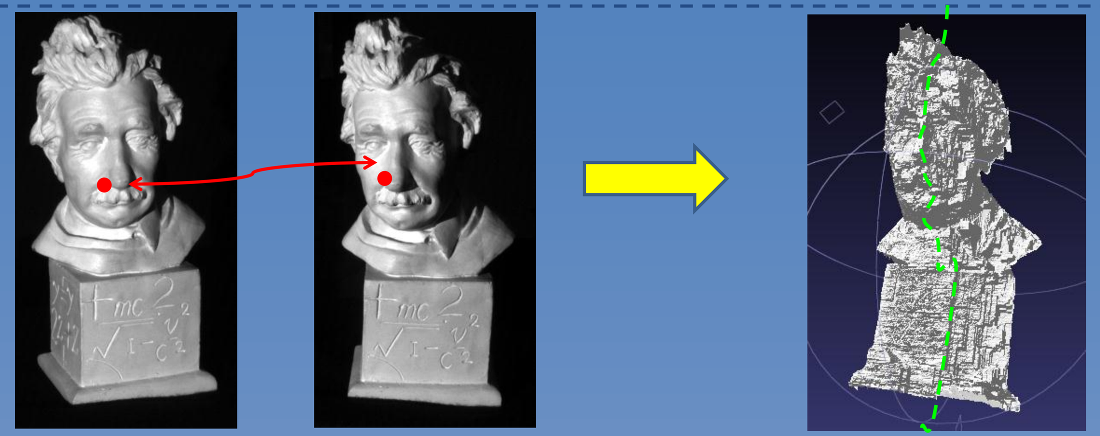
Binocular Photometric Stereo

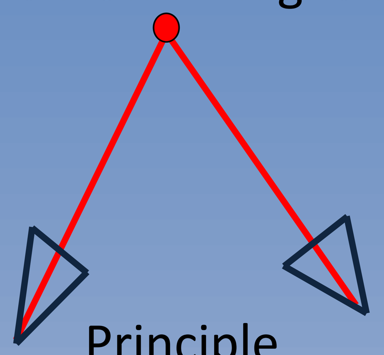
Hao Du[1,3], Dan B Goldman[2] and Steven M. Seitz[1,3]
[1]University of Washington, Seattle, WA , USA [2]Adobe Systems, Seattle, WA, USA [3]Google Inc., USA

