Advances in Computer Vision and Pattern Recognition

Founding editor

Sameer Singh, Rail Vision, Castle Donington, UK

Series editor

Sing Bing Kang, Microsoft Research, Redmond, WA, USA

Advisory Board

Horst Bischof, Graz University of Technology, Austria
Richard Bowden, University of Surrey, Guildford, UK
Sven Dickinson, University of Toronto, ON, Canada
Jiaya Jia, The Chinese University of Hong Kong, Hong Kong
Kyoung Mu Lee, Seoul National University, South Korea
Yoichi Sato, The University of Tokyo, Japan
Bernt Schiele, Max Planck Institute for Computer Science, Saarbrücken, Germany
Stan Sclaroff, Boston University, MA, USA

More information about this series at http://www.springer.com/series/4205

Eckart Michaelsen · Jochen Meidow

Hierarchical Perceptual Grouping for Object Recognition

Theoretical Views and Gestalt-Law Applications

Springer

Eckart Michaelsen
Fraunhofer IOSB
Ettlingen, Baden-Württemberg, Germany

Jochen Meidow
Fraunhofer IOSB
Ettlingen, Baden-Württemberg, Germany

ISSN 2191-6586 ISSN 2191-6594 (electronic)
Advances in Computer Vision and Pattern Recognition
ISBN 978-3-030-04039-0 ISBN 978-3-030-04040-6 (eBook)
https://doi.org/10.1007/978-3-030-04040-6

Library of Congress Control Number: 2018960737

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Around the year 2008, I realized that much of what we had published as knowledge-based methods for image analysis was actually perceptual grouping. Moreover, these perceptual grouping rules were those that turned out to be more robust than the actual automatic knowledge utilization part. Moreover, the same constructions were needed over and over again, through many modalities of sensing and tasks to be fulfilled. One main source of malfunction of rule-based systems was the threshold parameters: Should two straight lines be parallel if their orientation deviation is less than ten degrees? Or rather five degrees? It became evident that such hard thresholds should be replaced by using soft assessment functions.

At the International Conference on Pattern Recognition 2012 in Tsukuba, I discussed the issue with Vera Yashina of the algebraic branch of the pattern recognition community of the Russian Academy of Sciences. We agreed that such approach is not really a syntactic approach anymore, it is an algebraic formulation: The Gestalt algebra. 2012 happened to be the year of a major upheaval in pattern recognition and machine vision. It was realized that deep learning utilizing convolutional neural networks yields superior performance on object recognition from imagery. Almost nobody in the community seemed to like those machines with their vast number of parameters, but the facts could not be ignored. In the few years that have passed since 2012, this neural network approach has been adapted to almost any task in machine perception and artificial intelligence with remarkable success. So isn't perceptual grouping utilizing Gestalt laws, and knowledge-based machine inference an outdated topic?

Neural network approaches existed before 2012. Their superior performance nowadays results from the training data amounts which are at hand now and from the advances in computing machinery. Still, anything that must not be learned, because it is already known, helps in concentrating these precious resources on learning the unknown things. The laws of seeing are known for more than one hundred years. Seeing must not be learned, it can be coded by implementing these laws in computing machinery. There has been enough knowledge about this topic published in numerous papers, and also in several very recommendable textbooks. Why then yet another book on Gestalt laws?

Because the aspect of hierarchical grouping has been hardly treated in the existing literature, e.g., a window sash may be made of a lattice of 12 small sub-windows, and two such sashes make a reflection symmetric window aggregate, and several of these are repeated as a frieze on a facade, and the building on which the facade is seen, is repeated along a road. It is much more likely that we, or our machines, encounter images containing such deep hierarchies through the scales, than that the images contain only random noise and clutter. The Gestalt algebra has been deliberately designed for such hierarchical patterns.

