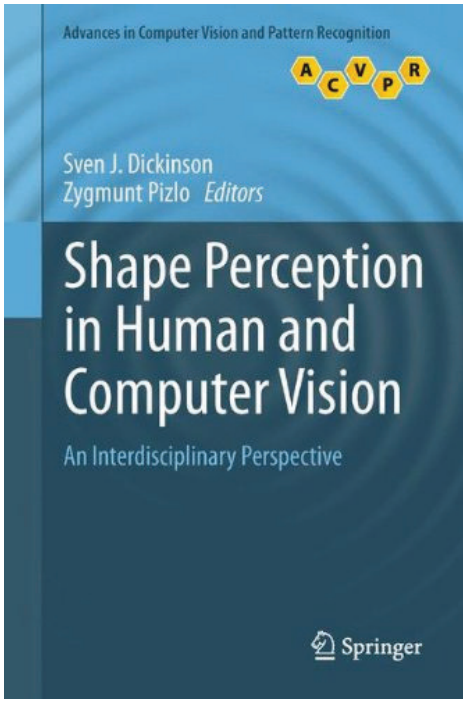
Shape Perception in Human and Computer Vision: An Interdisciplinary Perspective
Series: Advances in Computer Vision and Pattern Recognition
By Sven J. Dickinson and Zygmunt Pizlo (Eds.)

Springer, 2013

Reviewed by Dima Damen (University of Bristol, UK)

Published by Springer last year as part of the 'Advances in Computer Vision and Pattern Recognition' series, Dickinson and Pizlo edited thirty-four chapters written by seventy-three contributors. As their introduction states, the book is the result of a series of successful workshops on the topic (ECCV08, ECVP09, ECCV10 and VSS11). I recall attending the ECCV08 workshop in Marseilles, and despite the perfect sunshine, I found it more rewarding and enjoyable to attend the SPHCV workshop for a full day of interaction between human visual perception and computer vision experts. The book, as were the workshops, is an inter-disciplinary venture attempting to highlight similarities and identify differences in concepts and approaches between the two wealthy research domains. A comprehensive review of this book is quite difficult, with many chapters presenting state-of-the-art research and experimentation that are indeed worth reading. This review, instead, attempts to focus more on the arguments and concepts proposed by the various authors in both human and computer perception. Five ideas are presented next, which I found extremely interesting to think of while reading the book.

1. Definition and representation of shape:
In the various chapters, authors unanimously agree on the



Click above to go the publisher's web site for these volumes.

importance of shape, with experiments evidently showing that humans rely on shape and texture (rather than luminance on colour) in natural scene understanding (e.g. Ch. 5). Interestingly, and probably wisely, the editors do not attempt to adopt or unify a definition of 'shape'. Probably my most favourite description of the term is accredited to Wagemans (Ch. 6) where he states that "Shape is a beautiful thing". The question of whether all objects have shapes is discussed, or more interestingly whether certain objects have "more of a shape" than others (Ch. 2). For example Li et al. wonder whether a crumpled paper has a shape (Ch. 2) while Koenderink and van Doorn question whether space has a shape (Ch. 10).

Other attempts to define shape can be found throughout the book. In Ch. 25, Ommer specifies that shape is what 'glues' local features into a "sound spatial layout". In Ch. 23, Shi refers to shape as "an expressive abstraction of visual patterns in natural images." More mathematically, probably, Tyler refers to shape as "the properties of a manifold embedded in higher-dimensional space," (Ch. 24).

While some authors adopt a definition of shape, others devote their focus to rejecting currently-used definitions. In Ch. 2, Li et al. explain why the "invariance under transformation" is certainly not related to human shape perception, even if computationally-motivated. Crucially, several human vision experts argue whether the goal of computer vision shape perception actually matches that of human shape perception. Singh and Hoffman take an evolutionary point of view, where shape is merely an "effective code for expected fitness that has been tuned by natural selection" (Ch. 12).

Human shape perception certainly does not target veridical metric reconstruction, as "neither perceptual judgements nor motor actions are based on a veridical metric analysis of the visual scene", but only encodes relative properties of the object structure (Ch. 20).

With a more practical goal, authors discuss possible representations of shape. Cremers (Ch. 7) distinguishes between explicit representations (of points on the object's boundary or surface) and implicit representations (of the object's interior). Explicit parametric representations are adopted in Ch. 7 and Ch. 15, where the outline is modelled as a parametrised curve or surface. An implicit representation of the interior as a triangular mesh is presented in Ch. 21 as a computationally efficient one.

Skeleton-based representations are discussed (Ch. 3-5), motivated

by the uniqueness of skeletons and the ability to reconstruct a boundary from its skeleton. Other representations based on formelets (Ch. 5) or arclets (Ch. 18) are also addressed.

