Shape From X

• Recovery of 3D (shape) from one or two (2D images).
Shape From X

- Stereo
- Motion
- Shading
- Photometric Stereo
- Texture
- Contours
- Silhouettes
- Defocus

Copyright Mubarak Shah 2003
Shape from Stereo

(a) 
(b) 
(c)
Shape from Shading
Lambertian Model

\[ f(x, y) = n \cdot L = \left( n_x, n_y, n_z \right) \cdot (l_x, l_y, l_z) \]

\[ f(x, y) = n \cdot L = \left( \frac{1}{\sqrt{p^2 + q^2 + 1}} (-p, -q, 1) \right) \cdot (l_x, l_y, l_z) \]
Sphere

\[ z = \sqrt{R^2 - x^2 - y^2} \]

\[ p = \frac{\partial z}{\partial x} = -\frac{x}{z} \]

\[ q = \frac{\partial z}{\partial y} = -\frac{y}{z} \]

\[ (n_x, n_y, n_z) = \frac{1}{R} (x, y, z) \]
Sphere

Copyright Mubarak Shah 2003
Vase

(1, 0, 1)

(-1, 1, 1)

(-1, -1, 1)
Shape from Shading

**Figure 12.2** Synthetic shape from shading (Zhang, Tsai, Cryer *et al.*, 1999) © 1999 IEEE: shaded images, (a–b) with light from in front (0, 0, 1) and (c–d) with light the front right (1, 0, 1); (e–f) corresponding shape from shading reconstructions using the technique of Tsai and Shah (1994).
Photometric Stereo

- Use images taken under different illumination conditions to recover 3D shape.
Photometric Stereo

\[ I(x, y) = B \vec{N} \cdot \vec{S}, \]

B= Albedo
N=surface normal
S=light source direction

Take Three images under different light sources

\[ I_1(x, y) = B \vec{S}^1 \cdot \vec{N}, \]
\[ I_2(x, y) = B \vec{S}^2 \cdot \vec{N}, \]
\[ I_3(x, y) = B \vec{S}^3 \cdot \vec{N}. \]
Photometric Stereo

\[
B \begin{bmatrix}
  s_x^1 & s_y^1 & s_z^1 \\
  s_x^2 & s_y^2 & s_z^2 \\
  s_x^3 & s_y^3 & s_z^3 \\
\end{bmatrix}
\begin{bmatrix}
  n_x \\
n_y \\
n_z \\
\end{bmatrix} = \begin{bmatrix}
  I_1 \\
  I_2 \\
  I_3 \\
\end{bmatrix}.
\]

\[
BAN = I
\]

\[
B\vec{N} = A^{-1}I,
\]

\[
B\|\vec{N}\| = \|A^{-1}I\|
\]

\[
\|n\| = 1
\]

\[
B = \|A^{-1}I\|.
\]

\[
\vec{N} = \frac{A^{-1}I}{B}
\]
Results

Figure 6.13: (a)-(c) Input images with light sources $(1, 0, 1)$, $(-1, 1, 1)$, $(-1, -1, 1)$ (d) Intensity image reconstructed using the surface normals computed by photometric stereo.
Structure from Motion

Copyright Mubarak Shah 2003
Shape from Motion: Moving Light Display
Shape from Motion
Problem

• Given optical flow or point correspondences, compute 3-D motion (translation and rotation) and shape (depth).
Structure from Motion

- S. Ullman
- Hanson & Riseman
- Webb & Aggarwal
- T. Huang
- Heeger and Jepson
- Chellappa
- Faugeras
- Zisserman
- Kanade
- Pentland
- Van Gool
- Pollefeys
- Seitz & Szeliski
- Shahsua
- Irani
- Vidal & Yi Ma
- Medioni
- Fleet
- Tian & Shah

Copyright Mubarak Shah 2003
Tomasi and Kanade
Factorization
Orthographic Projection
Assumptions