When asked to write a textbook on this topic, I realized that expertise in probability calculus, least squares estimation, and projective geometry would be needed, and I asked Jochen Meidow to join in. Together we revised the operations of Gestalt algebra and present them in the volume at hand. For each such Gestalt operation, there is a separate chapter, containing the definition, as well as examples of application, and some brief review of the corresponding literature. The most important chapter is the algebraic closure chapter, where all operations can participate in the construction of hierarchies of such aggregates. But the book would not be complete without a chapter connecting the method to the data—i.e., a chapter on the extraction of primitives from pictures, a chapter on the cooperation with machine-readable knowledge, and a chapter on cooperation with machine learning.

The book is intended for students, researchers, and engineers active in machine vision. We hope that the field may benefit from our methods and that some of our proposals may help to develop and improve future seeing machines. We thank the management of the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB in Ettlingen, Germany, for facilitating the work on it as an ancillary activity, while being committed to the day-to-day business.

Ettlingen, Germany Eckart Michaelsen
September 2018

Contents

1 Introduction 1
1.1 Examples of Pictures with Hierarchical Gestalt 1
1.2 The State of the Art of Automatic Symmetry and Gestalt Recognition 5
1.3 The Gestalt Domain 11
1.4 Assessments for Gestalten 14
1.5 Statistically Best Mean Direction or Axis 18
1.6 The Structure of this Book 19
References 21

2 Reflection Symmetry 23
2.1 Introduction to Reflection Symmetric Gestalten 23
2.2 The Reflection Symmetry Constraint as Defined for Extracted Primitive Objects 25
2.3 Reformulation of the Constraint as a Continuous Score Function 27
2.4 Optimal Fitting of Reflection Symmetry Aggregate Features . . . 29
2.5 The Role of Proximity in Evidence for Reflection Symmetry 31
2.6 The Role of Similarity in Evidence for Reflection Symmetry and How to Combine the Evidences 33
2.7 Nested Symmetries Reformulated as Successive Scoring on Rising Scale 35
2.8 Clustering Reflection Symmetric Gestalten with Similar Axes 41
2.9 The Theory of A Contrario Testing and its Application to Finding Reflection Symmetric Patches in Images 46

2.10 The Minimum Description Length Approach for Nested Reflection Symmetry . . . 48
2.11 Projective Symmetry . . . 48
References . . . 50

3 Good Continuation in Rows or Frieze Symmetry . . . 53
3.1 Related Work on Row Gestalt Grouping . . . 55
3.2 The Row Gestalt as Defined on Locations . . . 56
3.3 Proximity for Row Gestalten . . . 58
3.4 The Role of Similarity in Row Gestalten . . . 59
3.4.1 Vector Features . . . 60
3.4.2 Scale Features . . . 62
3.4.3 Orientation Features . . . 63
3.5 Sequential Search . . . 64
3.5.1 The Combinatorics of Row Gestalten . . . 64
3.5.2 Greedy Search for Row Prolongation . . . 65
3.6 The A Contrario Approach to Row Grouping . . . 67
3.7 Perspective Foreshortening of Rows . . . 67
References . . . 69

4 Rotational Symmetry . . . 71
4.1 The Rotational Gestalt Law as Defined on Locations . . . 72
4.2 Fusion with Other Gestalt Laws . . . 75
4.2.1 Proximity Assessments for Rotational Gestalten . . . 75
4.2.2 Similarity Assessments for Rotational Gestalten . . . 77
4.3 Search for Rotational Gestalten . . . 78
4.3.1 Greedy Search for Rotational Gestalten . . . 78
4.3.2 A Practical Example with Rotational Gestalten of Level 1 . . . 79
4.4 The Rotational Group and the Dihedral on Group . . . 82
4.5 Perspective Foreshortening of Rotational Gestalts . . . 82
References . . . 84

5 Closure—Hierarchies of Gestalten . . . 85
5.1 Gestalt Algebra . . . 86
5.2 Empirical Experiments with Closure . . . 90
5.3 Transporting Evidence through Gestalt Algebra Terms . . . 92
5.3.1 Considering Additional Features . . . 93
5.3.2 Propagation of Adjustments through the Hierarchy . . . 95
References . . . 100

