



BROWN

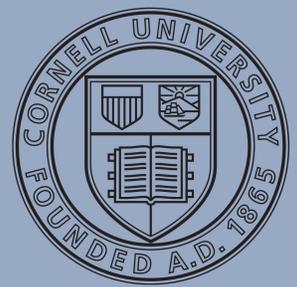
Shape from Depth Discontinuities under Orthographic Projection

Douglas Lanman*
Brown University

Daniel Cabrini Hauage†
Cornell University

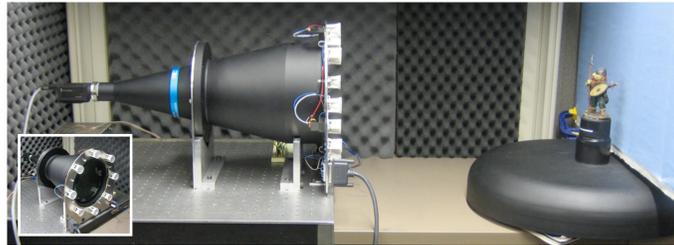
Gabriel Taubin*
Brown University

*{dlanman, taubin}@brown.edu †hauage@cs.cornell.edu



Abstract

- We present a new shape completion algorithm that can be used to detect and fill local concavities in the surface recovered from the visual motion of depth discontinuities viewed under orthographic projection.
- We analyze the properties of orthographic multi-flash cameras for depth edge detection, using either near-field point sources or directional illumination.
- We describe a calibration method for orthographic cameras using at least four images of a planar pattern augmented with a single point above its surface.
- We present and analyze the performance of an experimental prototype, which is the first to exploit the unique properties of orthographic multi-flash imaging to reconstruct the 3-D shape of solid surfaces.



Related Work

Epipolar-Plane Image Analysis

One of the earliest studies about EPIs was published by Bolles [1], in which he considers the case of linear motion for parallel cameras. In this case, a single scene point maps to a line in the EPI, with a slope corresponding to the distance of the point from the camera. Lines corresponding to points closer to the camera overlap those for points further away, allowing reconstruction without explicit feature matching [2]. This model is extended by Baker and Bolles [3] to deal with non-parallel cameras. Feldmann et al. [4] describe the properties of EPI curves for a circular motion path due to camera rotation. Their parameterized curves cannot be applied to our system, as they model texture features rather than depth discontinuities.

Optical Shape Capture Methods

In this work we propose a shape capture method inspired by the work of Crispell et al. [5]. In contrast to photometric stereo [6], in which lights are placed far from a camera, Raskar et al. [7] propose placing light sources close to the center of projection to estimate the set of visible depth discontinuities. Crispell et al. [5] show that such multi-flash cameras can be used to measure the visual motion of depth discontinuities as an object undergoes rigid rotation, allowing surface reconstruction using the differential method of Cipolla and Giblin [8].

Orthographic Imaging and Illumination

As demonstrated by Watanabe and Nayar [9], a telecentric lens can be fashioned from a conventional lens by placing an aperture at a specific location (e.g., at a focal point for a thin lens).

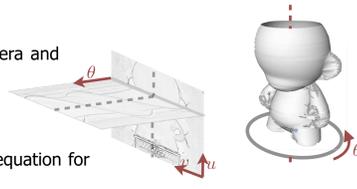
Curve and Surface Completion

Shape completion has been extensively studied in 2-D and 3-D. Numerous 2-D curve completion schemes have been proposed [10, 11, 12, 13]; generally, two position and tangent constraints are specified. As an infinite number of curves satisfy such boundary conditions, additional constraints have been proposed to obtain a unique solution. Ullman [10] proposes a curve of minimum total curvature formed by two circular arcs, tangent at both ends, meeting in the center. Horn [13] further analyzes the curvature energy when the number of circular arcs increases, proving that the internal energy is smaller than that of an Euler spiral or a simple circle. Kimia et al. [12] propose minimizing variation of curvature, yielding completions based on the Euler spiral. Pauly et al. [14] complete meshes by detecting and replicating repeated patterns. In a closely-related work, Crispell et al. [5] fill gaps using the implicit surface defined by an oriented point cloud. Finally, Curless and Levoy [15] propose a volumetric method for fitting a signed distance function to a set of range images.

Principles

Setup Overview

Our setup consists of an orthographic camera and a turntable. Depth discontinuities are obtained by the use of multiple flashes distributed around the lenses (Crispell et al. [5]). By tracking depth discontinuities over a sequence of images we obtain the equation for v , the line tangent to the 3-D surface.



