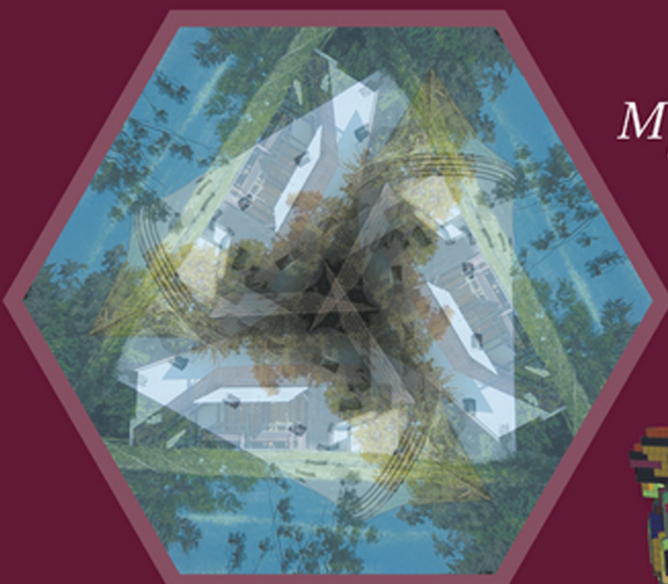
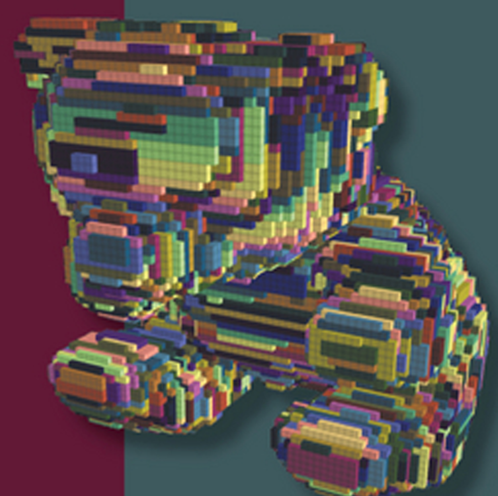
2D and 3D

Image Analysis by Moments



$$M_p^{(f)} = \int_\Omega \pi_p(x) f(x) dx$$



Jan Flusser
Tomáš Suk
Barbara Zitová

with website

WILEY

2D AND 3D IMAGE ANALYSIS BY MOMENTS

2D AND 3D IMAGE ANALYSIS BY MOMENTS

Jan Flusser, Tomáš Suk and Barbara Zitová

Institute of Information Theory and Automation,
Czech Academy of Sciences,
Prague,
Czech Republic

WILEY

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication data

Names: Flusser, Jan, author. | Suk, Tomáš, author. | Zitová, Barbara, author.
Title: 2D and 3D image analysis by moments / Jan Flusser, Tomáš Suk, Barbara Zitová.
Description: Chichester, UK ; Hoboken, NJ : John Wiley & Sons, 2016. | Includes bibliographical references and index.
Identifiers: LCCN 2016026517| ISBN 9781119039358 (cloth) | ISBN 9781119039372 (Adobe PDF) | ISBN 9781119039365 (epub)
Subjects: LCSH: Image analysis. | Moment problems (Mathematics) | Invariants.
Classification: LCC TA1637 F58 2016 | DDC 621.36/7015159–dc23 LC record available at https://lccn.loc.gov/2016026517

Set in 10/12pt, TimesLTStd by SPi Global, Chennai, India.

10 9 8 7 6 5 4 3 2 1

To my wife, Vlasta, my sons, Michal and Martin,
and my daughters, Jana and Veronika.

Jan Flusser

To my wife, Lenka, my daughter, Hana, and my son, Ondřej.

Tomáš Suk

To my husband, Pavel, my sons, Janek and Jakub,
and my daughter, Babetka.

Barbara Zitová

Contents

Preface

Seven years ago we published our first monograph on moments and their applications in image recognition: J. Flusser, T. Suk, and B. Zitová, *Moments and Moment Invariants in Pattern Recognition*, Wiley, 2009.