6 Search . . . 101
6.1 Stratified Search . . . 101
6.2 Recursive Search . . . 102

6.3 Monte Carlo Sampling with Preferences . . . 103
6.4 Any-time Search Using a Blackboard . . . 104
References . . . 105

7 Illusions . . . 107
7.1 Literature about Illusions in Seeing . . . 107
7.2 Deriving Illusion from Top-down Search . . . 108
7.3 Illusion as Tool to Counter Occlusion . . . 108
References . . . 109

8 Prolongation in Good Continuation . . . 111
8.1 Related Work on Contour Chaining, Line Prolongation, and Gap Filling . . . 112
8.2 Tensor Voting . . . 112
8.3 The Linear Prolongation Law and Corresponding Assessment Functions . . . 116
8.4 Greedy Search for Maximal Line Prolongation and Gap Closing . . . 121
8.5 Prolongation in Good Continuation as Control Problem . . . 121
8.6 Illusory Contours at Line Ends . . . 123
References . . . 125

9 Parallelism and Rectangularity . . . 127
9.1 Close Parallel Contours . . . 127
9.2 Drawing on Screens as Graphical User Interface . . . 129
9.3 Orthogonality and Parallelism for Polygons . . . 130
References . . . 133

10 Lattice Gestalten . . . 135
10.1 Related Work on Lattice Grouping . . . 136
10.2 The Lattice Gestalt as Defined on Locations . . . 136
10.3 The Role of Similarity in Lattice Gestalt Grouping . . . 138
10.4 Searching for Lattices . . . 139
10.5 An Example from SAR Scatterers . . . 141
10.6 Projective Distortion . . . 143
References . . . 143

11 Primitive Extraction . . . 145
11.1 Threshold Segmentation . . . 146
11.2 Super-Pixel Segmentation . . . 148
11.3 Maximally Stable Extremal Regions . . . 150
11.4 Scale-Invariant Feature Transform . . . 152
11.5 Multimodal Primitives . . . 154

11.6 Segmentation by Unsupervised Machine Learning 154
11.6.1 Learning Characteristic Colors from a Standard Three Bytes Per Pixel Image 155
11.6.2 Learning Characteristic Spectra from a Hyper-Spectral Image 156
11.7 Local Non-maxima Suppression 159
References 161

12 Knowledge and Gestalt Interaction 163
12.1 Visual Inference 163
12.2 A Small Review on Knowledge-Based Image Analysis 166
12.3 An Example from Remotely Sensed Hyper-spectral Imagery 169
12.4 An Example from Synthetic Aperture RADAR Imagery 171
References 173

13 Learning 175
13.1 Labeling of Imagery for Evaluation and Performance Improvement 175
13.2 Learning Assessment Weight Parameters 178
13.3 Learning Proximity Parameters with Reflection Ground Truth 179
13.4 Assembling Orientation Statistics with Frieze Ground Truth . . . 181
13.5 Estimating Parametric Mixture Distributions from Orientation Statistics 183
References 187

Appendix A: General Adjustment Model with Constraints 189

Index 193

Notations

Assessment Functions

a_ϕ	Assessment function w.r.t. orientation
a_s	Assessment function w.r.t. scale
a_d	Assessment function w.r.t. proximity (distance)
a_f	Assessment function w.r.t. periodicity
a	Assessment of gestalt

Objects and Sets

o, s	Object, e.g., Gestalt
$\mathcal{G}, \mathcal{S}$	Set of Gestalten
$\mathcal{X}$	Set of two-dimensional points

Features

ϕ	Orientation (direction)
e	Elongation
f	Periodicity
s	Size or scale
$\boldsymbol{x}$	Coordinates two-dimensional point (x, y)

Constants and Parameters

$t.$	Threshold
π	Constant
κ	Parameter von Mises distribution