charge is proportional

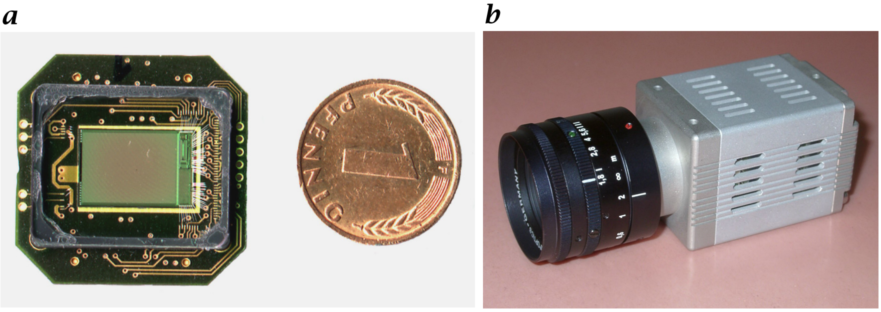

Figure 1.17: *Modern semiconductor cameras:* ***a*** *Complete CMOS camera on a chip with digital and analog output (image courtesy of K. Meier, Kirchhoff Institute for Physics, University of Heidelberg), [126].* ***b*** *High-end digital 12-bit CCD camera, Pixelfly (image courtesy of PCO GmbH, Germany).*

to the illumination. In the read-out phase, these charges are sequentially transported across the chip from sensor to sensor and finally converted to an electric voltage.

For quite some time, *CMOS image sensors* have been available. But only recently have these devices attracted significant attention because the image quality, especially the uniformity of the sensitivities of the individual sensor elements, now approaches the quality of CCD image sensors. CMOS imagers still do not reach up to the standards of CCD imagers in some features, especially at low illumination levels (higher dark current). They have, however, a number of significant advantages over CCD imagers. They consume significantly less power, subareas can be accessed quickly, and they can be added to circuits for image preprocessing and signal conversion. Indeed, it is possible to put a whole camera on a single chip (Fig. 1.17a). Last but not least, CMOS sensors can be manufactured more cheaply and thus open new application areas.

Generally, semiconductor imaging sensors are versatile and powerful devices:

- *Precise and stable geometry.* The individual sensor elements are precisely located on a regular grid. Geometric distortion is virtually absent. Moreover, the sensor is thermally stable in size due to the low linear thermal expansion coefficient of silicon ($2 \cdot 10^{-6}$/K). These features allow precise size and position measurements.
- *Small and rugged.* The sensors are small and insensitive to external influences such as magnetic fields and vibrations.
- *High sensitivity.* The *quantum efficiency*, i. e., the fraction of elementary charges generated per photon, can be close to one (≻ R2 and ≻ R1). Even standard imaging sensors, which are operated at room temperature, have a low noise level of only 10-100 electrons. Thus

they show an excellent sensitivity. Cooled imaging sensors can be used with exposure times of hours without showing a significant thermal signal.
However, commercial CCDs at room temperature cannot be used at low light levels because of the thermally generated electrons. But if CCD devices are cooled down to low temperatures, they can be exposed for hours. Such devices are commonly used in astronomy and are about one hundred times more sensitive than photographic material.

- *Wide variety.* Imaging sensors are available in a wide variety of resolutions and frame rates ($\succ$ R2 and $\succ$ R1). The largest built CCD sensor as of 2001 originates from Philips. In a modular design with $1k \times 1k$ sensor blocks, they built a $7k \times 9k$ sensor with $12 \times 12\,\mu$m pixels [68]. Among the fastest high-resolution imagers available is the 1280×1024 active-pixel CMOS sensor from Photobit with a peak frame rate of 500 Hz (660 MB/s data rate) [152].
- *Imaging beyond the visible.* Semiconductor imagers are not limited to the visible range of the electromagnetic spectrum. Standard silicon imagers can be made sensitive far beyond the visible wavelength range (400–700 nm) from 200 nm in the *ultraviolet* to 1100 nm in the near *infrared*. In the infrared range beyond 1100 nm, other semiconductors such an GaAs, InSb, HgCdTe are used ($\succ$ R3) since silicon becomes transparent. Towards shorter wavelengths, specially designed silicon imagers can be made sensitive well into the *x-ray* wavelength region.

