Undergraduate Topics in Computer Science

Undergraduate Topics in Computer Science (UTiCS) delivers high-quality instructional content for undergraduates studying in all areas of computing and information science. From core foundational and theoretical material to final-year topics and applications, UTiCS books take a fresh, concise, and modern approach and are ideal for self-study or for a one- or two-semester course. The texts are all authored by established experts in their fields, reviewed by an international advisory board, and contain numerous examples and problems. Many include fully worked solutions.

For further volumes:
www.springer.com/series/7592

Reinhard Klette

Concise Computer Vision

An Introduction into Theory and Algorithms

Springer

Reinhard Klette
Computer Science Department
University of Auckland
Auckland, New Zealand

ISSN 1863-7310 ISSN 2197-1781 (electronic)
Undergraduate Topics in Computer Science
ISBN 978-1-4471-6319-0 ISBN 978-1-4471-6320-6 (eBook)
DOI 10.1007/978-1-4471-6320-6
Springer London Heidelberg New York Dordrecht

Library of Congress Control Number: 2013958392

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Dedicated to all who have dreams

Computer vision may count the trees, estimate the distance to the islands, but it cannot detect the fantasies the people might have had who visited this bay

Preface

This is a textbook for a third- or fourth-year undergraduate course on Computer vision, which is a discipline in science and engineering.

Subject Area of the Book Computer Vision aims at using cameras for analysing or understanding scenes in the real world. This discipline studies methodological and algorithmic problems as well as topics related to the implementation of designed solutions.

In computer vision we may want to know how far away a building is to a camera, whether a vehicle drives in the middle of its lane, how many people are in a scene, or we even want to recognize a particular person—all to be answered based on recorded images or videos. Areas of application have expanded recently due to a solid progress in computer vision. There are significant advances in camera and computing technologies, but also in theoretical foundations of computer vision methodologies.

In recent years, computer vision became a key technology in many fields. For modern consumer products, see, for example apps for mobile phones, driver-assistance for cars, or user interaction with computer games. In industrial automation, computer vision is routinely used for quality or process control. There are significant contributions for the movie industry (e.g. the use of avatars or the creation of virtual worlds based on recorded images, the enhancement of historic video data, or high-quality presentations of movies). This is just mentioning a few application areas, which all come with particular image or video data, and particular needs to process or analyse those data.

Features of the Book This text book provides a general introduction into basics of computer vision, as potentially of use for many diverse areas of applications. Mathematical subjects play an important role, and the book also discusses algorithms. The book is not addressing particular applications.

Inserts (grey boxes) in the book provide historic context information, references or sources for presented material, and particular hints on mathematical subjects discussed first time at a given location. They are *additional* readings to the baseline material provided.

The book is *not* a guide on current research in computer vision, and it provides only a *very few* references; the reader can locate more easily on the net by searching for keywords of interest. The field of computer vision is actually so vivid, with countless references, such that any attempt would fail to insert in the given limited space a reasonable collection of references. But here is one hint at least: visit homepages.inf.ed.ac.uk/rbf/CVonline/ for a web-based introduction into topics in computer vision.

Target Audiences This text book provides material for an introductory course at third- or fourth-year level in an Engineering or Science undergraduate programme. Having some prior knowledge in image processing, image analysis, or computer graphics is of benefit, but the first two chapters of this text book also provide a first-time introduction into computational imaging.

Previous Uses of the Material Parts of the presented materials have been used in my lectures in the Mechatronics and Computer Science programmes at The University of Auckland, New Zealand, at CIMAT Guanajuato, Mexico, at Freiburg and Göttingen University, Germany, at the Technical University Cordoba, Argentina, at the Taiwan National Normal University, Taiwan, and at Wuhan University, China.

The presented material also benefits from four earlier book publications, [R. Klette and P. Zamperoni. Handbook of Image Processing Operators. Wiley, Chichester, 1996], [R. Klette, K. Schlüns, and A. Koschan. Computer Vision. Springer, Singapore, 1998], [R. Klette and A. Rosenfeld. Digital Geometry. Morgan Kaufmann, San Francisco, 2004], and [F. Huang, R. Klette, and K. Scheibe. Panoramic Imaging. Wiley, West Sussex, 2008].

The first two of those four books accompanied computer vision lectures of the author in Germany and New Zealand in the 1990s and early 2000s, and the third one also more recent lectures.

Notes to the Instructor and Suggested Uses The book contains more material than what can be covered in a one-semester course. An instructor should select according to given context such as prior knowledge of students and research focus in subsequent courses.

Each chapter ends with some exercises, including programming exercises. The book does not favour any particular implementation environment. Using procedures from systems such as `OpenCV` will typically simplify the solution. Programming exercises are intentionally formulated in a way to offer students a wide range of options for answering them. For example, for Exercise 2.5 in Chap. 2, you can use Java applets to visualize the results (but the text does not ask for it), you can use small- or large-sized images (the text does not specify it), and you can limit cursor movement to a central part of the input image such that the 11×11 square around location p is always completely contained in your image (or you can also cover special cases when moving the cursor also closer to the image border). As a result, every student should come up with her/his individual solution to programming exercises, and creativity in the designed solution should also be honoured.