That book (referred to as MMIPR) was motivated by the need for a monograph covering theoretical aspects of moments and moment invariants and their relationship to practical image recognition problems. Long before 2009, object recognition had become an established discipline inside image analysis. Moments and moment invariants, introduced to the image analysis community in the early 1960s, have played a very important role as features in invariant recognition. Nevertheless, such a book had not been available before 2009[1].

The development of moment invariants after 2009 was even more rapid than before. In SCOPUS, which is probably the most widely-used publication database, we have received 16,000 search results as the response to the "image moment" keyword and 6,000 results of the "moment invariants" search[2]. There has been an overlap of about 2,000 papers, which results in 20,000 relevant papers in total. This huge number of entries illustrates how a large and important area of computer science has been formed by the methods based on image moments. In Figure 1 we can observe the development in time. A relatively slow growth in the last century was followed by a rapid increase of the number of publications in 2009–2010 (we believe that the appearance of MMIPR at least slightly contributed to the growing interest in moment invariants). Since then, the annual number of publications has slightly fluctuated, reaching another local maximum in 2014. In 2014, a new multi-authored book edited by G. A. Papakostas[3] appeared on the market. Although the editor did a very good job, this book suffers from a common weakness of multi-authored books – the topics of individual chapters had been selected by their authors according to their interest, which made some areas overstressed, while some others, remained unmentioned despite their importance. The Papakostas book reflects recent developments in some areas but can hardly be used as a course textbook.

[1] The very first moment-focused book by R. Mukundan and K. R. Ramakrishnan, *Moment Functions in Image Analysis*, World Scientific, 1998, is just a short introduction to this field. The second book by M. Pawlak, *Image Analysis by Moments: Reconstruction and Computational Aspects*, Wroclaw, Poland, 2006, is focused narrowly on numerical aspects of image moments, without providing a broader context of invariant image recognition and of practical applications.

[2] The search was performed within the title and abstract of the papers.

[3] G. A. Papakostas ed., *Moments and Moment Invariants – Theory and Applications*, Science Gate Publishing, 2014.

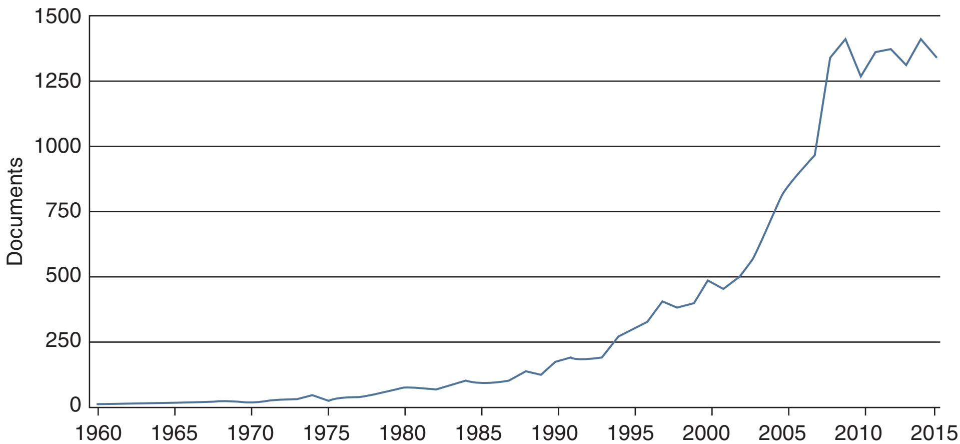


Figure 1 The number of moment-related publications as found in SCOPUS.

The great number of publications that have appeared since 2009 led us to the idea of writing a new, more comprehensive book on this topic. The abundant positive feedback we have received from the MMIPR readers and from our students was another factor which has strengthened our intentions. In 2014, the MMIPR was even translated into Chinese[4].