3D Reconstruction of Point from Depth Discontinuity

We can express the coordinates of the 3-D point p touched by the tangent through the following equation

$$p = q + \lambda v,$$

$$\lambda = -\frac{n^T q}{n^T v} \quad (1)$$

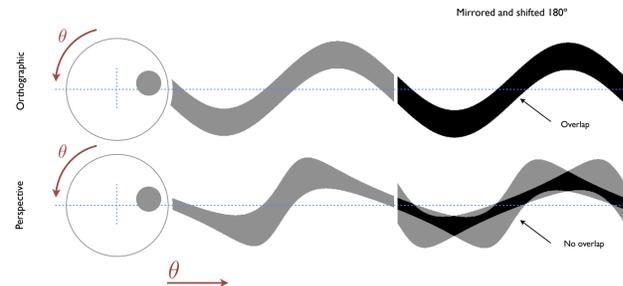
where q is the camera center of projection (a circular trajectory for the turntable case) and λ is the point depth with respect to the camera (our unknown). By manipulating this equation, and incorporating the surface normal n , it can be shown (Cipolla and Giblin [8]) that

Curve Symmetries in Orthographic Epipolar Slice Images

If we consider the case of a cylinder of radius r , centered at a distance R from the turntable rotation axis, the curves it describes on the EPI can be expressed as

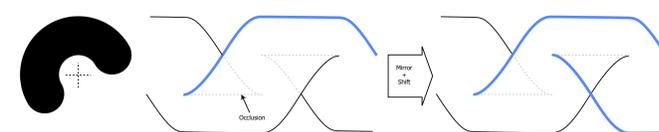
$$u_{\pm}(\theta) = \alpha R \cos(\theta) \pm \alpha r,$$

which exhibits interesting symmetries. If reflected about the rotation axis and shifted 180° we obtain the same curve. This symmetry is not present in the perspective case due to compression of the curve when the point is close to the camera and expansion when they are far from the camera. Locally, the surface boundary can be approximated by an osculating circle; as a result, we expect the EPI image to consist of intersecting sinusoidal contours.



Occlusion in the Epipolar Slice Images

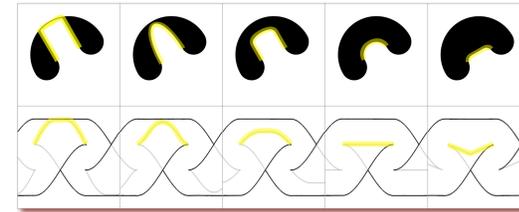
Occlusions limit the extent of u_{\pm} that can be recovered. By using the symmetry property present in orthographic projections we can extend the curve up to concavities.



Concavities and the Epipolar Slice Images

A set of cusps will be present in the EPI containing both visible and hidden depth discontinuities. These cusps correspond to positions where locally convex visible points on the surface transition into locally concave hidden points. Every concavity in the primal curve maps to a "fishtail" structure in the EPI. In other words, for a simple concavity, a T-junction will be present in the EPI, corresponding to a point of

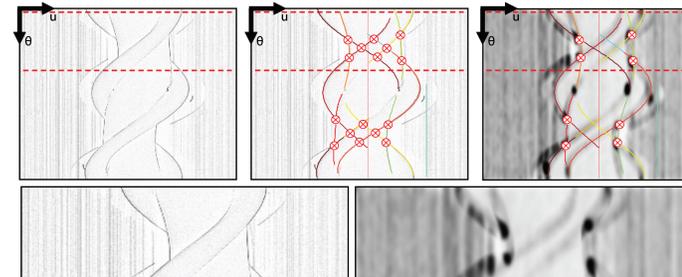
bitangency between the viewing ray and the surface [1]. As a result, T-junctions can be used to isolate corresponding points on either side of a concavity. We note that higher-order junctions, while improbable, correspond to points of multiple-tangency and can be processed in a similar manner. The proposed surface reconstruction method cannot recover locally concave points in deep concavities. As a result, a shape completion method is required to fill remaining gaps.



Implementation

Reconstruction Algorithm

1. To each EPI ridge fit a trigonometric polynomial $u_i(\theta)$
2. Form the additional set $\{\tilde{u}_i(\theta)\}$ of EPI curves by shifting each tracked contour by 180° and reflecting about the rotation axis.
3. Join curves in $\{\tilde{u}_i(\theta)\}$ and $\{u_i(\theta)\}$ if there is significant overlap, call this $\{u'_i(\theta)\}$.
4. Searching for T-junctions between curves in $\{u'_i(\theta)\}$.
5. To "hallucinate" the points inside concavities we fit a cubic Hermite interpolating polynomial to both sides of the concavity.
6. Points modeled by $\{u'_i(\theta)\}$ are reconstructed using equation (1).



Orthographic Camera Calibration

The method Zhang [17] proposes does not work with orthographic cameras since the intrinsic parameters matrix is singular for this case. We propose a similar technique, in which intrinsic and extrinsic parameters are separately estimated from multiple images of a planar checkerboard pattern.