Supplemental Resources The book is accompanied by supplemental material (data, sources, examples, presentations) on a website. See www.cs.auckland.ac.nz/~rklette/Books/K2014/.

Acknowledgements In alphabetical order of surnames, I am thanking the following colleagues, former or current students, and friends (if I am just mentioning a figure, then I am actually thanking for joint work or contacts about a subject related to that figure):

A-Kn *Ali Al-Sarraf* (Fig. 2.32), *Hernan Badino* (Fig. 9.25), *Anko Börner* (various comments on drafts of the book, and also contributions to Sect. 5.4.2), *Hugo Carlos* (support while writing the book at CIMAT), *Diego Caudillo* (Figs. 1.9, 5.28, and 5.29), *Gilberto Chávez* (Figs. 3.39 and 5.36, top row), *Chia-Yen Chen* (Figs. 6.21 and 7.25), *Kaihua Chen* (Fig. 3.33), *Ting-Yen Chen* (Fig. 5.35, contributions to Sect. 2.4, to Chap. 5, and provision of sources), *Eduardo Destefanis* (contribution to Example 9.1 and Fig. 9.5), *Uwe Franke* (Figs. 3.36, 6.3, and bottom, right, in 9.23), *Stefan Gehrig* (comments on stereo analysis parts and Fig. 9.25), *Roberto Guzmán* (Fig. 5.36, bottom row), *Wang Han* (having his students involved in checking a draft of the book), *Ralf Haeusler* (contributions to Sect. 8.1.5), *Gabriel Hartmann* (Fig. 9.24), *Simon Hermann* (contributions to Sects. 5.4.2 and 8.1.2, Figs. 4.16 and 7.5), *Václav Hlaváč* (suggestions for improving the contents of Chaps. 1 and 2), *Heiko Hirschmüller* (Fig. 7.1), *Wolfgang Huber* (Fig. 4.12, bottom, right), *Fay Huang* (contributions to Chap. 6, in particular to Sect. 6.1.4), *Ruyi Jiang* (contributions to Sect. 9.3.3), *Waqar Khan* (Fig. 7.17), *Ron Kimmel* (presentation suggestions on local operators and optic flow—which I need to keep mainly as a project for a future revision of the text), *Karsten Knoeppel* (contributions to Sect. 9.3.4),

Ko-Sc *Andreas Koschan* (comments on various parts of the book and Fig. 7.18, right), *Vladimir Kovalevsky* (Fig. 2.15), *Peter Kovesi* (contributions to Chaps. 1 and 2 regarding phase congruency, including the permission to reproduce figures), *Walter Kropatsch* (suggestions to Chaps. 2 and 3), *Richard Lewis-Shell* (Fig. 4.12, bottom, left), *Fajie Li* (Exercise 5.9), *Juan Lin* (contributions to Sect. 10.3), *Yizhe Lin* (Fig. 6.19), *Dongwei Liu* (Fig. 2.16), *Yan Liu* (permission to publish Fig. 1.6), *Rocío Lizárraga* (permission to publish Fig. 5.2, bottom row), *Peter Meer* (comments on Sect. 2.4.2), *James Milburn* (contributions to Sect. 4.4). *Pedro Real* (comments on geometric and topologic subjects), *Mahdi Rezaei* (contributions to face detection in Chap. 10, including text and figures, and Exercise 10.2), *Bodo Rosenhahn* (Fig. 7.9, right), *John Rugis* (definition of similarity curvature and Exercises 7.2 and 7.6), *James Russell* (contributions to Sect. 5.1.1), *Jorge Sanchez* (contribution to Example 9.1, Figs. 9.1, right, and 9.5), *Konstantin Schauwecker* (comments on feature detectors and RANSAC plane detection, Figs. 6.10, right, 7.19, 9.9, and 2.23), *Karsten Scheibe* (contributions to Chap. 6, in particular to Sect. 6.1.4), and Fig. 7.1), *Karsten Schlüns* (contributions to Sect. 7.4),