1.7.2 Image Acquisition and Display

A frame grabber converts the electrical signal from the camera into a digital image that can be processed by a computer. Image display and processing nowadays no longer require any special hardware. With the advent of graphical user interfaces, image display has become an integral part of a personal computer or workstation. Besides the display of grayscale images with up to 256 shades (8 bit), also true-color images with up to 16.7 million colors (3 channels with 8 bits each), can be displayed on inexpensive PC graphic display systems with a resolution of up to 1600×1200 pixels.

Consequently, a modern frame grabber no longer requires its own image display unit. It only needs circuits to digitize the electrical signal from the imaging sensor and to store the image in the memory of the computer. The direct transfer of image data from a frame grabber to the memory (RAM) of a microcomputer has become possible since 1995 with the introduction of fast peripheral bus systems such as the PCI bus. This 32-bit wide and 33 Mhz fast bus has a peak transfer rate of

132 MB/s. Depending on the PCI bus controller on the frame grabber and the chipset on the motherboard of the computer, sustained transfer rates between 15 and 80 MB/s have been reported. This is sufficient to transfer image sequences in real time to the main memory, even for color images and fast frame rate images. The second generation 64-bit, 66 MHz PCI bus quadruples the data transfer rates to a peak transfer rate of 512 MB/s. Digital cameras that transfer image data directly to the PC via standardized digital interfaces such as *Firewire* (*IEEE 1394*), *Camera link*, or even fast *Ethernet* will further simplify the image input to computers.

The transfer rates to standard hard disks, however, are considerably lower. Sustained transfer rates are typically lower than 10 MB/s. This is inadequate for uncompressed real-time image sequence storage to disk. Real-time transfer of image data with sustained data rates between 10 and 30 MB/s is, however, possible with *RAID arrays*.

1.7.3 Computer Hardware for Fast Image Processing

The tremendous progress of computer technology in the past 20 years has brought digital image processing to the desk of every scientist and engineer. For a general-purpose computer to be useful for image processing, four key demands must be met: high-resolution image display, sufficient memory transfer bandwidth, sufficient storage space, and sufficient computing power. In all four areas, a critical level of performance has been reached that makes it possible to process images on standard hardware. In the near future, it can be expected that general-purpose computers can handle volumetric images and/or image sequences without difficulties. In the following, we will briefly outline these key areas.

General-purpose computers now include sufficient random access memory (RAM) to store multiple images. A 32-bit computer can address up to 4 GB of memory. This is sufficient to handle complex image processing tasks even with large images. Nowadays, also 64-bit computer systems are available. They provide enough RAM even for demanding applications with image sequences and volumetric images.

While in the early days of personal computers hard disks had a capacity of just 5–10 MB, nowadays disk systems with more than ten thousand times more storage capacity (40–200 GB) are standard. Thus, a large number of images can be stored on a disk, which is an important requirement for scientific image processing. For permanent data storage and PC exchange, the *DVD* is playing an important role as a cheap and versatile storage medium. One DVD can hold almost 5 GB of image data that can be read independent of the operating system on MS Windows, Macintosh, and UNIX platforms. Cheap DVD writers allow anyone to produce DVDs.

Within the short history of microprocessors and personal computers, computing power has increased tremendously. From 1978 to 2001 the clock rate has increased from 4.7 MHz to 1.6 GHz by a factor of 300. The speed of elementary operations such as floating-point addition and multiplication has increased even more because on modern CPUs these operations have now a throughput of only a few clocks instead of about 100 on early processors. Thus, in less than 25 years, the speed of floating-point computations on a single microprocessor increased more than a factor of 10 000.

Image processing could benefit from this development only partly. On modern 32-bit processors it became increasingly inefficient to transfer and process 8-bit and 16-bit image data. This changed only in 1997 with the integration of multimedia techniques into PCs and workstations. The basic idea of fast image data processing is very simple. It makes use of the 64-bit data paths in modern processors for quick transfer and processing of multiple image data in parallel. This approach to parallel computing is a form of the *single instruction multiple data* (*SIMD*) concept. In 64-bit machines, eight 8-bit, four 16-bit or two 32-bit data can be processed together.