• The camera model is orthographic.
• The positions of “p” points in “F” frames (F>=3), which are not all coplanar, and have been tracked.
• The entire sequence has been acquired before starting (batch mode).
• Camera calibration not needed, if we accept 3D points up to a scale factor.
Input

Images

KLT Tracks
Feature Points

Image points \( \{ (u_{fp}, v_{fp}) \mid f = 1, \ldots, F, p = 1, \ldots, P \} \)

\[
W = \begin{bmatrix}
  u_{11} & \cdots & u_{1P} \\
  \vdots & & \vdots \\
  u_{F1} & \cdots & u_{FP} \\
  v_{11} & \cdots & v_{1P} \\
  \vdots & & \vdots \\
  v_{F1} & \cdots & v_{FP}
\end{bmatrix}
\]
Mean Normalize Feature Points

\[ a_f = \frac{1}{P} \sum_{p=1}^{P} u_p \]
\[ b_f = \frac{1}{P} \sum_{p=1}^{P} v_p \]

\[ \tilde{u}_{fP} = u_{fP} - a_{fP} \quad (A) \]
\[ \tilde{v}_{fP} = v_{fP} - b_{fP} \]
Orthographic Projection
Orthographic Projection: Feature Points

\[ s_p = (X_p, Y_p, Z_p) \]  

3D world point

\[ u_{fP} = i^T_f (s_p - t_f) \]  

Orthographic projection

\[ v_{fP} = j^T_f (s_p - t_f) \]

\[ k_f = i_f \times j_f \]

i, j, k are unit vectors along X, Y, Z

Copyright Mubarak Shah 2003
Orthographic Projection: Mean Normalized Feature Points

\[ \tilde{\mathbf{u}}_{fp} = \mathbf{u}_{fp} - \alpha_f \]

\[ \alpha_f = \frac{1}{P} \sum_{n=1}^{P} u_p \]

\[ = i_f^T (s_p - t_f) - \]

\[ = i_f^T \left[ s_p - \frac{1}{P} \sum_{q=1}^{P} s_q \right] \]

\[ = i_f^T s_P \]

If Origin of world is at the centroid of object points
Orthographic Projection: Mean Normalized Feature Points

\[ \tilde{\mathbf{u}}_{fp} = \mathbf{u}_{fp} - \alpha_f \]

\[ \alpha_f = \frac{1}{P} \sum_{p=1}^{P} u_p \]

\[ = i_f^T (s_p - t_f) - \frac{1}{P} \sum_{q=1}^{P} i_f^T (s_q - t_f) \]

\[ = i_f^T \left[ s_p - \frac{1}{P} \sum_{q=1}^{P} s_q \right] \]

If Origin of world is at the centroid of object points

Copyright Mubarak Shah 2003
Registered Measurement Matrix

\[ \tilde{u}_{fP} = i_f^T S_P \]

\[ \tilde{v}_{fP} = j_f^T S_P \]

\[ \tilde{W} = \begin{bmatrix} \tilde{U} \\ - \tilde{V} \end{bmatrix} \]
Registered Measurement Matrix

\[
\begin{align*}
\tilde{u}_{fP} &= i_f^T S_P \\
\tilde{v}_{fP} &= j_f^T S_P \\
\tilde{W} &= \begin{bmatrix}
\tilde{U} \\
- \tilde{V}
\end{bmatrix}
\end{align*}
\]

\[
\tilde{W} = \begin{bmatrix}
i_1^T \\
\vdots \\
i_F^T \\
j_1^T \\
\vdots \\
j_F^T
\end{bmatrix}
\begin{bmatrix}
s_1 \\
\vdots \\
s_P
\end{bmatrix} = RS
\]

Rank of \(S\) is 3, because points in 3D space are not Co-planar.

