Shape from Varying Illumination and Viewpoint

Neel Joshi David J. Kriegman
University of California, San Diego

UCSD

Goal

- Given views of an object rotating in front of a light and a camera, where the lighting and object rotation is unknown, recover the structure

Contributions

- We recover shape under varying viewpoint and lighting by combining multi-view and photometric stereo
- We derive a multi-ocular photometric matching cost. By minimizing this using a graphcut, we find correspondence
- Once correspondence is found, photometric stereo recovers normals, and the depth map from correspondence and the normals are fused to recover high quality 3-D shape
- Relative to previous related work, we produce higher-quality reconstructions using a non-iterative method with fewer images

Summary

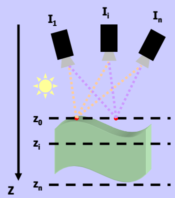
- Tomasi-Kanade Factorization to get projection matrices for each image
- Search over depths for each (x,y) point
  - Project point from the reference view to every other image
  - Get image patch centered around each projected point
  - Minimize Rank-3 approximation error using a graphcut
  - Perform photometric stereo and use the normals and the depth map to recover the final surface
- Key idea:
  - Rank-3 error is higher at the incorrect depth due to patches having inconsistent shading and albedo

Combining Multi-view Stereo and Photometric Stereo

- Fixed Light Location
- Multiple Camera Views
- Correspondence Search
- Shape is often noisy

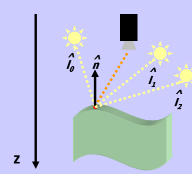
- Multiple Light Location
- Fixed Camera View
- Potential Ambiguity (GBR)

Multi-view and Photometric Stereo



- Search over depth range
- Find depths where the cost function is zero
- Results in a dense correspondence map

Multi-view Stereo



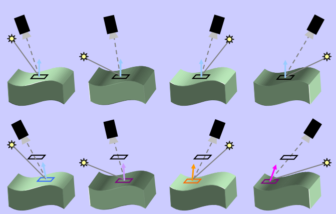
$$I = [\hat{i}_0 \quad \hat{i}_1 \quad \hat{i}_2]$$

$$L = [\hat{l}_0 \quad \hat{l}_1 \quad \hat{l}_2]$$

$$I = NL$$

$$N = IL^{-1}$$

Photometric Stereo

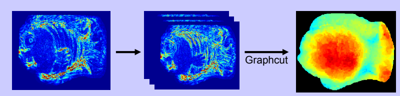


Rank-3 Approximation for pixel neighborhood:

$$I = [\hat{i}_0 \quad \dots \quad \hat{i}_n] \qquad [U \; S \; V] = svd(I)$$

$$U' = U(:,1:3) \qquad V' = V(1:3,:)$$

$$S' = S(1:3,1:3) \qquad I' = U'S'V'^T$$

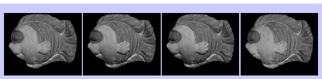


Rank-3 Approximation error for the pixel neighborhood:

$$c = \|I - I'\|^2$$

Compute for all pixels over a range of depths and find the depth map by minimizing the cost using a graphcut

Computing the Depth Map using a Photometric Cost Function



- Use depth map to align input images
- Perform Un-calibrated Photometric stereo:

$\tilde{N}$ = Pseudo-Normals $\tilde{L}$ = Pseudo-Lighting

- Compute linear transform A such that:

$$\frac{\tilde{n}}{|\tilde{n}|} A = \nabla(\hat{z}) \qquad \tilde{N}A = N$$

Recovering Normals

Error Function Minimized to Solve for the Final Surface:

$$J(S) = E^P + E^N + E^S$$

Position Error:

$$E^P = \lambda_1 \Sigma_{k=1}^m [S_k - z_k]^2$$

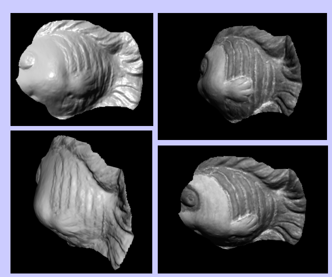
Normal Error:

$$E^N = (1 - \lambda_1) \left[\Sigma_{k=1}^m [T_k^x \cdot N_k^C]^2 + \Sigma_{k=1}^m [T_k^y \cdot N_k^C]^2\right]$$

Smoothness Constraint:

$$E^S = \lambda_2 \Sigma_{k=1}^m \left[(\nabla^2 S)(\sigma_k)\right]$$

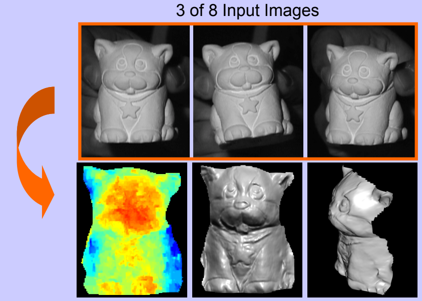
Linear Least Squares



Solving for the Final Surface

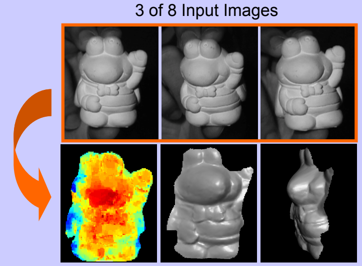
Results

3 of 8 Input Images



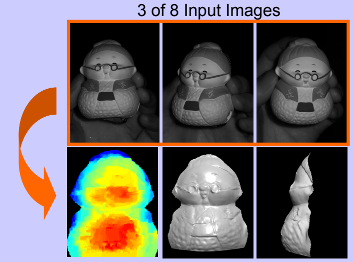
Depth Map (graphcut) Final Surface

3 of 8 Input Images



Depth Map (graphcut) Final Surface

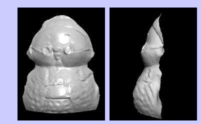
3 of 8 Input Images

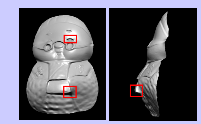


Depth Map (graphcut) Final Surface

Comparison

Our Results





From Lim et al. 2005

Selected References

- L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. ICCV 2003.
- J. Lim, J. Ho, M.-H. Yang, and D. Kriegman. Passive photometric stereo from motion. ICCV 2005.
- A. Maki, M. Watanabe, and C. Wiles. Geotensity: Combining motion and lighting for 3d surface reconstruction. IJCV, 48(2):75–90, 2002.
- D. Simakov, D. Frolova, and R. Basri. Dense shape reconstruction of a moving object under arbitrary, unknown lighting. ICCV 2003
- H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. JOSA A (11):3079–3089, Nov 1994.
- D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and normals for precise 3D geometry. SIGGRAPH 2005.