Sun was the first to integrate the SIMD concept into a general-purpose computer architecture with the *visual instruction set* (*VIS*) on the UltraSparc architecture [139]. In January 1997 Intel introduced the *Multimedia Instruction Set Extension* (*MMX*) for the next generation of Pentium processors (P55C). The SIMD concept was quickly adopted by other processor manufacturers. Motorola, for instance, developed the *AltiVec* instruction set. It has also become an integral part of new 64-bit architectures such as in *IA-64* architecture from Intel and the *x86-64* architecture from AMD.

Thus, it is evident that SIMD-processing of image data has become a standard part of future microprocessor architectures. More and more image processing tasks can be processed in real time on standard microprocessors without the need for any expensive and awkward special hardware. However, significant progress for compilers is still required before SIMD techniques can be used by the general programmer. Today, the user either depends on libraries that are optimized by the hardware manufacturers for specific hardware platforms or he is forced to dive into the details of hardware architectures for optimized programming.

1.7.4 Software and Algorithms

The rapid progress of computer hardware may distract us from the importance of software and the mathematical foundation of the basic concepts for image processing. In the early days, image processing may have been characterized more as an “art” than as a science. It was like tapping in the dark, empirically searching for a solution. Once an algorithm worked for a certain task, you could be sure that it would not work with other images and you would not even know why. Fortunately, this is gradually changing. Image processing is about to mature to a well-developed science. The deeper understanding has also led to a more realistic assessment of today's capabilities of image processing and analysis, which in many respects is still worlds away from the capability of human vision.

It is a widespread misconception that a better mathematical foundation for image processing is of interest only to the theoreticians and has no real consequences for the applications. The contrary is true. The advantages are tremendous. In the first place, mathematical analysis allows a distinction between image processing problems that can and those that cannot be solved. This is already very helpful. Image processing algorithms become predictable and accurate, and in some cases optimal results are known. New mathematical methods often result in novel approaches that can solve previously intractable problems or that are much faster or more accurate than previous approaches. Often the speed up that can be gained by a fast algorithm is considerable. In some cases it can reach up to several orders of magnitude. Thus fast algorithms make many image processing techniques applicable and reduce the hardware costs considerably.

1.8 *Exercises*

1.1: Image sequence viewer

Interactive viewing and inspection of all image sequences and volumetric images used throughout this textbook (dip6ex01.01).

1.2: ***Image processing tasks**

Figure 1.13 contains a systematic summary of the hierarchy of image processing operations from illumination to the analysis of objects extracted from the images taken. Investigate, which of the operations in this diagram are required for the following tasks.

1. Measurement of the size distribution of color pigments (Section 1.2.1, Fig. 1.1c)
2. Detection of a brain tumor in a volumetric magnetic resonance tomography image (Section 1.2.2, Fig. 1.5) and measurement of its size and shape
3. Investigation of the diurnal cycle of the growth of plant leaves (Section 1.2.3, Fig. 1.6)
4. Character recognition (OCR): Reading of the label on an integrated circuit (Section 1.2.4, Fig. 1.10a)
5. Partitioning of galaxies according to their form and spectrum into different classes (Section 1.2.4, Fig. 1.12)

1.3: *****Interdisziplinary nature of image processing**

1. Which other sciences contribute methods that are used in digital image processing?
2. Which areas of science and technology use digital image processing techniques?

1.4: ****Comparison of computer vision and biological vision**

In Section 1.7 we discuss the components of a digital image processing system. Try to identify the corresponding components of a biological vision system. Is there a one-to-one correspondence or do you see fundamental differences? Are there biological components that are not yet realized in computer vision systems and vice versa?

1.5: ***Amounts of data in digital image processing**

In digital image processing significantly larger amounts of data are required to be processed as this is normally the case with the analysis of time series. In order to get a feeling of the amount of data, estimate the amount of data that is to be processed in the following typical real-world applications.