The main goal of the book you are now holding in your hands is to be a comprehensive monograph covering the current state of the art of moment-based image analysis and presenting the latest developments in this field in a consistent form. In this book, the reader will find a survey of all important theoretical results as well as a description of how to use them in practical image analysis tasks. In particular, our aims were

- To review the development of moments and moment invariants after 2009;
- To cover the area of 3D moments and invariants, which were mostly omitted in the MMIPR but have become important in the last few years;
- To present some of the latest unpublished results, especially in the field of blur invariants; and
- To provide readers with an introduction to a broader context, showing the role of the moments in image analysis and recognition and reviewing other relevant areas and approaches.

At the same time, we aimed to write a self-contained book. This led us to the decision (with the kind permission of Wiley) to include also the core parts of the MMIPR, of course in enhanced and extended/up-to-date form. Attentive readers may realize that about one half of this book was already treated in the MMIPR in some form, while the other half is original. In particular, Chapters 1, 2, 4, and 6 are completely or mostly original. Chapters 7, 8, 9, and 10 are substantially extended versions of their ancestors in the MMIPR, and Chapters 3 and 5 were adopted from the MMIPR after minor updates.

[4] Published by John Wiley & Univ. of Science and Technology of China Press, see `http://library.utia.cas.cz/separaty/2015/ZOI/flusser-0444327-cover.jpg`.

The book is based on our deep experience with moments and moment invariants gained from twenty-five years of research in this area, from teaching graduate courses on moment invariants and related fields at the Czech Technical University and at the Charles University, Prague, Czech Republic, and from presenting several tutorials on moments at major international conferences.

The target readership includes academic researchers and R&D engineers from all application areas who need to recognize 2D and 3D objects extracted from binary/graylevel/color images and who look for invariant and robust object descriptors, as well as specialists in moment-based image analysis interested in a new development on this field. Last but not least, the book is also intended for university lecturers and graduate students of image analysis and pattern recognition. It can be used as textbook for advanced graduate courses on Invariant Pattern Recognition. The first two chapters can be even utilized as supplementary reading to undergraduate courses on Pattern Recognition and Image Analysis.

We created an accompanying website at `http://zoi.utia.cas.cz/moment_invariants2` containing selected Matlab codes, the complete lists of the invariants, the slides for those who wish to use this book for educational purposes, and Errata (if any). This website is free for the book readers (the password can be found in the book) and is going to be regularly updated.

Authors' biographies

Prof. Jan Flusser, PhD, DrSc, received the MSc degree in mathematical engineering from the Czech Technical University, Prague, Czech Republic, in 1985, the PhD degree in computer science from the Czechoslovak Academy of Sciences in 1990, and the DrSc degree in technical cybernetics in 2001. Since 1985 he has been with the Institute of Information Theory and Automation, Czech Academy of Sciences, Prague. In 1995–2007, he held the position of head of Department of Image Processing. Since 2007 he has been a Director of the Institute. He is a full professor of computer science at the Czech Technical University, Faculty of Nuclear Science and Physical Engineering, and at the Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic, where he gives undergraduate and graduate courses on Digital Image Processing, Pattern Recognition, and Moment Invariants and Wavelets. Jan Flusser's research interests cover moments and moment invariants, image registration, image fusion, multichannel blind deconvolution, and super-resolution imaging. He has authored and coauthored more than 200 research publications in these areas, including the monograph *Moments and Moment Invariants in Pattern Recognition* (Wiley, 2009), approximately 60 journal papers, and 20 tutorials and invited/keynote talks at major conferences (ICIP'05, ICCS'06, COMPSTAT'06, ICIP'07, DICTA'07, EUSIPCO'07, CVPR'08, CGIM'08, FUSION'08, SPPRA'09, SCIA'09, ICIP'09, CGIM'10, AIA'14, and CI'15). His publications have received about approximately 1000 citations.

