A Survey of Photometric Stereo Techniques

Jens Ackermann
TU Darmstadt
jens.ackermann@gris.informatik.tu-darmstadt.de

Michael Goesele
TU Darmstadt
michael.goesele@gris.informatik.tu-darmstadt.de
Editorial Scope

Topics

Foundations and Trends® in Computer Graphics and Vision publishes survey and tutorial articles in the following topics:

- Rendering
- Shape
- Mesh simplification
- Animation
- Sensors and sensing
- Image restoration and enhancement
- Segmentation and grouping
- Feature detection and selection
- Color processing
- Texture analysis and synthesis
- Illumination and reflectance modeling
- Shape representation
- Tracking
- Calibration
- Structure from motion
- Motion estimation and registration
- Stereo matching and reconstruction
- 3D reconstruction and image-based modeling
- Learning and statistical methods
- Appearance-based matching
- Object and scene recognition
- Face detection and recognition
- Activity and gesture recognition
- Image and video retrieval
- Video analysis and event recognition
- Medical image analysis
- Robot localization and navigation

Information for Librarians

Foundations and Trends® in Computer Graphics and Vision, 2013, Volume 9, 4 issues. ISSN paper version 1572-2740. ISSN online version 1572-2759. Also available as a combined paper and online subscription.
A Survey of Photometric Stereo Techniques

Jens Ackermann
TU Darmstadt
jens.ackermann@gris.informatik.tu-darmstadt.de

Michael Goesele
TU Darmstadt
michael.goesele@gris.informatik.tu-darmstadt.de
# Contents

1 Introduction
   1.1 Scope of this survey ........................................ 4
   1.2 Overview .................................................. 5

2 Background
   2.1 Light-Matter-Interaction ................................... 6
   2.2 Lambertian photometric stereo ................................ 7
   2.3 Error sources ............................................... 8
   2.4 Baseline quality ............................................ 16
   2.5 Integration of surface gradients ............................. 19

3 Foundations .................................................. 23

4 Unknown lighting
   4.1 Point light source ........................................... 29
   4.2 General light source ......................................... 31
   4.3 Other generalizations ........................................ 33

5 Unknown reflectance ......................................... 35
   5.1 Diffuse and specular components ............................ 36
   5.2 More complex BRDF models .................................. 38
   5.3 Invariants .................................................. 41

Full text available at: http://dx.doi.org/10.1561/0600000065
### 6 Outlier handling

- 6.1 Explicit treatment ............................................. 47
- 6.2 General deviations ............................................. 49

### 7 Unknown lighting and reflectance

- 7.1 Diffuse specular separation ................................. 53
- 7.2 Parametric models ............................................ 54
- 7.3 Appearance profiles ........................................... 54

### 8 Multi-view settings

- 8.1 Fixed illumination ............................................. 59
- 8.2 Differential motion ........................................... 60
- 8.3 Wide baseline .................................................. 62

### 9 Outside the lab

- 9.1 Internet data ................................................... 71
- 9.2 Natural illumination .......................................... 76

### 10 Conclusion

- 10.1 Summary ...................................................... 80
- 10.2 Discussion and outlook .................................... 83
- 10.3 Looking ahead ............................................... 85

### Acknowledgements .............................................. 87

### References ....................................................... 88
Abstract

Reconstructing the shape of an object from images is an important problem in computer vision that has led to a variety of solution strategies. This survey covers photometric stereo, i.e., techniques that exploit the observed intensity variations caused by illumination changes to recover the orientation of the surface. In the most basic setting, a diffuse surface is illuminated from at least three directions and captured with a static camera. Under some conditions, this allows to recover per-pixel surface normals. Modern approaches generalize photometric stereo in various ways, e.g., relaxing constraints on lighting, surface reflectance and camera placement or creating different types of local surface estimates.

Starting with an introduction for readers unfamiliar with the subject, we discuss the foundations of this field of research. We then summarize important trends and developments that emerged in the last three decades. We put a focus on approaches with the potential to be applied in a broad range of scenarios. This implies, e.g., simple capture setups, relaxed model assumptions, and increased robustness requirements. The goal of this review is to provide an overview of the diverse concepts and ideas on the way towards more general techniques than traditional photometric stereo.