Motivation

Binocular Stereo



Left View Right View Reconstruction



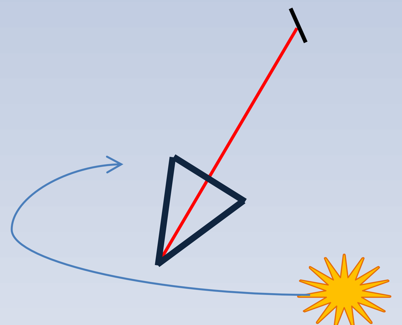
Principle

Pros.
- Metric depths

Cons.
- Limited surface quality at texture-less areas

Photometric Stereo

Captured Images Reconstruction



Principle

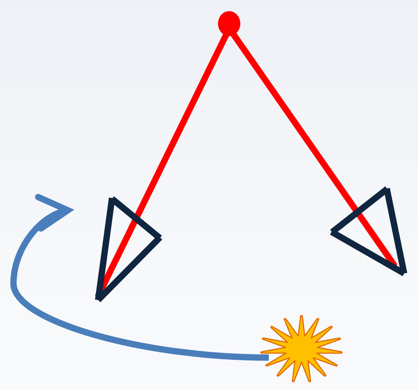
Pros.
- Detailed surface normals

Cons.
- Only relative depths;
- Handles discontinuities poorly

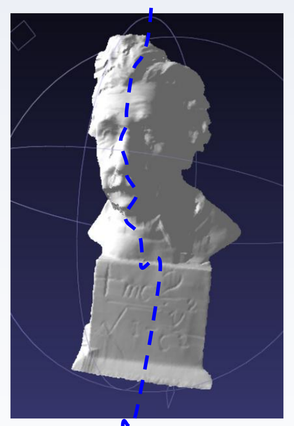
Our Work

Binocular Photometric Stereo

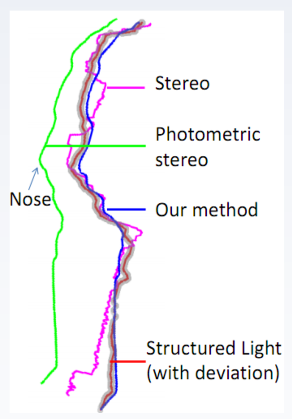
To achieve the best of both worlds!



Principle



Our Method



Cross Section

Setting:
Add a **second** camera for photometric stereo

Formulation:
A single **convex** optimization

Filter Flow

Filter Flow Background

S. Seitz and S. Baker, 2009
- A framework to model image transformation

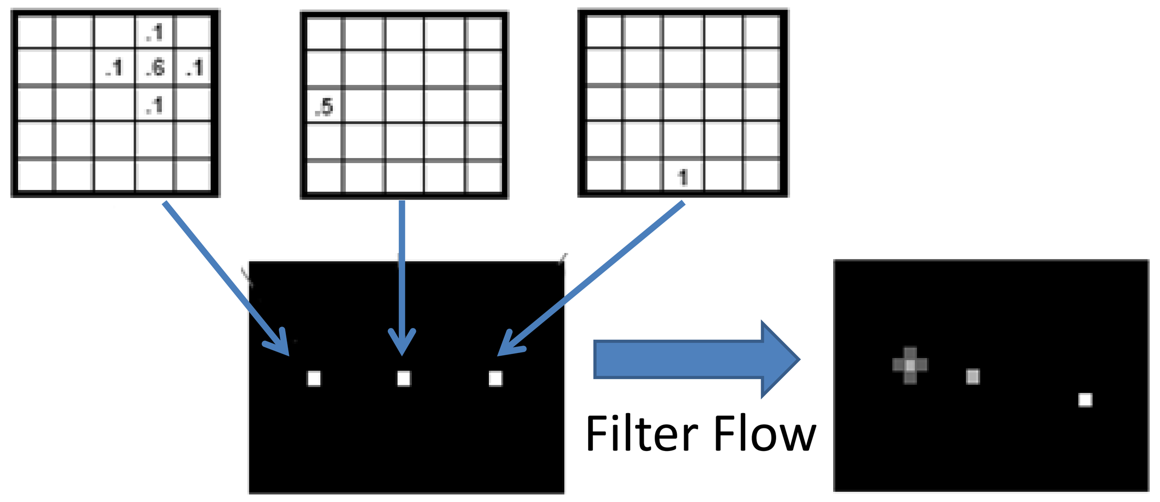
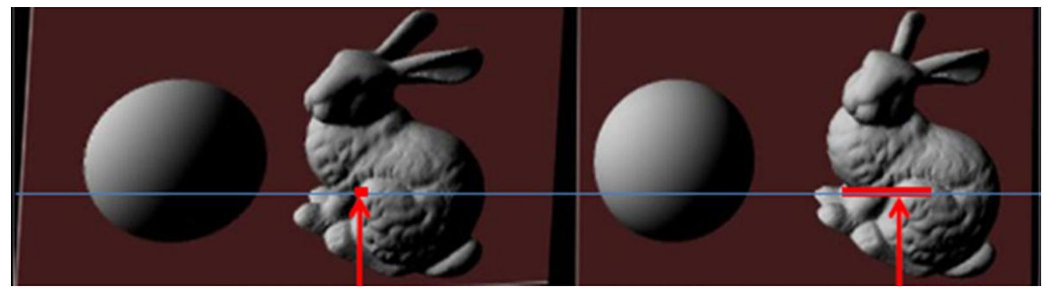


Image 1 Image 2

Filter flow models each pixel as a filtered version of the other image, with constraints on the entries of the filter kernel.