In 2007 Jan Flusser received the Award of the Chairman of the Czech Science Foundation for the best research project and won the Prize of the Czech Academy of Sciences for his contribution to image fusion theory. In 2010, he was awarded the SCOPUS 1000 Award presented by Elsevier. He received the Felber Medal of the Czech Technical University for excellent contribution to research and education in 2015. Personal webpage: `http://www.utia.cas.cz/people/flusser`

Tomáš Suk, PhD, received the MSc degree in technical cybernetics from the Czech Technical University, Prague, Czech Republic, in 1987 and the PhD degree in computer science from the Czechoslovak Academy of Sciences in 1992. Since 1992 he has been a research fellow with the Institute of Information Theory and Automation, Czech Academy of Sciences, Prague. His research interests include invariant features, moment and point-based invariants, color spaces, geometric transformations, and applications in botany, remote sensing, astronomy, medicine, and computer vision.

Tomáš Suk has authored and coauthored more than thirty journal papers and fifty conference papers in these areas, including tutorials on moment invariants held at the conferences ICIP'07 and SPPRA'09. He also coauthored the monograph *Moments and Moment Invariants in Pattern Recognition* (Wiley, 2009). His publications have received about approximately citations. In 2002 he received the Otto Wichterle Premium of the Czech Academy of Sciences for young scientists. Personal webpage: `http://zoi.utia.cas.cz/suk`

Barbara Zitová, PhD, received the MSc degree in computer science from the Charles University, Prague, Czech Republic, in 1995 and the Ph.D degree in software systems from the Charles University, Prague, Czech Republic, in 2000. Since 1995, she has been with the Institute of Information Theory and Automation, Czech Academy of Sciences, Prague. Since 2008 she has been the head of Department of Image Processing. She gives undergraduate and graduate courses on Digital Image Processing and Wavelets in Image Processing at the Czech Technical University and at the Charles University, Prague, Czech Republic. Barbara Zitová's research interests include geometric invariants, image enhancement, image registration, image fusion, and image processing in medical and in cultural heritage applications. She has authored or coauthored more than seventy research publications in these areas, including the monograph *Moments and Moment Invariants in Pattern Recognition* (Wiley, 2009) and tutorials at several major conferences. In 2003 Barbara Zitová received the Josef Hlavka Student Prize, in 2006 the Otto Wichterle Premium of the Czech Academy of Sciences for young scientists, and in 2010 she was awarded the prestigious SCOPUS 1000 Award for receiving more than 1000 citations of a single paper. Personal webpage: `http://zoi.utia.cas.cz/zitova`

Acknowledgements

First of all, we would like to express gratitude to our employer, the Institute of Information Theory and Automation (ÚTIA) and to its mother organization, the Czech Academy of Sciences, for creating an inspiring and friendly environment and for continuous support of our research. We also thank both universities where we teach, the Czech Technical University, Faculty of Nuclear Science and Physical Engineering, and the Charles University, Faculty of Mathematics and Physics, for their support and for administration of all our courses.

Many our colleagues, students, and friends contributed to the book in various ways. We are most grateful to all of them, especially to

Jiří Boldyš for his significant contribution to Chapters 4 and 6,
Bo Yang for his valuable contribution to Chapters 4 and 7,
Sajad Farokhi for participation in the experiments with Gaussian blur in Chapter 6,
Roxana Bujack for sharing her knowledge on vector field invariants,
Matteo Pedone for his contribution to blur-invariant image registration,
Jaroslav Kautsky for sharing his long-time experience with orthogonal polynomials,
Filip Šroubek for his valuable contribution to Chapter 6 and for providing test images and deconvolution codes,
Cyril Höschl IV for implementing the graph-based decomposition in Chapter 8,
Jitka Kostková for creating illustrations in Chapter 2,
Michal Šorel for his help with the optical flow experiment in Chapter 9,
Babak Mahdian for his contribution to the image forensics experiment in Chapter 9,
Michal Breznický for his comments on the blur invariants,
Stanislava Šimberová for providing the astronomical images,
Digitization Center of Cultural Heritage, Xanthi, Greece, for providing the 3D models of the ancient artifacts,
Alexander Kutka for his help with the focus measure experiment in Chapter 9,
Jan Kybic for providing us with his code for elastic matching, and
Jarmila Pánková for the graphical design of the book front cover.

