Third Edition

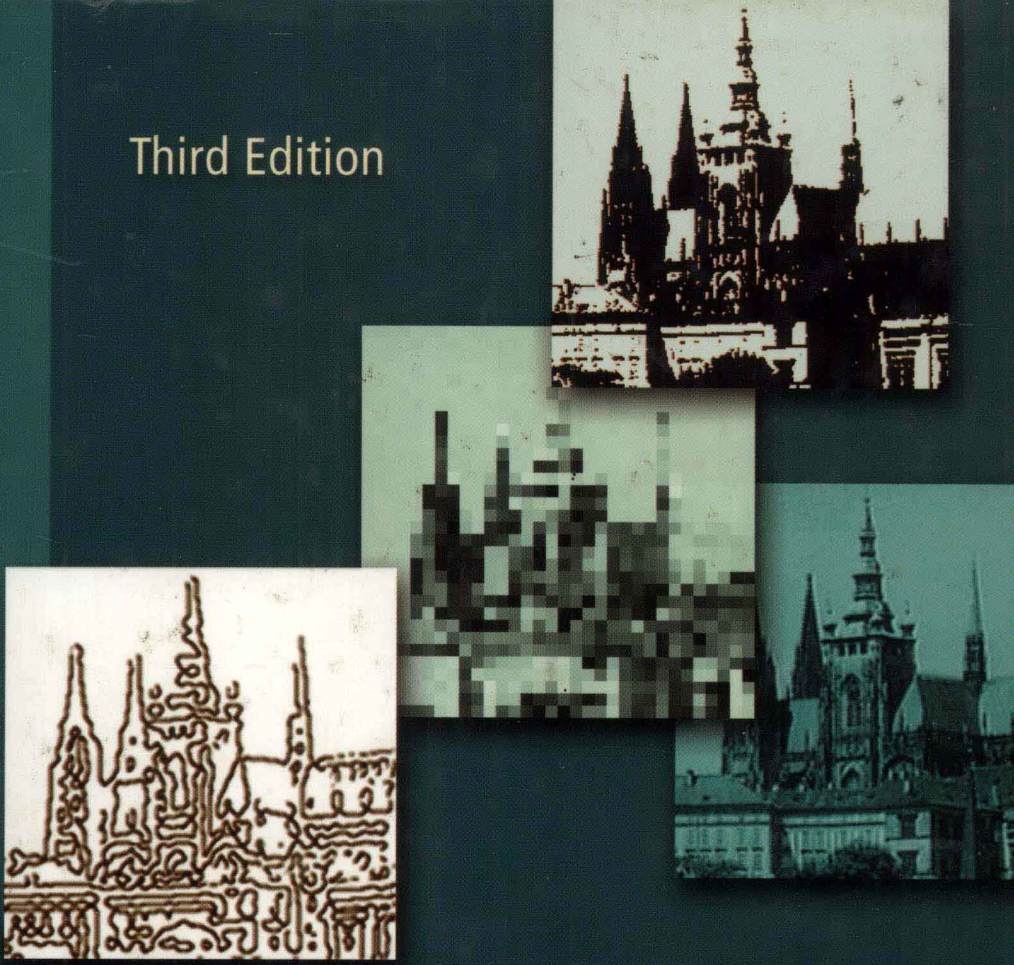


Image Processing, Analysis, and Machine Vision

Milan Sonka

Vaclav Hlavac

Roger Boyle

Image Processing, Analysis, and Machine Vision

Third Edition

Milan Sonka
The University of Iowa, Iowa City

Vaclav Hlavac
Czech Technical University, Prague

Roger Boyle
University of Leeds, Leeds

THOMSON

Australia Canada Mexico Singapore Spain United Kingdom United States

THOMSON

Image Processing, Analysis, and Machine Vision, Third Edition

by Milan Sonka, Vaclav Hlavac, Roger Boyle

General Manager:
Chris Carson

Developmental Editor:
Hilda Gowans

Permissions Coordinator:
Vicki Gould

Production Manager:
Renate McCloy

Interior Design:
Vit Zyka

Cover Design:
Andrew Adams

Compositor:
Vit Zyka

Printer:
Thomson/West

Printed and bound in the United States of America
1 2 3 4 10 09 08 07

For more information contact Thomson Learning, 1120 Birchmount Road, Toronto, Ontario, M1K 5G4. Or you can visit our Internet site at http://www.thomsonlearning.com