2. Fifty years of computer vision—a paradigm shift:
While shape has been the focus of research in Computer Vision during the 70s and 80s of last century, the trend has shifted dramatically recently, with fresh work on edge-based object detection rarely observed in major venues and journals. Several chapters discuss the poor performance of current detectors and texton-based descriptors on objects that have little or no texture (Ch. 23 and Ch. 33). In Ch. 25, Ommer presents a beautifully written historical review of shape based object/ visual recognition, a critique of developments in the field as well as an outlook into the future. Ch. 1 also discusses the paradigm shift and the influence of searching for a single category on formulating the shape perception problem.

3. Low-level, mid-level or high-level processing:
The book thoroughly discusses frameworks for the integration of low-level, mid-level and high-level processing towards shape perception. It is evident while reading this book that little belief is left in the hierarchical/pipeline approach to shape perception, in both humans and computers.

Several chapters in the book are devoted to discussing the feedforward feedback influences on shape analysis (Ch. 5, 6, 11, 25). In Ch. 11, Zucker clarifies that a deeper understanding of Neuroscience reveals a complex processing architecture. While separating segmentation (as a goal for early vision) from recognition (as a goal for higher-level vision) has always been problematic, trying to solve them jointly "makes it theoretically questionable about how to proceed" (Ch. 11).

In Ch. 5, Elder explains the argued involvement of feedback connections in the human brain during perception of complex natural scenes. Visual pathways communicated fragmented hypotheses to higher-order areas which generate more complete hypotheses of shape (via contour completion potentially). These hallucinations are subsequently tested (or confirmed) in earlier visual areas.

In Ch. 6, Wagemans adopts the viewpoint that low-level, mid-level and high-level analyses collaborate to achieve shape identification in human vision.

The influence of the high-level familiarity, or priors, of shape on low-level processing and the computation model for shape perception has also been tackled (Ch. 1, 4, 7, 30). In Ch. 30, Peterson and Cacciamani particularly focus on the link between perception and memory, proposing a dynamic view of object perception that strongly links the high and low levels of perception.

The book also devotes several chapters to the power of low-level processing on the perception of shape. Zaidi et al., for example, show how the orientation modulation in textured images, viewed monocularly, give a perception of the 3D shape (Ch. 22).

4. Perceptual grouping:
As expected, the power of and need for perceptual grouping is a primary topic in several chapters (Ch. 1, 4, 23, 25, 30), as shape parts are significantly less distinctive than texton-based features. "A single shape part, unlike a SIFT feature, carries very little distinctiveness. Only when shape primitives are non-accidentally grouped together do the resulting higher-order structures posses the indexing power ... to ... promising candidates", Dickinson et al. state (Ch. 1).

The topic of perceptual grouping is strongly linked to the Gestalt notion of 'foreground-segmentation'. This becomes more difficult when the object to clutter ratio (i.e. size of object in the scene) decreases (Ch. 23). In Ch. 30, Peterson and Cacciamani vividly argue against the Gestalt term "figure-ground segregation" as a prerequisite to perception in spite of ample evidence to support it. Despite the strong link, Feldman et al. state that "theoretical connections between shape and f/g [figure-ground] remain largely unexplored" (Ch. 4). As opposed to discussing whether foreground segmentation precedes shape perception, Wagemans state that "shape-perception, shape-based object identification and segmentation are all tasks that require perceptual organization" (Ch. 6).

Of the various approaches to perceptual grouping in this book, Ch. 1 uses region-based (using super-pixel) grouping guided by symmetry and contour closure. In Ch. 18, Kellman et al. discuss the notion of contour interpolation, guided by relatability, to unify regions across gaps towards a meaningful shape.

5. Two-dimensional (2D) versus three-dimensional (3D) representations:
While most chapters in the book discuss shape as a two-dimensional representation, other chapters focus on shape as a representation of a three-dimensional structure (Ch. 8, 9,

14). Hu et al. compare an object-centred representation of shape with a view-centred representation often using a “collection of 2D representations each covering a small portion of the modeled views” (Ch. 17).

They state that due to the difficulties of matching object-centred representation to images, most of the computational approaches favour the view-centred representation, while combining 2D and 3D primitives outperforms using 2D primitives solely (Ch. 17). Similarly, in Ch. 24 a framework is proposed to combine cues (depth, texture, ...) into both 2D and 3D shape representations. The effect of light on perception of three-dimensional shapes is reviewed in Ch. 8, 9 and 16. Evidence exists that humans use a prior on the lighting direction (particularly lighting from above) in shape perception (Ch. 16). Even beyond visual cues, Wallraven argues that shape perception is the result of multisensory processing, particularly considering the haptic modality (Ch. 23).