Special thanks go to the staff of the publisher John Wiley & Sons Limited, who encouraged us to start writing the book and who provided us with help and assistance during the publication process. We would also like to thank the many students in our courses and our tutorial attendees who have asked numerous challenging and stimulating questions.

The work on this book was partially supported by the Czech Science Foundation under the projects No. GA13-29225S in 2014 and No. GA15-16928S in 2015–16.

Last but not least, we would like to thank our families. Their encouragement and patience was instrumental in completing this book.

Jan Flusser
Tomáš Suk
Barbara Zitová
Institute of Information Theory and Automation of the CAS,
Prague, Czech Republic, November 2015.

About the companion website

Don't forget to visit the companion website for this book:

http://www.wiley.com/go/flusser/2d3dimageanalysis

There you will find valuable material designed to enhance your learning, including:

- 300 slides
- Matlab codes

Scan this QR code to visit the companion website

1

Motivation

In our everyday life, each of us constantly receives, processes, and analyzes a huge amount of information of various kinds, significance and quality, and has to make decisions based on this analysis. More than 95% of information we perceive is of a visual character. An image is a very powerful information medium and communication tool capable of representing complex scenes and processes in a compact and efficient way. Thanks to this, images are not only primary sources of information, but are also used for communication among people and for interaction between humans and machines. Common digital images contain enormous amounts of information. An image you can take and send to your friends using a smart phone in a few seconds contains as much information as hundreds of text pages. This is why in many application areas, such as robot vision, surveillance, medicine, remote sensing, and astronomy, there is an urgent need for automatic and powerful image analysis methods.

1.1 Image analysis by computers

Image analysis in a broad sense is a many-step process the input of which is an image while the output is a final (usually symbolic) piece of information or a decision. Typical examples are localization of human faces in the scene and recognition (against a list or a database of persons) who is who, finding and recognizing road signs in the visual field of the driver, and identification of suspect tissues in a CT or MRI image. A general image analysis flowchart is shown in Figure 1.1 and an example what the respective steps look like in the car licence plate recognition is shown in Figure 1.2.

The first three steps of image analysis—image acquisition, preprocessing, and object segmentation/detection—are comprehensively covered in classical image processing textbooks [1–5], in recent specialized monographs [6–8] and in thousands of research papers. We very briefly recall them below, but we do not deal with them further in this book. The topic of this book falls into the fourth step, the feature design. The book is devoted to one particular family of features which are based on image moments. The last step, object classification, is shortly reviewed in Chapter 2, but its deeper description is beyond the scope of this book.

In *image acquisition*, the main theoretical questions are how to choose the sampling scheme, the sampling frequency and the number of the quantization levels such that the artifacts caused by aliasing, moire, and quantization noise do not degrade the image much while keeping the

2D and 3D Image Analysis by Moments, First Edition. Jan Flusser, Tomáš Suk and Barbara Zitová.