Library of Congress Control Number 2007921908

ISBN: 10: 0-495-08252-X
ISBN: 13: 978-0-495-08252-1

North America
Thomson Learning
1120 Birchmount Road
Toronto, Ontario M1K 5G4
Canada

Asia
Thomson Learning
5 Shenton Way #01-01
UIC Building
Singapore 068808

Australia/New Zealand
Thomson Learning
102 Dodds Street
Southbank, Victoria
Australia 3006

Europe/Middle East/Africa
Thomson Learning
High Holborn House
50/51 Bedford Row
London WCIR 4LR
United Kingdom

Latin America
Thomson Learning
Seneca, 53
Colonia Polanco
11560 Mexico D.F.
Mexico

Spain
Paraninfo
Calle/Magallanes, 25
28015 Madrid, Spain

Abbreviations

1D	one dimension(al)
2D, 3D, ...	two dimension(al), three dimension(al), ...
AAM	active appearance model
AI	artificial intelligence
ASM	active shape model
B-rep	boundary representation
BBN	Bayesian belief network
CAD	computer-aided design
CCD	charge-coupled device
CONDENSATION	CONditional DENSity propagATION
CSG	constructive solid geometry
CT	computed tomography
dB	decibel, 20 times the decimal logarithm of a ratio
DCT	discrete cosine transform
dof	degrees of freedom
DWF	discrete wavelet frame
ECG	electro-cardiogram
EEG	electro-encephalogram
EM	expectation-maximization
FFT	fast Fourier transform
FOE	focus of expansion
GA	genetic algorithm
GB	Giga byte = 2^{30} bytes = 1,073,741,824 bytes
GMM	Gaussian mixture model
GVF	gradient vector flow
HMM	hidden Markov model
ICA	independent component analysis
IHS	intensity, hue, saturation
JPEG	Joint Photographic Experts Group
Kb	Kilo bit = 2^{10} bits = 1,024 bits
KB	Kilo byte = 2^{10} bytes = 1,024 bytes
Mb	Mega bit = 2^{20} bits = 1,048,576 bits
MB	Mega byte = 2^{20} bytes = 1,048,576 bytes
MDL	minimum description length
MR	magnetic resonance

MRI	magnetic resonance imaging
μs	microsecond
ms	millisecond
OCR	optical character recognition
OS	order statistics
PCA	principal component analysis
p.d.f.	probability density function
PDM	point distribution model
PET	positron emission tomography
PMF	Pollard-Mayhew-Frisby (correspondence algorithm)
RANSAC	RANdom SAmple Consensus
RGB	red, green, blue
RCT	reversible component transform
SNR	signal-to-noise ratio
SVD	singular value decomposition
TV	television

$\arg(x, y)$	angle (in radians) from x axis to the point (x, y)
$\underset{i}{\operatorname{argmax}}\big(\operatorname{expr}(i)\big)$	the value of i that causes $\operatorname{expr}(i)$ to be maximal
$\underset{i}{\operatorname{argmin}}\big(\operatorname{expr}(i)\big)$	the value of i that causes $\operatorname{expr}(i)$ to be minimal
div	integer division or divergence
mod	remainder after integer division
$\operatorname{round}(x)$	largest integer which is not bigger than $x + 0.5$
$\emptyset$	empty set
A^c	complement of set A
$A \subset B$, $B \supset A$	set A is included in set B
$A \cap B$	intersection between sets A and B
$A \cup B$	union of sets A and B
$A \mid B$	difference between sets A and B
$\mathbf{A}$	(uppercase bold) matrices
$\mathbf{x}$	(lowercase bold) vectors
$\lvert\mathbf{x}\rvert$	magnitude (or modulus) of vector $\mathbf{x}$
$\mathbf{x} \cdot \mathbf{y}$	scalar product between vectors $\mathbf{x}$ and $\mathbf{y}$
$\tilde{x}$	estimate of the value x
$\lvert x\rvert$	absolute value of a scalar
$\delta(x)$	Dirac function
Δx	small finite interval of x, difference
$\partial f / \partial x$	partial derivative of the function f with respect to x
$\nabla \mathbf{f}$, $\operatorname{grad} \mathbf{f}$	gradient of $\mathbf{f}$
$\nabla^2 \mathbf{f}$	Laplace operator applied to $\mathbf{f}$
$f * g$	convolution between functions f and g
$F .* G$	element-by-element of matrices F, G
D_E	Euclidean distance
D_4	city block distance
D_8	chessboard distance
F^*	complex conjugate of the complex function F
$\operatorname{rank}(A)$	rank of a matrix A
T^*	transformation dual to transformation T, also complex conjugate of T
$\mathcal{E}$	mean value operator
$\mathcal{L}$	linear operator
$\mathcal{O}$	origin of the coordinate system

