Deformable Surface 3D Reconstruction from Monocular Images

Synthesis Lectures on Computer Vision

Editors
Gérard Medioni, ***University of Southern California***
Sven Dickinson, ***University of Toronto***

Synthesis Lectures on Computer Vision publishes 50- to 150 page publications on topics pertaining to computer vision and pattern recognition. The scope largely follows the purview of premier computer science conferences, such as ICCV, CVPR, and ECCV.

Deformable Surface 3D Reconstruction from Monocular Images
Mathieu Salzmann and Pascal Fua
2010

Boosting-Based Face Detection and Adaptation
Cha Zhang, Zhengyou Zhang
2010

Image-Based Modeling of Plants and Trees
Sing Bing Kang, Long Quan
2009

© Springer Nature Switzerland AG 2022
Reprint of original edition © Morgan & Claypool 2011

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Deformable Surface 3D Reconstruction from Monocular Images

Mathieu Salzmann and Pascal Fua

ISBN: 978-3-031-00682-1 paperback
ISBN: 978-3-031-01810-7 ebook

DOI 10.1007/978-3-031-01810-7

A Publication in the Springer series
SYNTHESIS LECTURES ON COMPUTER VISION

Lecture #3
Series Editors: Gérard Medioni, *University of Southern California*
Sven Dickinson, *University of Toronto*
Series ISSN
Synthesis Lectures on Computer Vision
Print 2153-1056 Electronic 2153-1064

Deformable Surface 3D Reconstruction from Monocular Images

Mathieu Salzmann
Toyota Technological Institute at Chicago

Pascal Fua
École Polytechnique Fédérale de Lausanne

SYNTHESIS LECTURES ON COMPUTER VISION #3

ABSTRACT

Being able to recover the shape of 3D deformable surfaces from a single video stream would make it possible to field reconstruction systems that run on widely available hardware without requiring specialized devices. However, because many different 3D shapes can have virtually the same projection, such monocular shape recovery is inherently ambiguous.

In this survey, we will review the two main classes of techniques that have proved most effective so far: The template-based methods that rely on establishing correspondences with a reference image in which the shape is already known, and non-rigid structure-from-motion techniques that exploit points tracked across the sequences to reconstruct a completely unknown shape. In both cases, we will formalize the approach, discuss its inherent ambiguities, and present the practical solutions that have been proposed to resolve them. To conclude, we will suggest directions for future research.

KEYWORDS

computer vision, deformable surfaces, monocular 3D shape recovery, structure from motion

Contents

Acknowledgments . . . . . **ix**

Figure Credits . . . . . **xi**

1 Introduction **1**

2 Early Approaches to Non-Rigid Reconstruction **5**

2.1 Physics-Based Models 5
2.1.1 The Finite Element Method 6
2.1.2 Physics-Based Methods for Computer Graphics 7
2.1.3 Physics-Based Methods for Computer Vision 8
2.2 Learned Deformation Models 9
2.2.1 Statistical Learning Methods 10
2.2.2 Learned Models for Non-Rigid Modeling 11
2.3 Regularization via Shape Parameterization 13
2.4 Legacy of the Previous Approaches 14

3 Formalizing Template-Based Reconstruction **17**

3.1 Problem Definition 17
3.1.1 Motivation 17
3.1.2 Camera Models 18
3.2 3D-to-2D Correspondences 19
3.3 Linear Formulation 20
3.3.1 Ambiguities under Weak Perspective Projection 20
3.3.2 Ambiguities under Full Perspective Projection 23

4 Performing Template-Based Reconstruction **29**

4.1 Imposing Temporal Consistency 30
4.2 Imposing Geometric Constraints 33
4.2.1 Developable Surfaces 33
4.2.2 Smooth Surfaces 34
4.2.3 Distance Constraints 43

5 Formalizing Non-Rigid Structure from Motion 51
5.1 Problem Definition 51
5.2 NRSFM under Weak Perspective Projection 53
5.3 NRSFM under Full Perspective Projection 54
5.4 Ambiguities of NRSFM 56
5.5 The Missing Data Problem 57

6 Performing Non-Rigid Structure from Motion 59
6.1 Orthonormality Constraints 60
6.2 Imposing Temporal Consistency 61
6.2.1 From Basis Shapes to Basis Trajectories 65
6.3 Imposing Geometric Constraints 67
6.3.1 Global Constraints 67
6.3.2 Local Constraints 74
6.4 Splitting a Global Surface into Local Ones 76

7 Future Directions 81

Bibliography 83

Authors' Biographies 99

Acknowledgments

Parts of the authors' work described in this book was supported by grants from the Swiss National Science Foundation.

We also wish to acknowledge the invaluable help from many present and former colleagues from EPFL's Computer Vision Laboratory. We are especially indebted to Slobodan Ilic, Vincent Lepetit, Francesc Moreno-Noguer, Raquel Urtasun, and Aydin Varol.

Mathieu Salzmann and Pascal Fua
December 2010

Figure Credits

Figure 1.1 J. Pilet, V. Lepetit, and P. Fua. Fast Non-Rigid Surface Detection, Registration and Realistic Augmentation. *International Journal of Computer Vision*, 76(2), February 2008. With kind permission from Springer Science+Business Media, Copyright © 2008 Springer. DOI: 10.1007/s11263-006-0017-9

1.2 K. S. Bhat, C. D. Twigg, J. K. Hodgins, P. K. Khosla, Z. Popovic, and S. M. Seitz. Estimating Cloth Simulation Parameters from Video. In *ACM Symposium on Computer Animation*, 2003. Copyright (c) 2003, Association for Computing Machinery, Inc. Reprinted by permission.

