Unconstrained Shape from Shading using CNN

Bachelor-Thesis von Daniel José Ceballos Jung aus Stuttgart
Tag der Einreichung:

1. Gutachten: Stefan Roth, Ph.D.
2. Gutachten: Stephan Richter, M.Sc.

TECHNISCHE
UNIVERSITÄT
DARMSTADT

visual inference

Unconstrained Shape from Shading using CNN

Vorgelegte Bachelor-Thesis von Daniel José Ceballos Jung aus Stuttgart

1. Gutachten: Stefan Roth, Ph.D.
2. Gutachten: Stephan Richter, M.Sc.

Tag der Einreichung:

Erklärung zur Bachelor-Thesis

Hiermit versichere ich, die vorliegende Bachelor-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 23.07.2018

(Daniel Ceballos)

Acknowledgements

First of all, I want to thank Professor Stefan Roth for giving me the opportunity to write my thesis in his visual inference group.

Further, I thank my supervisor Stephan for being a great help, giving lots of input and guiding me through this thesis.

I would like to thank my friend Janne for investing a lot of time in proofreading and giving great input on writing.

Eventually, I thank my family for their trust and unconditional support.

Abstract

Contents

3 Background

In this chapter a brief intruduction is given about the basic techniques used in this thesis. Section 3.1 gives a short description of the classic shape-from-shading problem and the basic assumptions. In section 3.2 a short introduction on convolutional neuronal networks is given and how to make training deep neuronal network architectures more easy 3.2.1.

3.1 Shape recovery using shading

Three components are essential to the shape-from-shading (SFS) problem. The object whose shape should be estimated, the light source and the camera which produces the image of the surface (see figure 3.1). Considering a three-dimensional coordinate system $O(x,y,z)$ attached to the camera and $O(z)$ attached to the to the optical-axis then $O(x,y)$ coincides with the image plane:

$$\mathbf{x} = f\frac{x}{z} \qquad \mathbf{y} = f\frac{y}{z} \tag{3.1}$$

where $(\mathbf{x},\mathbf{y})$ are the image coordinates of a point (x,y,z) made by a camera with focal length f. As the image irradiance and the object irradiance are proportional, the shape-from-shading problem can then be formalized as the "image irradiance equation":

$$R(p,q) = E(x,y) \tag{3.2}$$

where $E(x,y)$ is the grayscale level (irradiance) of the image in point $(\mathbf{x},\mathbf{y})$ and $R(p,q)$ is the reflection function, giving the amount of re-emitted light from the surface depending on its orientation (p,q). The orientation of a surface can be specified as by its gradient $(p,q,-1)$, where $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$ are the first derivatives of z with respect to x and y. So to derive the irradiance of an image knowledge about the light source ,camera and the objects reflection function is necessary. With this set up the image irradiance equation can be used to analyze an images. So solving the first order non-linear partial differential equation in $\mathbf{x}$ and $\mathbf{y}$ will give the orientation (p,y) and thus the shape of the surface.

In view of the complexity of the SFS problem restrictions and assumptions have to be made. Therefore classic approaches assume a single point light source, a Lambertian (matt) surface and orthographic projection in order to make the problem solvable.

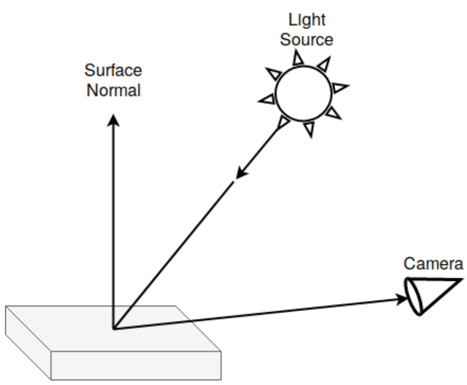

Figure 3.1: Image configuration for the shape-from-shading problem [16].

3.1.1 Orthographic and perspective projection

3.2 Machine learning usinng CNN

- erklären wie CNN funktionieren? - Idee für encoder decoder Architektur erklären

3.2.1 Deep Residual Learning

Deep residual neuronal networks In view of the fact that deeper neuronal networks are harder to train - erklären was ResNet ist
- wie funktionieren Residual - Architektur von ResNet18

List of Figures

List of Tables

Bibliography

[1] J. T. Barron and J. Malik. Shape, albedo, and illumination from a single image of an unknown object. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 334–341, 2012.

[2] J. Ben-Arie and D. Nandy. A neural network approach for reconstructing surface shape from shading. In *Proceedings 1998 International Conference on Image Processing*, pages 972–976 vol.2, 1998.

[3] Jean-Denis Durou, Maurizio Falcone, and Manuela Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, pages 22 – 43, 2008.

[4] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Optical Society of America A*, pages 2379–2394, 1987.

[5] Alan L. Yuille Heinrich H. Buelthoff. Shape-from-x: psychophysics and computation. *Proceedings of SPIE*, pages 1383 – 1383 – 12, 1991.

[6] B. K. P. Horn, R. S. Szeliski, and A. L. Yuille. Impossible shaded images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 166–170, 1993.

[7] B. K.P. Horn. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. Technical report, Cambridge, MA, USA, 1970.

[8] R. Huang and W. A. P. Smith. Shape-from-shading under complex natural illumination. *IEEE International Conference on Image Processing*, pages 13–16, 2011.

[9] Wenzel Jakob. Mitsuba renderer, 2010. http://www.mitsuba-renderer.org.

[10] M. K. Johnson and E. H. Adelson. Shape estimation in natural illumination. *IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[11] N. Khan, Lam Tran, and M. Tappen. Training many-parameter shape-from-shading models using a surface database. In *International Conference on Computer Vision Workshops*, pages 1433–1440, 2009.

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

[13] G. Oxholm and K. Nishino. Shape and reflectance from natural illumination. *European Conference on Computer Vision*, 2012.

[14] Vilayanur S. Ramachandran. Perceiving shape from shading. *Scientific American*, pages 76–83, 1988.

[15] Daniel L. Ruderman and William Bialek. Statistics of natural images: Scaling in the woods. *Advances in Neural Information Processing Systems 6*, pages 551–558, 1994.

[16] Robert James Woodham. Reflectance map techniques for analyzing surface defects in metal castings. Master's thesis, Massachusetts Institute of Technology, 1978.

[17] Jiajun Wu, Yifan Wang, Tianfan Xue, Xingyuan Sun, William T Freeman, and Joshua B Tenenbaum. MarrNet: 3D Shape Reconstruction via 2.5D Sketches. In *Advances In Neural Information Processing Systems*, 2017.

[18] Ruo Zhang, Ping-Sing Tsai, J. E. Cryer, and M. Shah. Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 690–706, 1999.