Symbols

$\#$	number of (e.g., pixels)
$\breve{B}$	point set symmetrical to point set B
$\oplus$	morphological dilation
$\ominus$	morphological erosion
$\circ$	morphological opening
$\bullet$	morphological closing
$\otimes$	morphological hit-or-miss transformation
$\oslash$	morphological thinning
$\odot$	morphological thickening
$\wedge$	logical and
$\vee$	logical or
trace	sum of elements on the matrix main diagonal
cov	covariance matrix
sec	secant, $\sec\alpha = 1/\cos\alpha$

Preface

Image processing, analysis, and machine vision represent an exciting and dynamic part of cognitive and computer science. Following an explosion of interest during the 1970s and the 1980s, the last three decades were characterized by a maturing of the field and significant growth of active applications; remote sensing, technical diagnostics, autonomous vehicle guidance, biomedical imaging (2D, 3D, and 4D) and automatic surveillance are the most rapidly developing areas. This progress can be seen in an increasing number of software and hardware products on the market—as a single example of many, the omnipresence of consumer-level digital cameras is striking. Reflecting this continuing development, the number of digital image processing and machine vision courses offered at universities worldwide is increasing rapidly.

There are many texts available in the areas we cover—many of them are referenced somewhere in this book. The subject suffers, however, from a shortage of texts which are 'complete' in the sense that they are accessible to the novice, of use to the educated, and up to date. Here we present the third edition of a text first published in 1993. We include many of the very rapid developments that have taken and are still taking place, which quickly age some of the very good textbooks produced in the recent past.

The target audience spans the range from the undergraduate with negligible experience in the area through to the Master's and research student seeking an advanced springboard in a particular topic. Every section of this text has been updated since the second version (particularly with respect to references). Chapters 2 and 3 were reorganized and enhanced to better present a broad yet not overwhelming foundation, which is used throughout the book. While the second edition published in 1998 provided a comprehensive treatment of 2D image processing and analysis, analysis of volumetric and thus inherently 3D image data has become a necessity. To keep up with the rapidly advancing field, a brand new Chapter 7 covers image segmentation methods and approaches with 3D (or higher dimension) capabilities such as mean shift segmentation, gradient vector flow snakes, level sets, direct graph cut segmentation, and optimal single and multiple surface detection. As a result, the book now has two chapters devoted to segmentation, clearly reflecting the importance of this area.

Many other new topics were added throughout the book. Wholly new sections are presented on: support vector classifiers; boosting approaches to pattern recognition; model fitting via random sample consensus; active appearance models; object detection using a boosted cascade of classifiers; coupled hidden Markov models; Bayesian belief networks; Gaussian mixture models; expectation-maximization; JPEG 2000 image compression; multiscale wavelet texture description; detection of specific motion patterns; background modeling for video tracking; kernel-based tracking; and particle filters for motion modeling. All in all, about 25% of this third edition consists of a newly written material presenting state-of-the-art methods and techniques that have already proven their importance in the field.

A carefully prepared set of exercises is of great use in assisting the reader in acquiring practical understanding. We have chosen to provide a stand-alone Matlab companion book [Svoboda et al., 2008] with an accompanying web page (http://www.engineering.thomsonlearning.com) rather than include exercises directly. This companion text contains short-answer questions and problems of varying difficulty, frequently requiring practical usage of computer tools and/or development of application programs. It concentrates on algorithmic aspects; the Matlab programming environment was chosen for this purpose as it allows quick insights and fast prototyping. Many of the algorithms presented here have their counterparts in the exercise companion book. Source code of programs and data used in examples are available on the web page, which will undoubtedly grow with contributions from users.