Conclusion: Reading Dickinson and Pizlo's compilations is both enjoyable and educational, due to the wide collection of contributions in a single volume. The book successfully addresses the balance between asking difficult questions, arguing certain answers and providing clues for future directions. The book would have benefited from a conclusion that unifies and compares the slightly fragmented ideas spread throughout the book. The editors though might have felt that this is the role of the reader. A recommended book to interested researchers working towards shape-based approaches to visual perception.

FREE BOOKS

The *IAPR Newsletter* is looking for reviewers for the books listed below. If you have interest and some knowledge in the topic, email us with your mailing address. We will send you a copy of the book—which you may keep—and will expect in return a review for the Newsletter. ~Zeeshan Zia, Associate Editor for Book Reviews

* ***Natural User Interfaces in Medical Image Analysis*** by Marek R. Ogiela and Tomasz Hachaj (ISBN 978-3-319-07799-4): *http://www.springer.com/computer/image+processing/book/978-3-319-07799-4*

* ***Fusion in Computer Vision*** by Bogdan Ionescu et al. (ISBN 978-3-319-05695-1): *http://www.springer.com/computer/image+processing/book/978-3-319-05695-1*

* ***Scalable Pattern Recognition Algorithms*** by Pradipta Maji and Sushmita Paul (ISBN 978-3-319-05629-6): *http://www.springer.com/computer/bioinformatics/book/978-3-319-05629-6*

* ***Biometrics in the New World*** by Julian Ashbourn (ISBN 978-3-319-04158-2): *http://www.springer.com/computer/image+processing/book/978-3-319-04158-2*

* ***Human Action Recognition with Depth Cameras*** by Jiang Wang et al. (ISBN 978-3-319-04560-3): *http://www.springer.com/computer/image+processing/book/978-3-319-04560-3*

* ***Image Blending Techniques and their Application in Underwater Mosaicing*** by Ricard Prados, Rafael Garcia, and László Neumann (ISBN 978-3-319-05558-9): *http://www.springer.com/computer/image+processing/book/978-3-319-05557-2*

* ***Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods*** by Chris Aldrich and Lidia Auret (ISBN 978-1-4471-5184-5): *http://www.springer.com/computer/ai/book/978-1-4471-5184-5*

* ***Visual Texture*** by Michal Haindl and Jiri Filip (ISBN 978-1-4471-4901-9): *http://www.springer.com/computer/image+processing/book/978-1-4471-4901-9*

* ***Imaging Spectroscopy for Scene Analysis*** by Antonio Robles-Kelly and Cong Phuoc Huynh (ISBN 978-1-4471-4651-3): *http://www.springer.com/computer/image+processing/book/978-1-4471-4651-3*

* ***Motion History Images for Action Recognition and Understanding*** by Md. Atiqur Rahman Ahad (ISBN 978-1-4471-4729-9): *http://www.springer.com/computer/image+processing/book/978-1-4471-4729-9*

* ***Introduction to Image Processing Using R*** by Alejandro C. Frery and Talita Perciano (ISBN 978-1-4471-4949-1): *http://www.springer.com/computer/image+processing/book/978-1-4471-4949-1*

* ***Time-of-Flight Cameras*** by Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Horaud(ISBN 978-1-4471-4657-5): *http://link.springer.com/book/10.1007%2F978-1-4471-4658-2*

* ***Real-Time Detection of Lines and Grids*** by Adam Herout, Markéta Dubská and Jiří Havel (ISBN 978-1-4471-4413-7): *http://link.springer.com/book/10.1007%2F978-1-4471-4414-4*

* ***Omnidirectional Vision Systems*** by Luis Puig and J.J. Guerrero (ISBN 978-1-4471-4946-0): *http://link.springer.com/book/10.1007%2F978-1-4471-4947-7*

* ***Annual Dynamics of Daylight Variability and Contrast*** by Siobhan Rockcastle and Marilyne Andersen (ISBN 978-1-4471-5232-3): *http://www.springer.com/computer/theoretical+computer+science/book/978-1-4471-5232-3*

* ***Efficient Algorithms for Discrete Wavelet Transform*** by K. K. Shukla and Arvind K. Tiwari (ISBN 978-1-4471-4940-8): *http://link.springer.com/book/10.1007%2F978-1-4471-4941-5*

* ***Graph-Based Clustering and Data Visualization Algorithms*** by Agnes Vathy-Fogarassy and Janos Abonyl (ISBN 978-1-4471-5157-9: *http://www.springer.com/computer/database+management+%26+information+retrieval/book/978-1-4471-5157-9*