Depth Super-Resolution Meets Uncalibrated Photometric Stereo

Songyou PENG  Bjoern HAEFNER  Yvain QUEAU  Daniel CREMERS

Computer Vision Group
Technical University of Munich

ICCV 2017 Color and Photometry in Computer Vision Workshop
Outline

1. Introduction
2. Background
3. Methodology
4. Evaluation and Results
5. Conclusion
Outline

1. Introduction
2. Background
3. Methodology
4. Evaluation and Results
5. Conclusion
Problem Statement

Example: RGB-D data from ASUS Xtion Pro Live

- Good quality
- High resolution
- Noisy & missing areas
- Low resolution
Goal

Objective:
Use high-resolution photometric clues in the RGB image to turn the low-resolution depth maps into a refined, high resolution one.
Contribution

Propose a novel variational model to:

- disambiguate depth super-resolution through high-resolution photometric clues;
- disambiguate uncalibrated photometric stereo through low-resolution depth cues.
Outline

1 Introduction

2 Background

3 Methodology

4 Evaluation and Results

5 Conclusion
Background

Depth Super-Resolution

\[ z^i_0 = Kz + \varepsilon^i_z, \quad \forall i \in \{1, \ldots, n\} \]

\( z^i_0 \): input LR depth maps
\( z \): output HR depth map
\( K \): down-sampling kernel
\( \varepsilon^i_z \): noise \( \sim \mathcal{N}(0, \sigma^2_z) \)

\[
\min_z \mathcal{R}_z(z) + \frac{1}{2n} \sum_{i=1}^{n} \|Kz - z^i_0\|_{\ell^2}^2
\]
Background

[Grosse et al., ICCV 2009]

Photometric Stereo

\[ l^i = \rho l^i \cdot \begin{bmatrix} n(z) \\ 1 \end{bmatrix} + \varepsilon^i, \ \forall i \in \{1, \ldots, n\} \]

- \( l^i \): images under various lightings
- \( l^i \): lighting vector \( \mathbb{R}^4 \)
- \( \rho \): albedo / reflectance
- \( n(z) \): surface normal

\[
\min_{z} \mathcal{R}_l(z) + \frac{1}{2n} \sum_{i=1}^{n} \| \rho l^i \cdot \begin{bmatrix} n(z) \\ 1 \end{bmatrix} - l^i \|_{\ell^2}^2
\]
Background

Depth Super-Resolution

\[ z'^i_0 = Kz + \varepsilon^i_z, \quad \forall i \in \{1, \ldots, n\} \]

\( z'^i_0 \): input LR depths
\( z^i \): output HR depth
\( K \): down-sampling kernel
\( \varepsilon^i_z \): noise \( \sim N(0, \sigma^2_z) \)

\[ \min_z \mathcal{R}_z(z) + \frac{1}{2n} \sum_{i=1}^{n} \|Kz - z'^i_0\|_2^2 \]

Photometric Stereo

\[ l^i = \rho l^i \cdot \begin{bmatrix} n(z) \\ 1 \end{bmatrix} + \varepsilon^i, \quad \forall i \in \{1, \ldots, n\} \]

\( l^i \): images under various lightings
\( l^i \): lighting vector \( \mathbb{R}^4 \)
\( \rho \): albedo / reflectance
\( n(z) \): surface normal

\[ \min_z \mathcal{R}_l(z) + \frac{1}{2n} \sum_{i=1}^{n} \|\rho l^i \cdot \begin{bmatrix} n(z) \\ 1 \end{bmatrix} - l^i\|_2^2 \]
Background

Depth Super-Resolution

\[ z'_0 = Kz + \varepsilon'_z, \quad \forall i \in \{1, \ldots, n\} \]

- \( z'_0 \): input LR depths
- \( z \): output HR depth
- \( K \): down-sampling kernal
- \( \varepsilon'_z \): noise \( \sim \mathcal{N}(0, \sigma_z^2) \)

Photometric Stereo

\[ I^i = \rho l^i \cdot \begin{bmatrix} n(z) \end{bmatrix}_1 + \varepsilon^i, \quad \forall i \in \{1, \ldots, n\} \]

- \( I^i \): images under various lightings
- \( l^i \): lighting vector \( \mathbb{R}^4 \)
- \( \rho \): albedo / reflectance
- \( n(z) \): surface normal

Proposed Model: \( \min_z \sum_{i=1}^n \left\{ \|Kz - z'_0\|_2^2 + \lambda \left\| \rho l^i \cdot \begin{bmatrix} n(z) \end{bmatrix}_1 - I^i \right\|_2^2 \} \)
Outline

1. Introduction
2. Background
3. Methodology
4. Evaluation and Results
5. Conclusion
Methodology

With \((i, \star, p)\) the indices of images, channel and pixel,

\[
I^i_\star(p) = \rho_\star(p) I^i_\star \cdot \begin{bmatrix} n(p) \\ 1 \end{bmatrix} + \varepsilon^i_\star(p)
\]
Methodology

With \((i, \star, p)\) the indices of images, channel and pixel,

\[
I^i_\star(p) = \rho_\star(p) I^i_\star \cdot \left[ \begin{array}{c} n(p) \\ 1 \end{array} \right] + \varepsilon^i_\star(p)
\]

\[
n(p) = \frac{1}{d(z)(p)} \left[ -z(p) - \nabla z(p) \cdot (p - p^0) \right]
\]

\(f\): focal length  
\(p^0\): principal point  
\(d(z)\): normalizer
Methodology