The exercise companion book is intended both for students and teachers. Students can learn from short-answer questions, formulated problems and from programmed examples. They are also likely to use pieces of the code provided in their own programs. Teachers will find the book useful for preparing examples in lectures, and assignments for their students. Our experience is that such material allows the teacher to concentrate a course on ideas and their demonstrations rather than on programming of simple algorithms. Solutions to problems will be available for teachers in the password protected part of the web page. The web page will also carry the official list of errata. The reader is encouraged to check this resource frequently.

This book reflects the authors' experience in teaching one- and two-semester undergraduate and graduate courses in Digital Image Processing, Digital Image Analysis, Image Understanding, Medical Imaging, Machine Vision, Pattern Recognition, and Intelligent Robotics at their respective institutions. We hope that this combined experience will give a thorough grounding to the beginner and provide material that is advanced enough to allow the more mature student to understand fully the relevant areas of the subject. We acknowledge that in a very short time the more active areas will have moved beyond this text.

This book could have been arranged in many ways. It begins with low-level processing and works its way up to higher levels of image interpretation; the authors have chosen this framework because they believe that image understanding originates from a common database of information. The book is formally divided into 16 chapters, beginning with low-level processing and working toward higher-level image representation, although this structure will be less apparent after Chapter 12, when we present mathematical morphology, image compression, texture, and motion analysis which are very useful but often special-purpose approaches that may not always be included in the processing chain.

Decimal section numbering is used, and equations and figures are numbered within each chapter. Each chapter is supported by an extensive list of references and exercises [Svoboda et al., 2008]. A selection of algorithms is summarized formally in a manner that should aid implementation—not all the algorithms discussed are presented in this way (this might have doubled the length of the book); we have chosen what we regard as the key, or most useful or illustrative, examples for this treatment. Each chapter further includes a concise Summary section.

Chapters present material from an introductory level through to an overview of current work; as such, it is unlikely that the beginner will, at the first reading, expect to absorb all of a given topic. Often it has been necessary to make reference to material in later chapters and sections, but when this is done an understanding of material in hand will not depend on an understanding of that which comes later. It is expected that the more advanced student will use the book as a reference text and signpost to current activity in the field—we believe at the time of going to press that the reference list is full in its indication of current directions, but record here our apologies to any work we have overlooked. The serious reader will note that the reference list contains citations of both the classic material that has survived the test of time as well as references that are very recent and represent what the authors consider promising new directions. Of course, before long, more relevant work will have been published that is not listed here.

This is a long book and therefore contains material sufficient for much more than one course. Clearly, there are many ways of using it, but for guidance we suggest an ordering that would generate five distinct modules:

Digital Image Processing I, an undergraduate course.

Digital Image Processing II, an undergraduate/graduate course, for which Digital Image Processing I may be regarded as prerequisite.

Computer Vision I, an undergraduate/graduate course, for which Digital Image Processing I may be regarded as prerequisite.

Computer Vision II, a graduate course, for which Computer Vision I may be regarded as prerequisite.

Image Analysis and Understanding, a graduate course, for which Computer Vision I may be regarded as prerequisite.

The important parts of a course, and necessary prerequisites, will naturally be specified locally; a suggestion for partitioning the contents follows this Preface.

Assignments should wherever possible make use of existing software; it is our experience that courses of this nature should not be seen as 'programming courses', but it is the case that the more direct practical experience the students have of the material discussed, the better is their understanding. Since the first edition was published, an explosion of web-based material has become available, permitting many of the exercises we present to be conducted without the necessity of implementing from scratch—we do not present explicit pointers to Web material, since they evolve so quickly; however, pointers to specific support materials for this book and others may be located via the designated book web page, http://www.engineering.thomsonlearning.com.

The book has been prepared using the LaTeX text processing system. Its completion would have been impossible without extensive usage of the Internet computer network and electronic mail. We would like to acknowledge the University of Iowa, the Czech Technical University, and the School of Computing at the University of Leeds for providing the environment in which this book was born and re-born.