Copyright Mubarak Shah 2003
Rank Theorem

Without noise, the registered measurement matrix $\tilde{W}$ is at most of rank three.

$$\tilde{W} = \begin{bmatrix} i_1^T & \cdots & i_F^T \\ j_1^T & \cdots & j_F^T \end{bmatrix} [s_1 \quad \cdots \quad s_p] = RS$$
How to find Translation?

\[
\tilde{u}_{fp} = u_{fp} - a_f \quad \text{(A)}
\]

\[
u_{fp} = \tilde{u}_{fp} + a_f \quad \tilde{u}_{fp} = i_f^T s_p \quad \text{(B)}
\]

\[
u_{fp} = i_f^T s_p + a_f \quad u_{fp} = i_f^T (s_p - t_f) \quad \text{(C)}
\]

\[
a_f = -t_f^T i_f \quad \text{(D)}
\]

\(a_f\) is projection of camera translation along x-axis.
Registered Measurement Matrix in terms of Motion and Shape

\[ \tilde{u}_{fP} = i_f^T s_P \]  \hspace{1cm} (B) \hspace{1cm} \tilde{v}_{fP} = j_f^T s_P \]

\[ \tilde{W} = \begin{bmatrix} \tilde{U} \\ - \tilde{V} \end{bmatrix} \]

\[ \tilde{W} = \begin{bmatrix} i_1^T \\ \vdots \\ i_F^T \\ j_1^T \\ \vdots \\ j_F^T \end{bmatrix} \begin{bmatrix} s_1 & \cdots & s_P \end{bmatrix} = RS \]

3XP

Rank of S is 3, because points in 3D space are not Co-planar
Measurement Matrix in terms of Motion and Shape

\[ u_{fp} = i_f^T (s_p - t_f) \]  From (C) \[ a_f = -t_f i_f^T \]  (D)

\[ u_{fp} = i_f s_p + a_f \]

\[ v_{fp} = j_f s_p + b_f \]

\[ W = RS + te_p^T \]

\[ t = (a_1, \ldots, a_f, b_1, \ldots, b_f)^T \]

\[ e_p^T = (1, \ldots, 1) \]

Copyright Mubarak Shah 2003
Measurement Matrix in terms of Motion and Shape

\[ u_{fp} = i_f s_p + a_f \quad v_{fp} = j_f s_p + b_f \]

\[ W = RS + te_p^T \]

\[ W = \begin{bmatrix} u_{11} & \cdots & u_{1p} \\ \vdots & & \vdots \\ u_{F1} & \cdots & u_{FP} \\ v_{11} & \cdots & v_{1p} \\ \vdots & & \vdots \\ v_{F1} & \cdots & v_{FP} \end{bmatrix} = \begin{bmatrix} i_1 & j_1 & k_1 \\ \vdots & \vdots & \vdots \\ X_1 & \cdots & X_P \\ Y_1 & \cdots & Y_P \\ Z_1 & \cdots & Z_P \end{bmatrix} + \begin{bmatrix} a_1 \\ \vdots \\ a_F \\ b_1 \\ \vdots \\ b_F \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \]

Copyright Mubarak Shah 2003
Projected camera translation can be computed:

\[- i_f^T t_f = a_f = \frac{1}{P} \sum_{p=1}^{P} u_p \]

\[- j_f^T t_f = b_f = \frac{1}{P} \sum_{p=1}^{P} v_p \]
Noisy Measurements

- Without noise, the matrix $\tilde{W}$ must be at most of rank 3. When noise corrupts the images, however, $\tilde{W}$ will not be rank 3. Rank theorem can be extended to the case of noisy measurements.
Approximate Rank

\[ \tilde{W} = O_1 \Sigma O_2 \]

SVD

Copyright  Mubarak Shah 2003
Singular Value Decomposition (SVD)

• For some linear systems $Ax=b$, Gaussian Elimination or LU decomposition does not work, because matrix $A$ is singular, or very close to singular. SVD will not only diagnose for you, but it will solve it.
Singular Value Decomposition (SVD)

Theorem: Any $m$ by $n$ matrix $A$, for which $m \geq n$, can be written as

$$A = O_1 \Sigma O_2$$

where $\Sigma$ is diagonal, $O_1$ and $O_2$ are orthogonal,

$$O_1^T O_1 = O_2^T O_2 = I$$
Singular Value Decomposition (SVD)