DOI: 10.1561/0600000065.
Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>normal</td>
</tr>
<tr>
<td>$X$</td>
<td>a 3D point</td>
</tr>
<tr>
<td>$\mathcal{P}$</td>
<td>set of 3D points (i.e. a patch or mesh)</td>
</tr>
<tr>
<td>$L$</td>
<td>radiance</td>
</tr>
<tr>
<td>$L_s$</td>
<td>incoming (source) radiance</td>
</tr>
<tr>
<td>$\omega$</td>
<td>direction vector</td>
</tr>
<tr>
<td>$D$</td>
<td>light matrix/vector</td>
</tr>
<tr>
<td>$\omega_{in/out}$</td>
<td>incoming/outgoing direction vector</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>solid angle</td>
</tr>
<tr>
<td>$d\omega$</td>
<td>differential solid angle</td>
</tr>
<tr>
<td>$\rho$</td>
<td>BRDF or diffuse constant</td>
</tr>
<tr>
<td>$c, C$</td>
<td>arbitrary constant (scalar or matrix)</td>
</tr>
<tr>
<td>$M$</td>
<td>number of images</td>
</tr>
<tr>
<td>$I$</td>
<td>image intensity</td>
</tr>
<tr>
<td>$f$</td>
<td>camera response</td>
</tr>
<tr>
<td>$\theta, \phi$</td>
<td>angles (typically zenith and azimuth)</td>
</tr>
<tr>
<td>$p, q$</td>
<td>coordinates of the gradient</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>standard deviation</td>
</tr>
<tr>
<td>$N$</td>
<td>normal distribution</td>
</tr>
<tr>
<td>$Z$</td>
<td>depth map/height field</td>
</tr>
<tr>
<td>$(u, v)$</td>
<td>2D coordinates</td>
</tr>
<tr>
<td>$N$</td>
<td>normal map/normal field</td>
</tr>
<tr>
<td>$E$</td>
<td>error (matrix, measure, etc.)</td>
</tr>
<tr>
<td>$R$</td>
<td>reflectance map or albedo matrix</td>
</tr>
<tr>
<td>$P$</td>
<td>number of pixels/patches</td>
</tr>
<tr>
<td>$S, T$</td>
<td>factorization of the radiance matrix</td>
</tr>
<tr>
<td>$A$</td>
<td>transformation matrix</td>
</tr>
<tr>
<td>$B$</td>
<td>arbitrary matrix</td>
</tr>
<tr>
<td>$B^\dagger$</td>
<td>pseudoinverse of $B$</td>
</tr>
<tr>
<td>$\alpha, \mu, \nu$</td>
<td>parameters</td>
</tr>
<tr>
<td>$\lambda, \gamma$</td>
<td>parameters</td>
</tr>
<tr>
<td>$r$</td>
<td>inner matrix dimension</td>
</tr>
<tr>
<td>$D_R$</td>
<td>reflection vector</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>angle</td>
</tr>
<tr>
<td>$\mathcal{D}, \mathcal{K}$</td>
<td>factorization of $L$</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>lobe function</td>
</tr>
<tr>
<td>$\mathcal{M}$</td>
<td>general model</td>
</tr>
<tr>
<td>$W$</td>
<td>reciprocal pair matrix</td>
</tr>
<tr>
<td>$o$</td>
<td>general observation</td>
</tr>
<tr>
<td>$K$</td>
<td>error free radiance matrix</td>
</tr>
<tr>
<td>$H$</td>
<td>indicator matrix</td>
</tr>
<tr>
<td>$v$</td>
<td>viewing direction</td>
</tr>
<tr>
<td>$W$</td>
<td>intensity profile matrix</td>
</tr>
<tr>
<td>$w$</td>
<td>intensity profile</td>
</tr>
<tr>
<td>$S$</td>
<td>the unit sphere in $\mathbb{R}^3$</td>
</tr>
<tr>
<td>$\mathcal{M}$</td>
<td>manifold</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>embedding</td>
</tr>
<tr>
<td>$\xi$</td>
<td>shadow function</td>
</tr>
<tr>
<td>$d$</td>
<td>distance score</td>
</tr>
<tr>
<td>$\tau$</td>
<td>projection operator</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>Haar-Wavelet</td>
</tr>
<tr>
<td>$h_i$</td>
<td>EMOR basis function</td>
</tr>
<tr>
<td>$\xi$</td>
<td>shadow function</td>
</tr>
</tbody>
</table>
The shape of an object, its reflectance, and the incoming illumination define the image that the object forms in our eyes or on a camera sensor. Even for a uniformly colored, diffuse object, the shading typically changes depending on the local orientation of the surface and the properties of the incoming light. Highlights form on glossy or specular surfaces, providing additional cues about the surface shape and reflectance as well as the illumination. Humans are remarkably good at deducing such information from images even under very general conditions (Thompson et al. [2011]). In contrast, recovering one or several of these individual components is a much more difficult task for a computer system.