Sh-Z *Bok-Suk Shin* (Latex editing suggestions, comments on various parts of the book, contributions to Sects. 3.4.1 and 5.1.1, and Fig. 9.23 with related comments), *Eric Song* (Fig. 5.6, left), *Zijiang Song* (contributions to Chap. 9, in particular to Sect. 9.2.4), *Kathrin Spiller* (contribution to 3D case in Sect. 7.2.2), *Junli Tao* (contributions to pedestrian detection in Chap. 10, including text and figures and Exercise 10.1, and comments about the structure of this chapter), *Akihiko Torii* (contributions to Sect. 6.1.4), *Johan VanHorebeek* (comments on Chap. 10), *Tobi Vaudrey* (contributions to Sect. 2.3.2 and Fig. 4.18, contributions to Sect. 9.3.4, and Exercise 9.6), *Mou Wei* (comments on Chap. 4), *Shou-Kang Wei* (joint work on subjects related to Sect. 6.1.4), *Tiangong Wei* (contributions to Sect. 7.4.3), *Jürgen Wiest* (Fig. 9.1, left), *Yihui Zheng* (contributions to Sect. 5.1.1), *Zezhong Xu* (contributions to Sect. 3.4.1 and Fig. 3.40), *Shenghai Yuan* (comments on Sects. 3.3.1 and 3.3.2), *Qi Zang* (Exercise 5.5, and Figs. 2.21, 5.37, and 10.1), *Yi Zeng* (Fig. 9.15), and *Joviša Žunić* (contributions to Sect. 3.3.2).

The author is, in particular, indebted to *Sandino Morales* (D.F., Mexico) for implementing and testing algorithms, providing many figures, contributions to Chaps. 4 and 8, and for numerous comments about various parts of the book, to *Władysław Skarbek* (Warsaw, Poland) for manifold suggestions for improving the contents, and for contributing Exercises 1.9, 2.10, 2.11, 3.12, 4.11, 5.7, 5.8, and 6.10, and to *Garry Tee* (Auckland, New Zealand) for careful reading, commenting, for parts of Insert 5.9, the footnote on p. 402, and many more valuable hints.

I thank my wife, *Gisela Klette*, for authoring Sect. 3.2.4 about the Euclidean distance transform and critical views on structure and details of the book while the book was written at CIMAT Guanajuato between mid July to beginning of November 2013 during a sabbatical leave from The University of Auckland, New Zealand.

Guanajuato, Mexico Reinhard Klette
3 November 2013

Contents

Symbols

$\lvert S\rvert$	Cardinality of a set S
$\lVert\mathbf{a}\rVert_1$	L_1 norm
$\lVert\mathbf{a}\rVert_2$	L_2 norm
$\wedge$	Logical 'and'
$\vee$	Logical 'or'
$\cap$	Intersection of sets
$\cup$	Union of sets
$\square$	End of proof
a, b, c	Real numbers
A	Adjacency set
$\mathscr{A}(\cdot)$	Area of a measurable set (as a function)
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	Matrices
α, β, γ	Angles
b	Base distance of a stereo camera system
$\mathbb{C}$	Set of complex numbers $a + i \cdot b$, with $i = \sqrt{-1}$ and $a, b \in \mathbb{R}$
d	Disparity
d_1	L_1 metric
d_2	L_2 metric, also known as the Euclidean metric
e	Real constant $e = \exp(1) \approx 2.7182818284$
ε	Real number greater than zero
f	Focal length
f, g, h	Functions
$G_{\max}$	Maximum grey level in an image
γ	Curve in a Euclidean space (e.g. a straight line, polyline, or smooth curve)
$\mathbf{H}$	Hessian matrix
i, j, k, l, m, n	Natural numbers; pixel coordinates (i, j) in a window
$I, I(.,.,t)$	Image, frame of a sequence, frame at time t
L	Length (as a real number)

$\mathscr{L}(\cdot)$	Length of a rectifiable curve (as a function)
λ	Real number; default: between 0 and 1
n	Natural number
N	Neighbourhood (in the image grid)
N_{cols}, N_{rows}	Number of columns, number of rows
$\mathbb{N}$	Set $\{0, 1, 2, \ldots\}$ of natural numbers
$\mathscr{O}(\cdot)$	Asymptotic upper bound
Ω	Image carrier, set of all $N_{cols} \times N_{rows}$ pixel locations
p, q	Points in $\mathbb{R}^2$, with coordinates x and y
P, Q, R	Points in $\mathbb{R}^3$, with coordinates X, Y, and Z
π	Real constant $\pi = 4 \times \arctan(1) \approx 3.14159265358979$
Π	Polyhedron
r	Radius of a disk or sphere; point in $\mathbb{R}^2$ or $\mathbb{R}^3$
$\mathbb{R}$	Set of real numbers
$\mathbf{R}$	Rotation matrix
ρ	Path with finite number of vertices
s	Point in $\mathbb{R}^2$ or $\mathbb{R}^3$
S	Set
t	Time; point in $\mathbb{R}^2$ or $\mathbb{R}^3$
$\mathbf{t}$	Translation vector
T, τ	Threshold (real number)
u, v	Components of optical flow; vertices or nodes; points in $\mathbb{R}^2$ or $\mathbb{R}^3$
$\mathbf{u}$	Optical flow vector with $\mathbf{u} = (u, v)$
W, W_p	Window in an image, window with reference pixel p
x, y	Real variables; pixel coordinates (x, y) in an image
X, Y, Z	Coordinates in $\mathbb{R}^3$
$\mathbb{Z}$	Set of integers