Companion website: www.wiley.com/go/flusser/2d3dimageanalysis

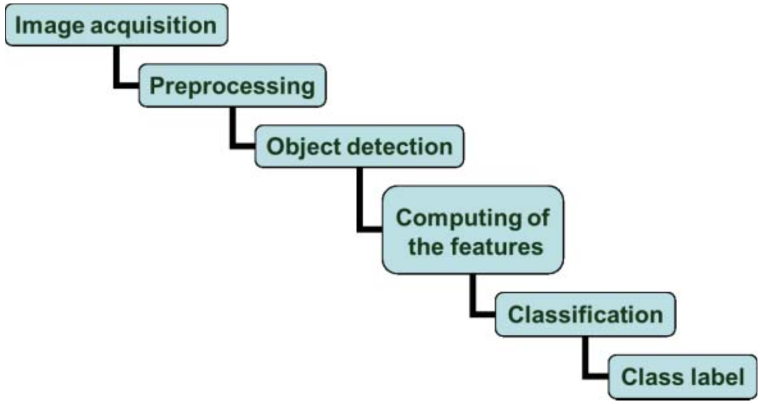


Figure 1.1 General image analysis flowchart

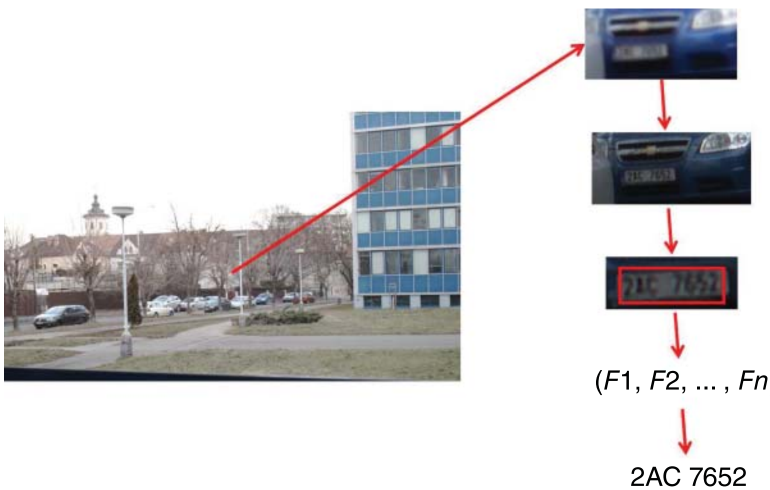


Figure 1.2 An example of the car licence plate recognition

image size reasonably low (there are of course also many technical questions about the appropriate choice of the camera and the spectral band, the objective, the memory card, the transmission line, the storage format, and so forth, which we do not discuss here).

Since real imaging systems as well as imaging conditions are usually imperfect, the acquired image represents only a degraded version of the original scene. Various kinds of degradations (geometric as well as graylevel/color) are introduced into the image during the acquisition process by such factors as imaging geometry, lens aberration, wrong focus, motion of the scene, systematic and random sensor errors, noise, etc. (see Figure 1.3 for the general scheme and an illustrative example). Removal or at least suppression of these degradations is a subject of *image preprocessing*. Historically, image preprocessing was one of the first topics systematically studied in digital image processing (already in the very first monograph [9] there was a chapter

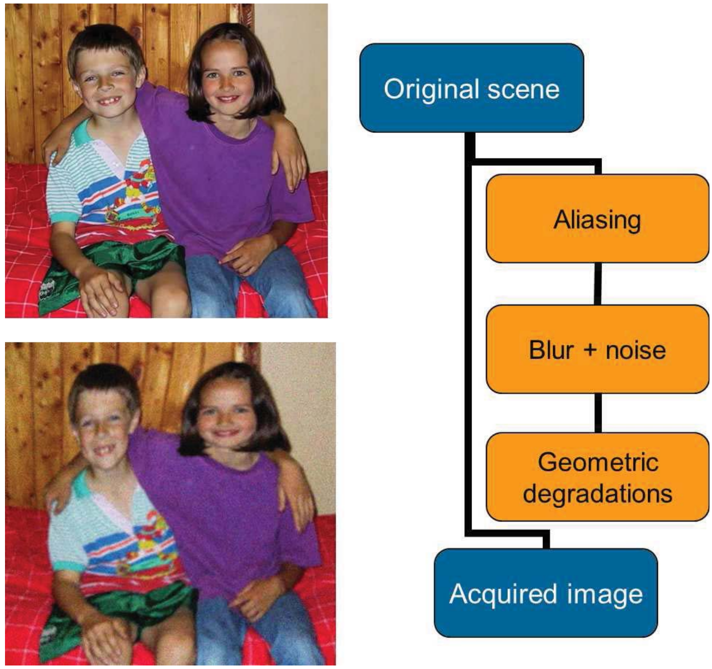


Figure 1.3 Image acquisition process with degradations

devoted to this topic) because even simple preprocessing methods were able to enhance the visual quality of the images and were feasible on old computers. The first two steps, image acquisition and preprocessing, are in the literature often categorized into *low-level processing*. The characteristic feature of the low-level methods is that both their input and output are digital images. On the contrary, in *high-level processing*, the input is a digital image (often an output of some preprocessing) while the output is a symbolic (i.e. high-level) information, such as the coordinates of the detected objects, the list of boundary pixels, etc.

