SIGGRAPH 99 Course on
3D Photography

Passive 3D Photography

Steve Seitz
Carnegie Mellon University

http://www.cs.cmu.edu/~seitz

Talk Outline

1. *Visual Cues*
2. *Classical Vision Algorithms*
3. *State of the Art (video)*

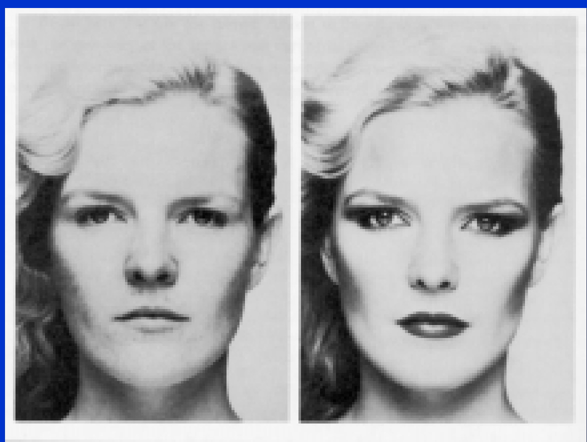
Visual Cues

Motion

Visual Cues

Motion

Shading



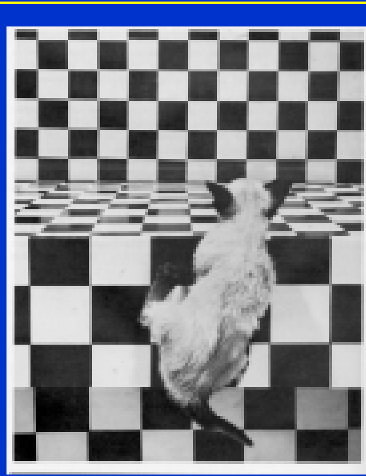
Merle Norman Cosmetics, Los Angeles

Visual Cues

Motion

Shading

Texture



The Visual Cliff, by William Vandivert, 1960

Visual Cues

Motion

Shading

Texture

Focus

From *The Art of Photography*, Canon

Visual Cues

Motion

Shading

Texture

Focus

Others:

- Highlights
- Shadows
- Silhouettes
- Inter-reflections
- Symmetry
- Light Polarization
- ...

Reconstruction Algorithms

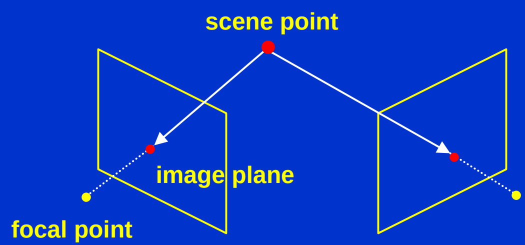
Shape From X

- ✔ Stereo (shape from parallax)
- ✔ Structure from motion
- Shape from shading
- Photometric stereo
- Shape from texture
- Shape from focus/defocus
- Shape from silhouettes, ...

Stereo

The Stereo Problem

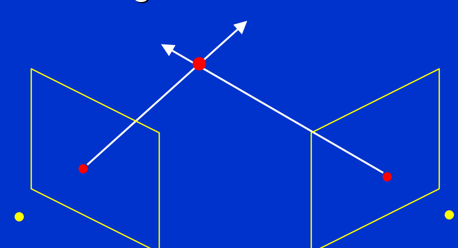
- Reconstruct scene geometry from two or more *calibrated* images



Stereo

The Stereo Problem

- Reconstruct scene geometry from two or more *calibrated* images



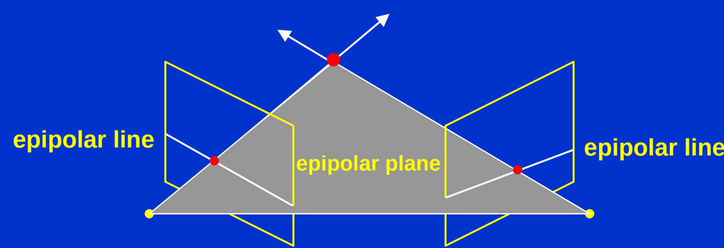
Basic Principle: Triangulation

- Gives reconstruction as intersection of two rays
- Requires *point correspondence*

Stereo Correspondence

Determine Pixel Correspondence

- Pairs of points that correspond to same scene point



Epipolar Constraint

- Reduces correspondence problem to 1D search along *conjugate epipolar lines*

Stereo Matching Algorithms

Match Pixels in Conjugate Epipolar Lines

- Assume color of point does not change
- Pitfalls
  - > specularities (non-Lambertian surfaces)
  - > ambiguity (low-contrast regions)
  - > missing data (occlusions)
  - > intensity error (quantization, sensor error)
  - > position error (camera calibration)
- Numerous approaches
  - > winner-take all
  - > dynamic programming [Ohta 85]
  - > smoothness functionals
  - > more images (trinocular, N-ocular) [Okutomi 93]