Milan Sonka is Professor of Electrical & Computer Engineering, Ophthalmology & Visual Sciences, and Radiation Oncology at the University of Iowa, Iowa City, Iowa, USA. His research interests include medical image analysis, computer-aided diagnosis, and machine vision. Václav Hlaváč is Professor of Cybernetics at the Czech Technical University, Prague. His research interests are knowledge-based image analysis, 3D model-based vision and relations between statistical and structural pattern recognition. Roger Boyle is Professor of Computing and Head of the School of Computing at the

University of Leeds, England, where his research interests are in low-level vision and pattern recognition. The first two authors first worked together as faculty members of the Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic from 1983 to 1990, and have been co-operating with the third since 1991. Interestingly enough, Boyle and Sonka, having collaborated for 15 years over the Internet, had their first face-to-face meeting only very recently, in the fall of 2006. We are happy to report that this meeting did not spoil their working relationship.

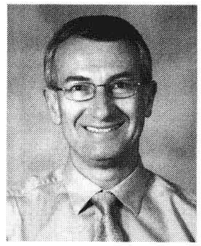
All authors have contributed throughout—the ordering on the cover corresponds to the weight of individual contribution. Any errors of fact are the joint responsibility of all.

For the first time, the book was not typeset by the authors. Professional help from Dr. Vít Zýka has resulted in a visually pleasing, clean, and well formatted manuscript. Despite long hours, and periods of the—as our wives call it—author's syndrome of limited communication with the outside world, we are already looking forward to future editions in which all omissions, typos, and errors get corrected.

Milan Sonka

The University of Iowa
Iowa City, Iowa, USA

milan-sonka@uiowa.edu
http://www.engineering.uiowa.edu/~sonka



Václav Hlaváč

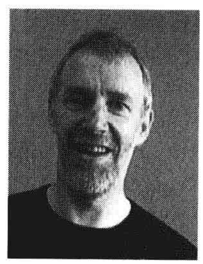
Czech Technical University
Prague, Czech Republic

hlavac@cmp.felk.cvut.cz
http://cmp.felk.cvut.cz/~hlavac

Roger Boyle

University of Leeds
Leeds, United Kingdom

roger@comp.leeds.ac.uk
http://www.comp.leeds.ac.uk/roger/



References

Svoboda T., Kybic J., and Hlavac V. *Image Processing, Analysis, and Machine Vision: A MATLAB Companion.* Thomson Engineering, 2008.

Possible course outlines

Here, one possible ordering of the material covered in the five courses proposed in the Preface is given. This should not, of course, be considered the only option—on the contrary, the possibilities for organizing Image Processing and Analysis courses are practically endless. Therefore, what follows shall only be regarded as suggestions, and instructors shall tailor content to fit the assumed knowledge, abilities, and needs of the students enrolled.

Figure 1 shows course pre-requisite dependencies of the proposed ordering. Figure 2 shows the mapping between the proposed course outlines and the material covered in the individual chapters and sections.

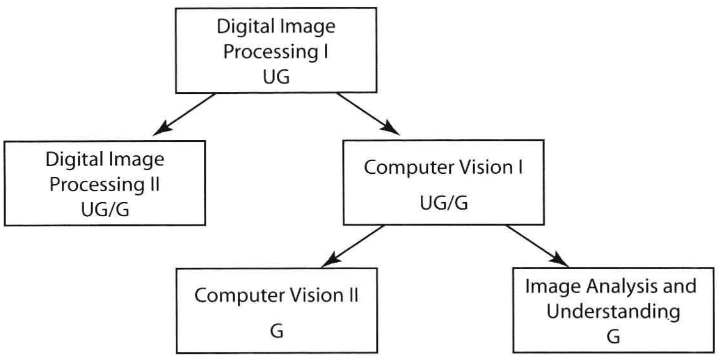


Figure 1: Pre-requisite dependencies of the proposed five courses. UG = undergraduate course, G = graduate course.

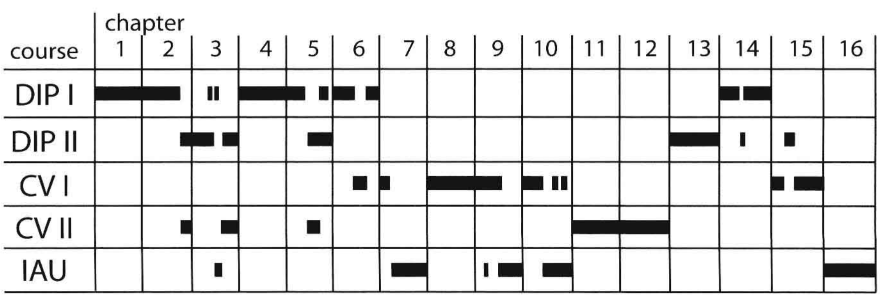


Figure 2: Mapping between the proposed course outlines and material covered in individual chapters and sections. See course outlines for details.