1.4 Courtesy of Mingxing Hu.

1.5 M. Salzmann, F. Moreno-Noguer, V. Lepetit, and P. Fua. Closed-Form Solution to Non-Rigid 3D Surface Registration. In *European Conference on Computer Vision*, October 2008. DOI: 10.1007/978-3-540-88693-8_43

2.1 V. Blanz and T. Vetter. A Morphable Model for the Synthesis of 3D Faces. In *ACM SIGGRAPH*, pages 187–194, August 1999. Copyright © 1999, Association for Computing Machinery, Inc. Reprinted by permission.

2.2 M. Dimitrijević, S. Ilić, and P. Fua. Accurate Face Models from Uncalibrated and Ill-Lit Video Sequences. In *Conference on Computer Vision and Pattern Recognition*, June 2004. Copyright © 2004 IEEE. Used with permission. DOI: 10.1109/CVPR.2004.26

4.2-4.4 M. Salzmann, R. Hartley, and P. Fua. Convex Optimization for Deformable Surface 3D Tracking. In *International Conference on Computer Vision*, October 2007. Copyright © 2007 IEEE. Used with permission. DOI: 10.1109/ICCV.2007.4409031

4.6 M. Perriollat and A. Bartoli. A Quasi-Minimal Model for Paper-Like Surfaces. In *BenCos: Workshop Towards Benchmarking Automated Calibration, Orientation and Surface Reconstruction from Images*, 2007. Copyright © 2007 IEEE. Used with permission. DOI: 10.1109/CVPR.2007.383356

4.10 V. Blanz and T. Vetter. A Morphable Model for the Synthesis of 3D Faces. In *ACM SIGGRAPH*, pages 187–194, August 1999. Copyright © 1999, Association for Computing Machinery, Inc. Reprinted by permission. DOI: 10.1145/311535.311556

4.11-4.12 M. Salzmann and P. Fua. Linear Local Models for Monocular Reconstruction of Deformable Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011. Copyright © 2011 IEEE. Used with permission. DOI: 10.1109/TPAMI.2010.158

4.13 A. Shaji, A. Varol, L. Torresani, and P. Fua. Simultaneous Point Matching and 3D Deformable Surface Reconstruction. In *Conference on Computer Vision and Pattern Recognition*, June 2010. Copyright © 2010 IEEE. Used with permission. DOI: 10.1109/CVPR.2010.5539827

4.15-4.16 M. Salzmann and P. Fua. Linear Local Models for Monocular Reconstruction of Deformable Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011. Copyright © 2011 IEEE. Used with permission. DOI: 10.1109/TPAMI.2010.158

6.1 I. Akhter, Y. Sheikh, and S. Khan. In Defense of Orthonormality Constraints for Nonrigid Structure from Motion. In *Conference on Computer Vision and Pattern Recognition*, June 2009. Copyright © 2009 IEEE. Used with permission. DOI: 10.1109/CVPR.2009.5206620

6.2 A. Shaji and S. Chandran. Riemannian Manifold Optimisation for Non-Rigid Structure from Motion. In *Conference on Computer Vision and Pattern Recognition*, 2008. Copyright © 2008 IEEE. Used with permission. DOI: 10.1145/15922.15903

6.3 V.Rabaud and S. Belongie. Linear Embeddings inNon-Rigid Structure fromMotion. In *Conference on Computer Vision and Pattern Recognition*, June 2009. Copyright © 2009 IEEE. Used with permission. DOI: 10.1109/CVPR.2009.5206628

6.4-6.5 I. Akhter, Y. Sheikh, S. Khan, and T. Kanade. Nonrigid Structure from Motion in Trajectory Space. In *Neural Information Processing Systems*, December 2008. Copyright © 2008 NIPS.

6.7-6.8 J. Xiao and T. Kanade. Uncalibrated Perspective Reconstruction of Deformable Structures. In *International Conference on Computer Vision*, 2005. Copyright © 2005 IEEE. Used with Permission. DOI: 10.1109/TPAMI.2007.70752

6.9-6.10 V. Rabaud and S. Belongie. Re-Thinking Non-Rigid Structure from Motion. In *Conference on Computer Vision and Pattern Recognition*, June 2008. Copyright © 2008 IEEE. Used with Permission. DOI: 10.1109/CVPR.2008.4587679

6.11-6.12 J. Fayad, A. Del Bue, L. Agapito, and P. M. Q. Aguiar. Non-Rigid Structure from Motion Using Quadratic Deformation Models. In *British Machine Vision Conference*, 2009.

6.13 X. Llado, A. Del Bue, and L. Agapito. Non-Rigid Metric Reconstruction from Perspective Cameras. *Image Vision Computing*, 28:1339–1353, 2010. Copyright © 2010, with permission from Elsevier. DOI: 10.1016/j.imavis.2010.01.014

6.14 A. Varol, M. Salzmann, E. Tola, and P. Fua. Template-Free Monocular Reconstruction of Deformable Surfaces. In *International Conference on ComputerVision*, September 2009. Copyright © 2009 IEEE. Used with Permission. DOI: 10.1109/ICCV.2009.5459403 DOI: 10.1109/ICCV.2009.5459403

6.15 J. Fayad, A. Del Bue, L. Agapito, and P. M. Q. Aguiar. Non-Rigid Structure from Motion Using Quadratic Deformation Models. In *British Machine Vision Conference*, 2009. 73, 74, 77. DOI: 10.1016/j.imavis.2005.10.004

6.16 J. Taylor, A. D. Jepson, and K. N. Kutulakos. Non-Rigid Structure from Locally-Rigid Motion. In *Conference on Computer Vision and Pattern Recognition*, June 2010. Copyright © 2010 IEEE. Used with Permission. DOI: 10.1109/CVPR.2010.5540002