Modeling Stereo with Filter Flow

1D Filters



$\mathbf{u} = (u, v)$ $M^{\mathbf{u}}$

1.9 Pixel shift to the left 0 0 0 0.9 0.1 0 0 0 0 0 0

Disparity is the centroid of filter kernel.
- Disparity and depth are nonlinearly related.
- Normal and depth are linearly related.

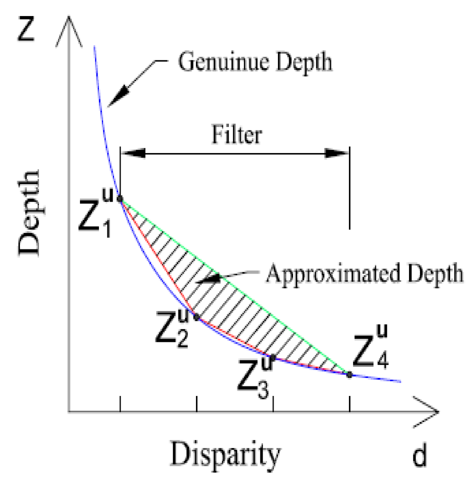
Linear Optimization Objectives

- Data Objective (DO): (Stereo intensity match)

$$\sum_{\mathbf{u}} \left\| I_l^{\,i}(u,v) - \sum_j M_j^{\mathbf{u}} I_r^{\,i}(u+j,v) \right\|_1$$

- Normal Objective (NO): (Normal match)

$$\left\|\hat{T}_u^{\mathbf{u}} \cdot N^{\mathbf{u}}\right\|_1 + \left\|\hat{T}_v^{\mathbf{u}} \cdot N^{\mathbf{u}}\right\|_1 ,$$



where

- Tangents: $\hat{T}_u^{\mathbf{u}} = \frac{\partial \hat{P}^{\mathbf{u}}}{\partial u}$ $\hat{T}_v^{\mathbf{u}} = \frac{\partial \hat{P}^{\mathbf{u}}}{\partial v}$
- 3D Pos.: $\hat{P}^{\mathbf{u}} = (\hat{X}^{\mathbf{u}}, \hat{Y}^{\mathbf{u}}, \hat{Z}^{\mathbf{u}})$
- Depth $\hat{Z}^{\mathbf{u}} = \sum_j M_j^{\mathbf{u}} Z_j^{\mathbf{u}} = \sum_j M_j^{\mathbf{u}} \frac{fb}{j}$

Constraints

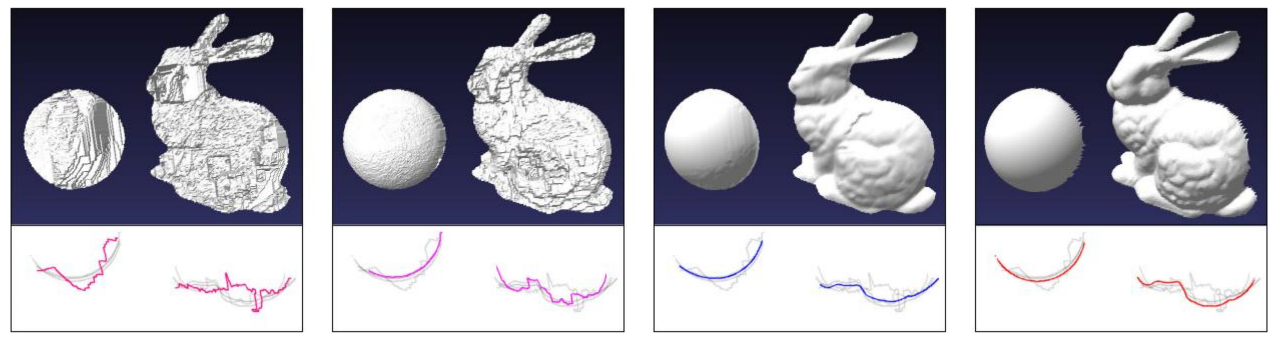
- Intensity preservation
  - Non-negative: $M_j^{\mathbf{u}} \geq 0 | \forall j$ (POS-M)
  - Sum-to-one: $\sum M_j^{\mathbf{u}} = 1$ (SUM1-M)