If $A$ is square, then $O_1, O_2, \Sigma$ are all square.

\[
O_1^{-1} = O_1^T \\
O_2^{-1} = O_2^T \\
\Sigma^{-1} = \text{diag}\left(\frac{1}{w_j}\right)
\]

\[
A = O_1 \Sigma O_2
\]
Approximate Rank

\[ \tilde{W} = O_1 \Sigma O_2 \]

\[ O_1 = \begin{bmatrix} O_1' & O_1'' \end{bmatrix} \]

\[ \Sigma = \begin{bmatrix} \Sigma' & 0 \\ 0 & \Sigma'' \end{bmatrix} \]

\[ O_2 = \begin{bmatrix} O_2' \\ O_2'' \end{bmatrix} \]

Copyright  Mubarak Shah 2003
Approximate Rank

\[ \hat{W} = O_1 \Sigma O_2 = O'_1 \Sigma' O'_2 + O''_1 \Sigma'' O''_2 \]

\[ \hat{W} = O'_1 \Sigma' O'_2 \]

The best rank 3 approximation to the ideal registered measurement matrix.
Rank Theorem for noisy measurement

The best possible shape and rotation estimate is obtained by considering only 3 greatest singular values of $\tilde{W}$ together with the corresponding left, right eigenvectors.
Approximate Rank

\[ \hat{R} = O_1' \left[ \Sigma' \right]^{1/2} \]
Approximate Rotation matrix

\[ \hat{S} = \left[ \Sigma' \right]^{1/2} O_2' \]
Approximate Shape matrix

\[ \hat{W} = \hat{R} \hat{S} \]
This decomposition is not unique

\[ \hat{W} = (\hat{R} Q)(Q^{-1} \hat{S}) \]
\( Q \) is any 3X3 invertible matrix

Copyright  Mubarak Shah 2003
Approximate Rank

\[ R = \hat{R}Q \]
\[ S = Q^{-1}\hat{S} \]

\( R \) and \( S \) are linear transformation of approximate Rotation and shape matrices.

How to determine \( Q \)?

\[ \hat{i}_f^TQQ^T\hat{i}_f = 1 \]
\[ \hat{j}_f^TQQ^T\hat{j}_f = 1 \]
\[ \hat{i}_f^TQQ^T\hat{j}_f = 0 \]

Rows of rotation matrix are unit vectors and orthogonal.
How to determine $Q$: Newton’s Method

\[ f_1(q) = \hat{i}_1^T QQ^T \hat{i}_1^T - 1 = 0 \]
\[ f_2(q) = \hat{j}_1^T QQ^T \hat{j}_1^T - 1 = 0 \]
\[ f_3(q) = \hat{i}_1^T QQ^T \hat{j}_1^T = 0 \]
\[ \vdots \]
\[ f_{3F-2}(q) = \hat{i}_f^T QQ^T \hat{i}_f^T - 1 = 0 \]
\[ f_{3F-1}(q) = \hat{j}_f^T QQ^T \hat{j}_f^T - 1 = 0 \]
\[ f_{3F}(q) = \hat{i}_f^T QQ^T \hat{j}_f^T = 0 \]

\[ M\Delta q = \varepsilon \]

\[ \Delta q = [\Delta q_1, \ldots, \Delta q_9] \]

\[ M_{ij} = \frac{\partial f_i}{\partial q_j} \]

Jacobian

\[ \varepsilon \text{ is error} \]
Algorithm

• Compute SVD of \( \tilde{W} = O_1 \Sigma O_2 \)

• define \( \hat{R} = O_1' [\Sigma']^{\frac{1}{2}} \)

• Compute \( \hat{S} = [\Sigma']^{\frac{1}{2}} O_2' \)

• Compute \( Q \)

• Compute \( R = \hat{R} Q \)

\( S = Q^{-1} \hat{S} \)
Hotel Sequence
Results (rotations)
Selected Features
Comparison
House Sequence
Reconstructed Walls
Paper