Structure from Motion

The SFM Problem

- **Reconstruct scene geometry and camera motion from two or more images**

Assume

- **Pixel correspondence**
  > via tracking
- **Projection model**
  > classic methods are orthographic

Orthographic Projection

$$\underset{2\times1}{\mathbf{u}} = \underset{2\times3}{\Pi}\ \underset{3\times1}{\mathbf{X}} + \underset{2\times1}{\mathbf{t}}$$

image point (u), projection matrix (Π), scene point (X), image offset (t)

Trick

- **Choose scene origin to be centroid of 3D points**
- **Choose image origins to be centroid of 2D points**
- **Allows us to drop the camera translation:**

$$\underset{2\times1}{\mathbf{u}} = \underset{2\times3}{\Pi}\ \underset{3\times1}{\mathbf{X}}$$

Shape by Factorization [Tomasi & Kanade, 92]

projection of *n* features in one image:

$$\underset{2\times n}{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}} = \underset{2\times n}{\Pi} \underset{3\times n}{\begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_n \end{bmatrix}}$$

projection of *n* features in *f* images

$$\underset{2f\times n}{\begin{bmatrix} \mathbf{u}_1^1 & \mathbf{u}_2^1 & \cdots & \mathbf{u}_n^1 \\ \mathbf{u}_1^2 & \mathbf{u}_2^2 & \cdots & \mathbf{u}_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_1^f & \mathbf{u}_2^f & \cdots & \mathbf{u}_n^f \end{bmatrix}} = \underset{2f\times 3}{\begin{bmatrix} \Pi^1 \\ \Pi^2 \\ \vdots \\ \Pi^f \end{bmatrix}} \underset{3\times n}{\begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_n \end{bmatrix}}$$

W measurement **M** motion **S** shape

Shape by Factorization [Tomasi & Kanade, 92]

known — $\underset{2f\times n}{\mathbf{W}} = \underset{2f\times 3}{\mathbf{M}} \underset{3\times n}{\mathbf{S}}$ — solve for

Factorization Technique

- *W* is at most rank 3 (assuming no noise)
- We can use *singular value decomposition* to factor W:

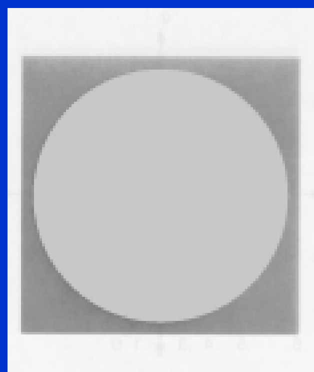
$$\underset{2f\times n}{\mathbf{W}} = \underset{2f\times 3}{\mathbf{M}'} \underset{3\times n}{\mathbf{S}'}$$

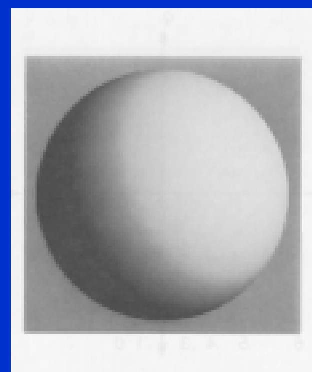
- *S'* differs from *S* by a linear transformation *A*:

$$\mathbf{W} = \mathbf{M}'\mathbf{S}' = (\mathbf{M}\mathbf{A}^{-1})(\mathbf{A}\mathbf{S})$$

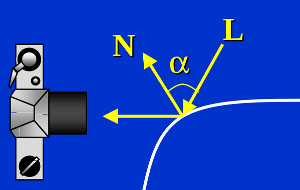
- Solve for *A* by enforcing constraints on *M*

Shape from Shading





Shape from Shading [Horn, 1970]