The Convex Optimization

DO + NO s.t. POS-M and SUM1-M
- Solve for filter entries $M_j^{\mathbf{u}}$
- Recover depths according to the approximation

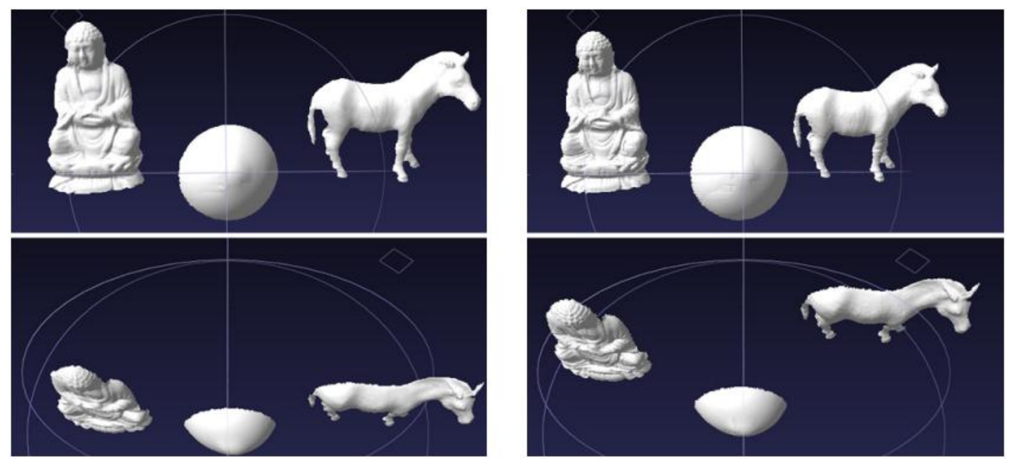
Results

Simulated Data - Comparing with Binocular Stereo

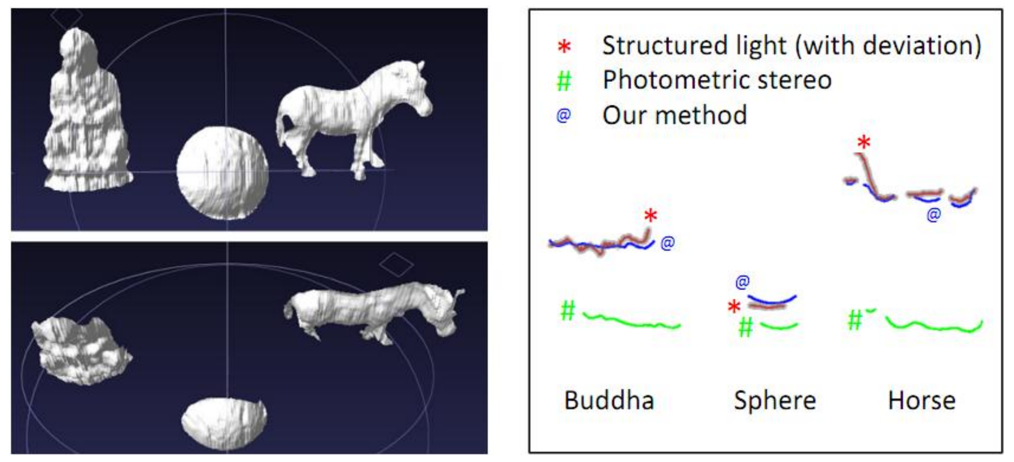


2 view stereo (filter flow) | 2 view spacetime st. | Our method | Ground truth

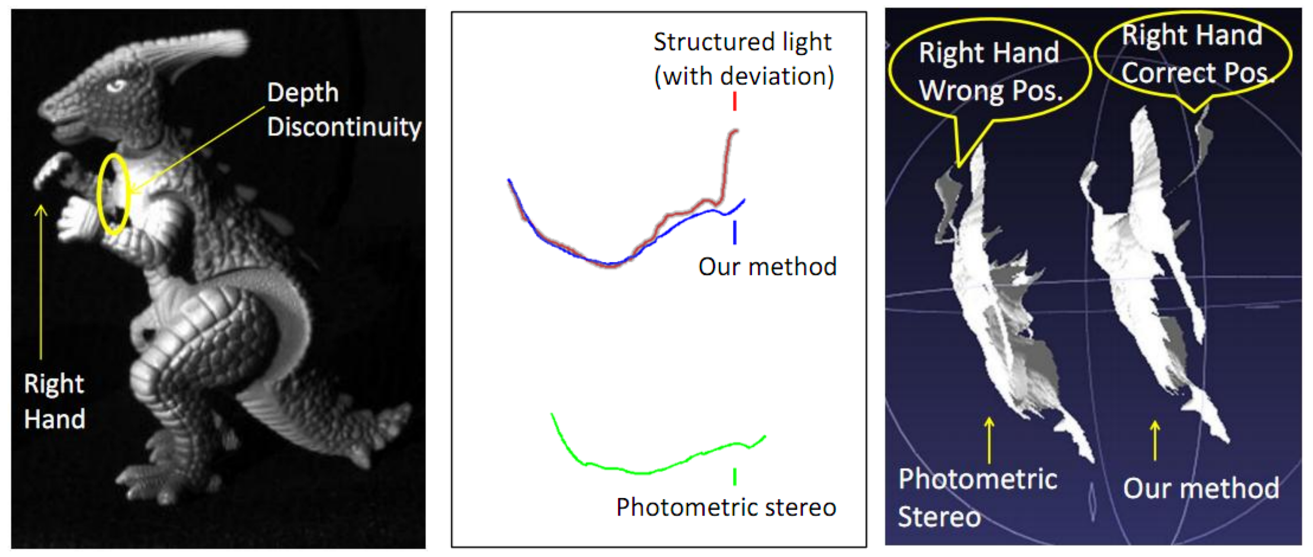
Real Data - Comparing with Photometric Stereo



Photometric stereo Our method



Structured light Cross section



An input image Cross section Top view

Real Data - Comparing with [Nehab'05].

Our method Nehab et al. Cross section

Our method simultaneously optimizes correspondence and normal cues, therefore it operates effectively at low textured flat areas.

Optional Compactness Objective (non-linear)

$$\sum_j M_j^{\mathbf{u}} \|j - \bar{M}^{\mathbf{u}}\|_2^2$$

Compactness Obj.

$$\sum_j M_j^{\mathbf{u}} [\max(0, \|j - \bar{M}^{\mathbf{u}}\|_2^2 - c^2)]$$

Soft-Compactness Obj.

Compactness Soft Compactness Cross section

	Our method	Compactness	Soft-CP
Bunny	0.199	0.191	0.112
Einstein	0.184	0.182	0.136
Three Objects	0.253	0.244	0.138
Dinosaur	0.170	0.170	0.083

Metric errors (in unit length)

Related Methods

Integrating with sparse stereo samples

These methods compute shape from photometric stereo as a separate step. [e.g. Lee 91]

Integrating with dense scanned shape

These methods use a laser scanned shape as input. [e.g. Neheb 05]

Multiview photometric stereo

These methods do not use surface appearance cues;, and use multiview [e.g. Hernandez 08, Vlasic 09]