To solve such inverse problems, different techniques exist such as shape from defocus (Pentland [1987]), shape from texture (Blostein and Ahuja [1989]), or (multi-view) stereo (Seitz et al. [2006]). In this survey, we direct our attention to photometric approaches. These exploit the intensity variations caused by illumination changes. Typically, the input is a set of images of an object or scene under a varying set of lighting conditions, and the output is the 3D shape, e.g. encoded as a field of surface normals.
The goal of classic photometric stereo is to recover the surface orientation from a known combination of reflectance and lighting in multiple images. Starting from this, a rich body of literature was developed that generalizes the problem and its solution in various directions. Many of these approaches assume a controlled capture setup because well-behaved input data is crucial for inverse problems. These requirements constrain the application mostly to the research community and production settings. In this survey, we put a focus on approaches with a potential to broaden this applicability, e.g., through simple capture setups or relaxed requirements.

1.1 Scope of this survey

Shape and appearance reconstruction have connections to different fields in computer vision, computer graphics, optimization theory, statistics, optics, etc. We will only cover those areas that are most relevant for this survey. These are approaches that rely on varying illumination in multiple images to recover at least the surface orientation and possibly even reflectance and illumination.

That excludes for example the large area of shape from shading approaches that operate on single images (Zhang et al. [1999], Johnson and Adelson [2011], Oxholm and Nishino [2012], Han et al. [2013]) and the related works on intrinsic image decomposition (e.g. Barron and Malik [2012, 2013]). These are highly ill-posed problems and their solution requires strong regularization. Using multiple images provides more information and better constrains the result space. Some works (e.g. Magda et al. [2001], Koppal and Narasimhan [2007], Liao et al. [2007]) exploit the fall-off in radiance with the distance from a near point light or more general changes in the apparent source intensity (e.g. Davis et al. [2005]). Our focus will be on illumination variations caused by directional changes in the incident light. We also do not consider purely specular surfaces (Healey and Binford [1986], Bonfort and Sturm [2003], Tarini et al. [2005], Chen et al. [2006], Nehab et al. [2008], Weinmann et al. [2013]) or the specialized approaches for face (Debevec et al. [2000], Zhou et al. [2007], Ghosh et al. [2011]) and dy-
namic performance (Ahmed et al. 2008, Vlasic et al. 2009, Wu et al. 2012) capturing. Similarly, we will mention only a few approaches that require highly complex capture setups because they have less potential to be adapted for unconstrained environments.

In many modern works ideas from different disciplines get combined, e.g. when silhouette constraints are merged with photometric cues in certain multi-view approaches. This makes a clear distinction to other fields difficult at times. We include such cases in our discussion if it suits the overall perspective.

1.2 Overview

We begin this survey with an introduction for readers unfamiliar with photometric stereo methods in Chapter 2. It serves as a tutorial and to promote awareness for the challenges present in practical applications of such techniques. Before proceeding to the main part, we briefly discuss some early works in Chapter 3 that laid the foundations for present-day research.

Finding an optimal categorization of photometric approaches to structure this survey is not trivial. We decide for a selection scheme based on common challenges among the algorithms such as unknown lighting in Chapter 4, complex reflectance in Chapter 5, or extremely uncontrolled conditions in Chapter 9. Within each category we aim at presenting relevant representatives instead of cluttering the exposition with lists of similar approaches.

Finally, we conclude the survey in Chapter 10 with a summary of the most important concepts and our views on future developments.


References


References


References


References


Yudeog Han, Joon-Young Lee, and In So Kweon. High quality shape from a single RGB-D image under uncalibrated natural illumination. In IEEE International Conference on Computer Vision, 2013.


References


References


Alexandros Panagopoulos, Chaohui Wang, Dimitris Samaras, and Nikos Paragios. Illumination estimation and cast shadow detection through a higher-order graphical model. In IEEE Conference on Computer Vision and Pattern Recognition, 2011.


James A. Paterson, David Claus, and Andrew W. Fitzgibbon. BRDF and geometry capture from extended inhomogeneous samples using flash photography. In Eurographics, 2005.


References


References