Object detection is a typical example of high-level processing. The goal is to localize the objects of interest in the image and separate (segment) them from the other objects and from the background[1]. Hundreds of segmentation methods have been described in the literature. Some of them are universal, but most of them were designed for specific families of objects such as characters, logos, cars, faces, human silhouettes, roads, etc. A good basic survey of object detection and segmentation methods can be found in [5] and in the references therein.

Feature definition and computing is probably the most challenging part of image analysis. The features should provide an accurate and unambiguous quantitative description of the

[1] Formally, the terms such as “object”, “boundary”, “interior”, “background”, and others are subject of the discrete topology we are working in. Various discrete topologies differ from one another by the notion of neighboring pixels and hence by the notion of connectivity. However, here we use the above terms intuitively without specifying the particular topology.

objects[2]. The feature values are elements of the *feature space* which should be for the sake of efficient computation of low dimensionality. The design of the features is highly dependent on the type of objects, on the conditions under which the images have been acquired, on the type and the quality of preprocessing, and on the application area. There is no unique "optimal" solution.

Classification/recognition of the object is the last stage of the image analysis pipeline. It is entirely performed in the feature space. Each unknown object, now being represented by a point in the feature space, is classified as an element of a certain class. The classes can be specified in advance by their representative samples, which create a *training set*; in such a case we speak about *supervised classification*. Alternatively, if no training set is available, the classes are formed from the unknown objects based on their distribution in the feature space. This case is called *unsupervised classification* or *clustering*, and in visual object recognition this approach is rare. Unlike the feature design, the classification algorithms are application independent—they consider neither the nature of the original data nor the physical meaning of the features. Classification methods are comprehensively reviewed in the famous Duda-Hart-Stork book [10] and in the most recent monograph [11]. The use of the classification methods is not restricted to image analysis. We can find numerous applications in artificial intelligence, decision making, social sciences, statistical data analysis, and in many other areas beyond the scope of this book.

1.2 Humans, computers, and object recognition

Why is image analysis so easy and natural for human beings and so difficult for machines? There are several reasons for this performance difference. Incorporating lifetime experience, applying information fusion, the ability of contextual perception, and perceptual robustness are the key elements. Current research directions are often motivated by these factors, but up to now the results have been achieved only in laboratory conditions and have not reached the quality of human beings yet. Probably the most serious difference is that humans incorporate their lifetime knowledge into the recognition process and do so in a very efficient way. If you see an object you have already met (even a long time ago), you are able to "retrieve" this information from your brain almost immediately (it appears that the "retrieval time" does not depend on the number of the objects stored in the brain) and use it as a hint for the classification. Your brain is also able to supply the missing information if only a part of the object is visible. This ability helps us a lot when recognizing partially occluded objects. Computers could in principle do the same, but it would require massive learning on large-scale databases, which would lead not only to a time-consuming learning stage but also to relatively slow search/retrieval of the learned objects. To overcome that, several projects have been started recently which are trying to substitute human life experience by shared knowledge available at search engines (see [12] for example) but this research is still at the beginning stages.

Our ability to combine supplementary information of other kinds in parallel with vision makes human perception efficient, too. We also have other senses (hearing, smell, taste, and

[2] There exists an alternative way of object description by its structure. The respective recognition methods are then called structural or syntactic ones. They use the apparatus of formal language theory. Each object class is given by a language, and each object is a word. An object is assigned to the class if its word belongs to the class language. This approach is useful if the structural information is more appropriate than a quantitative description.