Digital Image Processing I (DIP I)

An undergraduate course.

1 Introduction
2 The image, its representation and properties
- 2.1 Image representations
- 2.2 Image digitization
- 2.3 Digital image properties

4 Data structures for image analysis
5 Image pre-processing
- 5.1 Pixel brightness transformations
- 5.2 Geometric transformations
- 5.3 Local pre-processing (except 5.3.6–5.3.7, 5.3.9, 5.3.10, limited coverage of 5.3.4, 5.3.5, 5.3.5)
- 5.4 Image restoration (except 5.4.3)

6 Segmentation
- 6.1 Thresholding (except 6.1.3)
- 6.2 Edge-based segmentation (except 6.2.4, 6.2.5, 6.2.7, 6.2.8)
- 6.3 Region growing segmentation (except 6.3.4)
- 6.4 Matching
- 6.5 Evaluation issues in segmentation

3 The image, its mathematical and physical background
- 3.2 Linear integral transforms (3.2.1–3.2.4, 3.2.6 only)

14 Image data compression (except wavelet compression)
Practical image processing projects

Digital Image Processing II (DIP II)

An undergraduate/graduate course, for which Digital Image Processing I may be regarded as prerequisite.

1 Introduction (brief review)
2 The image, its representation and properties
- 2.4 Color images
- 2.5 Cameras

3 The image, its mathematical and physical background (except 3.2.8–3.2.10)
5 Image pre-processing
- 5.3.4 Scale in image processing
- 5.3.5 Canny edge detection
- 5.3.6 Parametric edge models
- 5.3.7 Edges in multi-spectral images
- 5.3.8 Pre-processing in frequency domain
- 5.3.9 Line detection
- 5.3.10 Corner detection
- 5.3.11 Maximally stable extremal regions
- 5.4 Image restoration

14 Image compression (14.2, 14.9 only)
15 Texture (15.1.6, 15.1.7 only)

13 Mathematical morphology
Practical image processing projects

Computer Vision I (CV I)

An undergraduate/graduate course, for which Digital Image Processing I may be regarded as prerequisite.

1 Introduction (brief review)
2 The image, its representation and properties (brief review)
6 Segmentation I
 6.2.4 Border detection as graph searching
 6.2.5 Border detection as dynamic programming
 6.3.4 Watershed segmentation
7 Segmentation II
 7.1 Mean shift segmentation
 7.2 Active contour models
8 Shape representation and description
9 Object recognition
 9.1 Knowledge representation
 9.2 Statistical pattern recognition (except 9.2.4)
 9.3 Neural networks
 9.4 Syntactic pattern recognition
10 Image understanding (except 10.4, 10.6, 10.9, 10.10)
15 Texture (except 15.1.6, 15.1.7)
Practical computer vision projects

Computer Vision II (CV II)

A graduate course, for which Computer Vision I may be regarded as prerequisite.

2 The image, its representation and properties
 2.4 Color images
 2.5 Cameras
3 The image, its mathematical and physical background
 3.4 Image formation physics
5 Image pre-processing
 5.3.4 Scale in image processing
 5.3.5 Canny edge detection
 5.3.6 Parametric edge models
 5.3.7 Edges in multi-spectral images
 5.3.9 Line detection
 5.3.10 Corner detection
 5.3.11 Maximally stable extremal regions
11 3D Vision, geometry and radiometry
12 Use of 3D vision
Practical 3D vision projects

Image Analysis and Understanding (IAU)

A graduate course, for which Computer Vision I may be regarded as prerequisite.

7 Segmentation (except 7.1, 7.2)
9 Object recognition
- 9.2.4 Support vector machines
- 9.5 Recognition as graph matching
- 9.6 Optimization techniques in recognition
- 9.7 Fuzzy systems
- 9.8 Boosting in pattern recognition

3 The image, its mathematical and physical background (3.2.8–3.2.10 only)
10 Image understanding
- 10.4 Active appearance models
- 10.6 Boosted cascade of classifiers
- 10.9 Hidden Markov models
- 10.10 Gaussian mixture models and expectation maximization

16 Motion analysis
Practical image understanding projects

Contents