With \((i, \star, p)\) the indices of images, channel and pixel,

\[
\begin{align*}
l^i_\star(p) &= \rho_\star(p) \cdot \mathbf{n}(p) + \varepsilon^i_\star(p) \\
n(p) &= \frac{1}{d(z)(p)} \begin{bmatrix} f \nabla z(p) \\
-z(p) - \nabla z(p) \cdot (p - p^0) \end{bmatrix}
\end{align*}
\]

where

- \(f\): focal length
- \(p^0\): principal point
- \(d(z)\): normalizer

\[
A^i(z, \rho, l^i)^\top \begin{bmatrix} \nabla z \\
z \end{bmatrix} = b^i(\rho, l^i) + \varepsilon^i
\]
Proposed Variational Model

Here we have:

- **depth super-resolution cue:** \( z_0^i = Kz + \varepsilon_z^i, \quad \forall i \in \{1, \ldots, n\} \)

- **photometric stereo cue:** \( A^i(z, \rho, l^i) \top \begin{bmatrix} \nabla z \\ z \end{bmatrix} = b^i(\rho, l^i) + \varepsilon^i \)

The final variational model is acquired from maximum likelihood:

\[
\min_{z, \rho, \{l^i\}_i} \left\{ \sum_{i=1}^n \|Kz - z_0^i\|_2^2 + \lambda \sum_{i=1}^n \left\| A^i(z, \rho, l^i) \top \begin{bmatrix} \nabla z \\ z \end{bmatrix} - b^i(\rho, l^i) \right\|_2^2 \right\}
\]
Alternating Optimization Workflow

\[
\min_{z, \rho, \{l^i\}_i} \left\{ \sum_{i=1}^{n} \|Kz - z_i^0\|_2^2 + \lambda \sum_{i=1}^{n} \left\| A^i(z, \rho, l^i)^\top \left[ \nabla z \right] - b^i(\rho, l^i) \right\|_2^2 \right\}
\]
Alternating Optimization Workflow

Input LR depth
Output HR depth
Details

S. Peng, B. Haefner, Y. Quéau, D. Cremers: Depth Super-Resolution Meets Uncalibrated Photometric Stereo
Outline

1. Introduction
2. Background
3. Methodology
4. Evaluation and Results
5. Conclusion
Synthetic Data

3D shape  
Ground truth HR depth  
LR noisy depth

HR albedo map

HR photometric stereo images

1Source: https://mtex-toolbox.github.io/files/doc/EBSDSpatialPlots.html
Quantitative Evaluation

Number of images

$n \in [10, 30]$ is a good compromise between accuracy and speed
Quantitative Evaluation

Parameter tuning

\[
\min_{z, \rho, \{l^i\}_i} \left\{ \sum_{i=1}^{n} \|Kz - z^i_0\|_2^2 + \lambda \sum_{i=1}^{n} \left\| A^i(z, \rho, l^i) \begin{bmatrix} \nabla z \\ z \end{bmatrix} - b^i(\rho, l^i) \right\|_2^2 \right\}
\]

\[\lambda \in [10^{-2}, 10^1] \text{ provide satisfactory results}\]
Quantitative Evaluation

RMSE = 0.0579
MAE = 65.7150

Input depth

RMSE = 0.0728
MAE = 34.4129

Depth super-resolution with TV
Quantitative Evaluation

RMSE = 0.0579
MAE = 65.7150

Input depth

RMSE = 0.9199
MAE = 41.8041

LDR Photometric Stereo
[Papadhimitri and Favaro, IJCV 2014]
Quantitative Evaluation

RMSE = 0.0579
MAE = 65.7150
Input depth

RMSE = 0.1655
MAE = 38.9316
RGBD-Fusion

[Or-El et al., CVPR 2015]
Quantitative Evaluation

RMSE = 0.0579
MAE = 65.7150
Input depth

RMSE = 0.0314
MAE = 1.45280
Ours
Quantitative Evaluation

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>0.0579</td>
<td>65.7150</td>
</tr>
<tr>
<td>Depth SR</td>
<td>0.0728</td>
<td>34.4129</td>
</tr>
<tr>
<td>LDR PS</td>
<td>0.0919</td>
<td>41.8041</td>
</tr>
<tr>
<td>RGBD-Fusion</td>
<td>0.1655</td>
<td>38.9316</td>
</tr>
<tr>
<td>Ours</td>
<td>0.0314</td>
<td>1.45280</td>
</tr>
</tbody>
</table>
Qualitative Evaluation

Input RGB image
Input depth
Depth SR
LDR-PS
RGBD-Fusion
Ours
Qualitative Evaluation

Input RGB image

Input depth

Depth SR

LDR-PS

RGBD-Fusion

Ours
Outline

1. Introduction
2. Background
3. Methodology
4. Evaluation and Results
5. Conclusion
Conclusions and Future work

- We proposed a novel variational framework for joint depth super-resolution and reflectance/light estimation
- Our method can be used out-of-the-box with common devices
- Theoretical analysis of this approach will be the next step

Data and codes are available on https://github.com/pengsongyou/SRmeetsPS
Depth Super-Resolution Meets Uncalibrated Photometric Stereo

Songyou PENG  Bjoern HAEFNER  Yvain QUEAU  Daniel CREMERS

Computer Vision Group
Technical University of Munich

ICCV 2017 Color and Photometry in Computer Vision Workshop
Shape from shading ambiguity

(a) An image  
(b) A possible explanation

(c) painter’s  
(d) sculptor’s

(e) Lighting designer’s

[Adelson and Pentland, 1996]
Generalized Bas-Relief (GBR)

[belhumeur et al., IJCV 99]