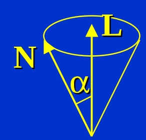


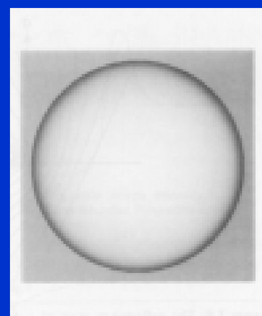
Classical Approach

- **Suppose reflected light depends only on α**

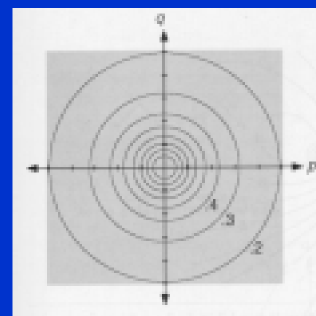
$$radiance = k \cos\alpha$$

The Reflectance Map





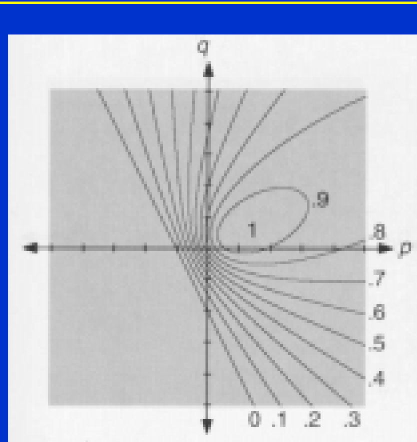
Image

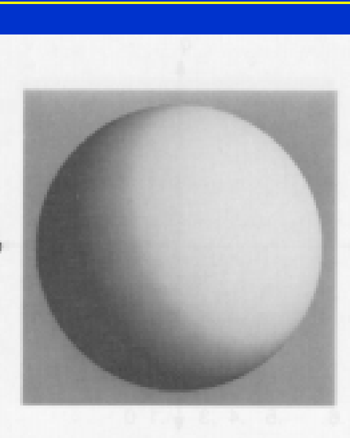


Reflectance Map: R

$$\mathbf{N} = [p \quad q \quad -1]$$

The Reflectance Map





Reflectance Map

Image

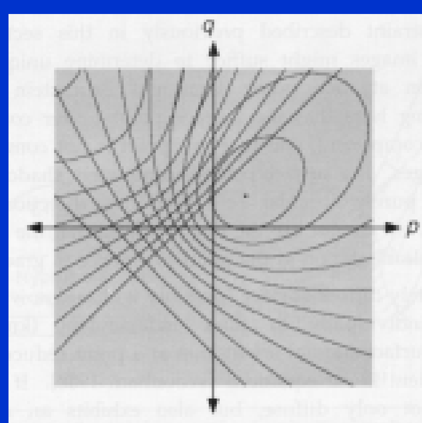
Finding a Unique Solution

Three Approaches

- **Characteristic Strip Method [Horn, 77]**
  - > select a few points where normal is known
  - > grow solution by moving direction of $\nabla\mathbf{R}$
- **Variational Method [Ikeuchi & Horn, 81]**
  - > start with an initial guess of surface shape
  - > define energy function
  - > refine to minimize energy function
- **Photometric Stereo [Woodham 80]**
  - > use more images

Photometric Stereo

Two Images Under Different Lighting



Need Three Images for Unique Solution

Photometric Stereo: Matrix Formulation

Write Equations in Matrix Form

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{L}}_1^T \\ \hat{\mathbf{L}}_2^T \\ \hat{\mathbf{L}}_3^T \end{bmatrix} \bullet \begin{bmatrix} k\hat{\mathbf{N}} \\ k\hat{\mathbf{N}} \\ k\hat{\mathbf{N}} \end{bmatrix}$$

$$\mathbf{I}_{3\times 1} \quad \mathbf{L}_{3\times 3} \quad \tilde{\mathbf{N}}_{3\times 1}$$

$$\tilde{\mathbf{N}} = \mathbf{L}_1^{-1}\,\mathbf{I}$$

$$k = \left\|\tilde{\mathbf{N}}\right\|$$

Advantage:

- Can solve for variable reflectance *k*

Resources

Computer Vision Home Page

- http://www.cs.cmu.edu/afs/cs/project/cil/ftp/html/vision.html

Computer Vision Textbooks

- D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, 1982.
- O. Faugeras, *Three-Dimensional Computer Vision*, MIT Press, 1993.
- B. K. P. Horn, *Robot Vision*, McGraw-Hill, 1986.
- R. Jain, R. Kasturi and B. G. Schunck, *Machine Vision*, McGraw-Hill, 1995.
- R. Klette, K. Schluns and A. Koschan, *Computer Vision: Three-Dimensional Data from Images*, Springer-Verlag, 1998.
- V. S. Nalwa, *A Guided Tour of Computer Vision*, Addison-Wesley, 1993.
- M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis, and Machine Vision*, Brooks/Cole Publishing, 1999.
- E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998.
- D. Marr, *Vision*, Freeman, 1982.
- J. Koenderink, *Solid Shape*, MIT Press, 1990.

Bibliography

Stereo

- Yuichi Ohta & Takeo Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming", IEEE Trans. on Pattern Analysis and Machine Intelligence, 7(2), 1985, pp. 129-154.
- Masatoshi Okutomi & Takeo Kanade, "A Multiple-Baseline Stereo", IEEE Trans. on Pattern Analysis and Machine Intelligence", 15(4), 353-363, 1985.

Structure-from-Motion

- Carlo Tomasi & Takeo Kanade, "Shape and Motion from Image Streams Under Orthography: A Factorization Method", Int. Journal of Computer Vision, 9(2), 1992, pp. 137-154.

Shape from Shading

- B. Horn and M. Brooks, "Shape from Shading", 1989, MIT Press.
- L. Wolff, S. Shafer, and G. E. Healey, "Physics-Based Vision: Shape Recovery", 1992, Jones and Bartlett.
- R. J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images", Optical Engineering, 1980, pp. 139-144.

Video