We can express the projection of a point on a checkerboard (with coordinate $z = 0$) as

$$\tilde{p} = \begin{bmatrix} \alpha & \gamma & 0 & 0 \\ \beta & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \tilde{P}$$

$$\tilde{p} = \begin{bmatrix} \alpha & \gamma & 0 \\ \beta & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ 1 & 1 & 1 \end{bmatrix} \tilde{P} = \begin{bmatrix} \alpha & \gamma & 0 \\ \beta & 0 & 1 \end{bmatrix} \begin{bmatrix} R_s & T_s \\ 0 & 1 \end{bmatrix} \tilde{P}$$

As was done in Zhang we compute a homography between the points on the image and the true coordinates

$$\tilde{p} = H\tilde{P} \Rightarrow H = K_s E_s$$

The expression on the right can be manipulated to obtain

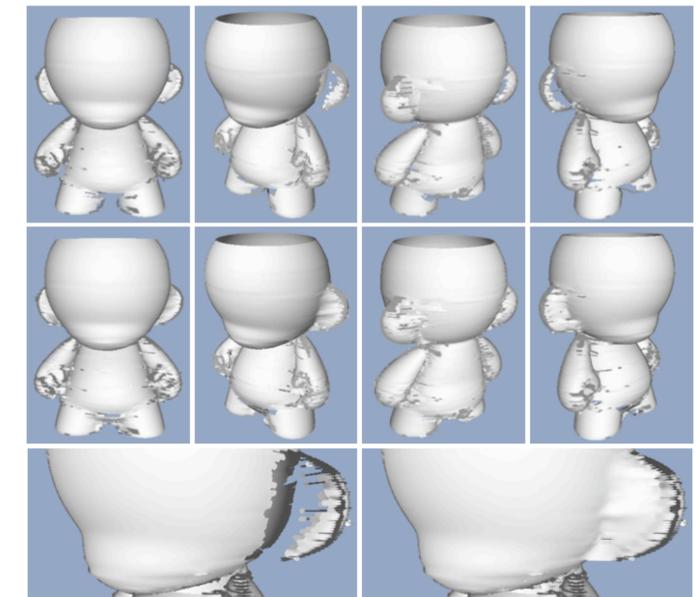
$$H^T K_s^{-T} K_s^{-1} H = \begin{bmatrix} R_s^T R_s & R_s^T T_s \\ T_s^T R_s & T_s^T T_s \end{bmatrix}$$

We can obtain the camera's intrinsic parameters by noting that

$$\det(R_s^T R_s - I_2) = 0$$

Results

A typical capture sequence consists of 670 viewpoints, separated by a rotation of approximated 0.527 degrees. Each images had a resolution of 1600x1200 at 8bits for each color channel. For each viewpoint four images are recorded in which the scene is sequentially illuminated by the top, bottom, left, and right flashes.



References

- [1] R. C. Bolles, H. H. Baker, and D. H. Marimon. "Epipolar-plane image analysis: An approach to determining structure from motion." in *IJCV*, 1987.
- [2] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. "A comparison and evaluation of multi-view stereo reconstruction algorithms." 2006.
- [3] H. Baker and R. Bolles. "Generalizing epipolar-plane image analysis on the spatiotemporal surface." in *IJCV*, 1989.
- [4] I. Feldmann, P. Eisert, and P. Kauff. "Extension of epipolar image analysis to circular camera movements." in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 3, 2003.
- [5] D. Crispell, D. Lanman, P. Sibley, Y. Zhao, and G. Taubin. "Beyond silhouettes: Surface reconstruction using multi-flash photography." in *International Symposium on 3D Data Processing, Visualization and Transmission*, vol. 1, 2006.
- [6] R. Woodham. "Photometric method for determining surface orientation from multiple images." *Optical engineering*, vol. 19, no. 1, 1980.
- [7] R. Raskar, K.-H. Tan, R. Feris, J. Yu, and M. Turk. "Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging." *ACM Trans. Graph.*, vol. 23, no. 3, 2004.
- [8] R. Cipolla and P. Giblin. *Visual Motion of Curves and Surfaces*. Cambridge University Press, 2000.
- [9] M. Watanabe and S. K. Nayar. "Telecentric optics for computational vision." in *ECCV*, 1996.
- [10] S. Ullman. "Filling-in the gaps: The shape of subjective contours and a model for their generation." *Biological Cybernetics*, vol. 25, no. 1, 1978.
- [11] D. Crispell, D. Lanman, P. Sibley, Y. Zhao, and G. Taubin. "Mathematical typography." *American Mathematical Society*, vol. 1, no. 2, 1979.
- [12] B. Kimia, I. Frankel, and A. Popescu. "Euler spiral for shape completion." in *IJCV*, 2003.
- [13] H. Horn. "The curve of least energy." *ACM Trans. Math. Softw.*, vol. 9, no. 4, 1983.
- [14] M. Pauly, N. J. Mitra, J. Wallner, H. Pottmann, and L. Guibas. "Discovering structural regularity in 3D geometry." *ACM Trans. Graph.*, vol. 27, no. 3, 2008.
- [15] B. Curless and M. Levoy. "A volumetric method for building complex models from range images." in *SIGGRAPH*, 1996.
- [16] N. Apostoloff and A. Fitzgibbon. "Learning spatiotemporal t-junctions for occlusion detection." in *CVPR*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 553-559.
- [17] Z. Zhang. "Flexible camera calibration by viewing a plane from unknown orientations." in *ICCV*, 1999.