1. **Water wave image sequences.** In a wind/wave facility image sequences are taken from wind waves at the surface of the water (Section 1.2.3, Fig. 1.8). Two camera systems are in use. Each of them takes image sequences with a spatial resolution of 640×480 pixel, 200 frames/s and 8 bit data resolution. A sequence of measurements runs over six hours. Every 15 minutes a sequence of 5 minutes is taken simultaneously with both cameras. How large is the data rate for real-time recording? How much data needs to be stored for the whole six hour run?
2. **Industrial inspection system for laser welding.** The welding of parts in an industrial production line is inspected by a high-speed camera system. The camera takes 256×256 large images with a rate of 1000 frames/s and a resolution of 16 bit per pixel for one second in order to control the welding of one part. One thousand parts are inspected per hour. The production line runs around the clock and includes six inspection places in total. Per hour 1000 parts are inspected. The line runs around the clock and includes six inspection places. Which amount of image data must be processed per day and year, respectively?
3. **Driver assistance system.** A driver assistance system detects the road lane and traffic signs with a camera system, which has a spatial resolution of 640×480 pixel and takes 25 frames/s. The camera delivers color images with the three color channels red, green, and blue. Which rate of image data (MB/s) must be processed in real time?
4. **Medical volumetric image sequences.** A fast computer tomographic systems for dynamic medical diagnosis takes volumetric images with a spatial resolution of $256 \times 256 \times 256$ and a repetition rate of 10 frames/s. The data are 16 bit deep. Which rate of data (MB/s) must be processed?

1.9 *Further Readings*

In this section, we give some hints on further readings in image processing.

Elementary textbooks. "The Image Processing Handbook" by Russ [173] is an excellent elementary introduction to image processing with a wealth of application examples and illustrations. Another excellent elementary textbook is Nalwa [144]. He gives — as the title indicates — a guided tour of computer vision.

Advanced textbooks. Still worthwhile to read is the classical, now almost twenty year old textbook "Digital Picture Processing" from Rosenfeld and Kak [172]. Another classical, but now somewhat outdated textbook is Jain [97]. From other classical textbooks new editions were published recently: Pratt [157]and Gonzalez and Woods [62]. The textbook of van der Heijden [205] discusses image-based measurements including parameter estimation and object recognition.

Textbooks covering special topics. Because of the cross-disciplinary nature of image processing (Section 1.5), image processing can be treated from quite different points of view. A collection of monographs is listed here that focus on one or the other aspect of image processing:

Topic	References
Image sensors	Holst [77], Howell [82], Janesick [99]
MR imaging	Haacke et al. [67], Liang and Lauterbur [122], Mitchell and Cohen [138]
Geometrical aspects of computer vision	Faugeras [42], Faugeras and Luong [43]
Perception	Mallot [129], Wandell [210]
Machine vision	Jain et al. [98], Demant et al. [31]
Robot vision and computer vision	Horn [81], Shapiro and Stockman [186], Forsyth and Ponce [54]
Signal processing	Granlund and Knutsson [64], Lim [124]
Satellite imaging and remote sensing	Richards and Jia [167], Schott [181]
Micro structure analysis	Ohser and Mücklich [147]
Industrial image processing	Demant et al. [31]
Object classification and pattern recognition	Duda et al. [38], Schürmann [182], Bishop [10], Schölkopf and Smola [180]
High-level vision	Ullman [202]

Human vision and computer vision. This topic is discussed in detail by Levine [121]. An excellent and up-to-date reference is also the monograph from Wandell [210].

Collection of articles. An excellent overview of image processing with direct access to some key original articles is given by the following collections of articles: "Digital Image Processing" by Chelappa [22], "Readings in Computer Vision: Issues, Problems, Principles, and Paradigms" by Fischler and Firschein [47], and "Computer Vision: Principles and Advances and Applications" by Kasturi and Jain [103, 104].

Handbooks. The "Practical Handbook on Image Processing for Scientific Applications" by Jähne [89] provides a task-oriented approach with many practical procedures and tips. A state-of-the-art survey of computer vision is given by the three-volume "Handbook of Computer Vision and Applications by Jähne et al. [94]. Algorithms for image processing and computer vision are provided by Voss and Süße [209], Pitas [154], Parker [150], Umbaugh [203], and Wilson and Ritter [217].