Slide 1

Introduction

The goal of 3D reconstruction is to recover the 3D **properties** of a geometric entity from its 2D images. Depending on the types of geometric entities, the 3D properties may include 3D coordinates of a 3D point or its depth (z coordinate) or in general the 3D shape (as characterized by the orientation of each point) of an object.

Slide 2

Introduction (cont'd)

The shape (range) of an object can be represented either as a list of 3-D coordinates in a given reference frame or a matrix of depth values along the directions of row and column image axes or a list of orientations for each 3D point. The image data representing the shape of a 3D object are also referred to as *range* image, *depth* image, *depth* maps, *surface profiles*, or *2.5-D images.*

Slide 3

Techniques

- Shape from monocular images.
- Shape from multiple images (stereo and motion).
- Shape from range sensors (e.g. RADAR or LADAR).

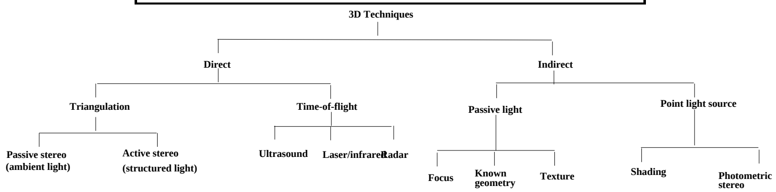
Slide 4

Range from Intensity Images

Collectively referred to as **Shape from X** techniques, the techniques for recovering 3D shape from moncular intensity images infer range data from monocular images in conjunction with other visual cues or geometric properties. Stereo approach recovers range data using multiple images via triangulation.

Slide 5

A Taxonomy of 3D Techniques



Slide 6

Stereo Vision

Stereo vision is a technique for the reconstruction of the three-dimensional description of a scene from images observed from multiple viewpoints.

Slide 7

Stereo Vision (cont'd)

Passive stereo (for a review see Poggio [?])

Active stereo

Slide 8

Passive Stereo

- Token detection.
- Token matching (e.g. pointwise matching).
- 3D reconstruction using the matched tokens.

Slide 9

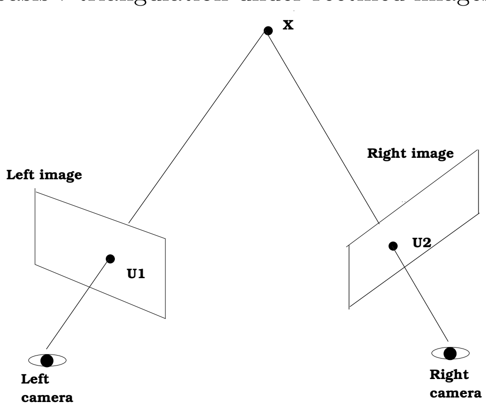
Different tokens

- Points
- Lines
- Conic curves

Slide 10

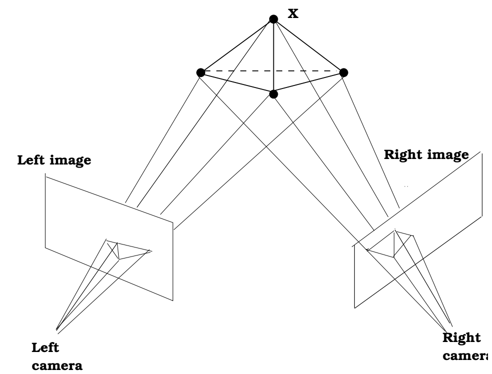
Passive Stereo Using 2D/3D Points

- Theoretical basis : triangulation under rectified images



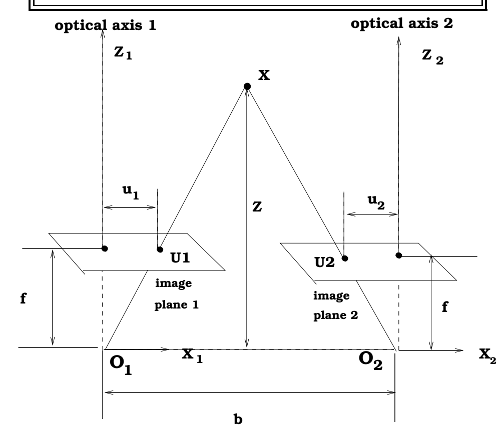
Slide 11

3D Triangulation Example



Slide 12

Theoretical basis : triangulation



Since triangle $XU1U2$ is similar to triangle XO_1O_2, we have

Slide 13

$$\frac{Z-f}{Z} = \frac{b-u_1+u_2}{b}$$

Hence,

$$Z = \frac{fb}{u_1-u_2}$$

where b is called the **base distance**, $u_1 - u_2$ the **disparity**. It is clear disparity is **inversely** proportional to depth. As depth approaches infinity, disparity approaches zero.

Slide 14

Major Issues

- Establish correspondences
- Reconstruct 3D points from matched 2D points via triangulation.

Slide 15

Establishing Correspondences

Establishing correspondences amounts to matching points from one image to points in another image such that each matched pair is generated by the same 3D point.

Slide 16

Establishing Correspondences: an example

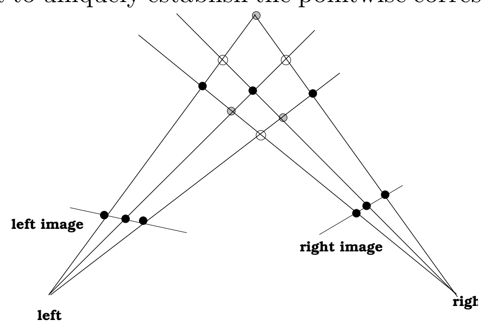
see Figure 7.1 of Trucco's book

Slide 17

Establishing Correspondences (cont'd)

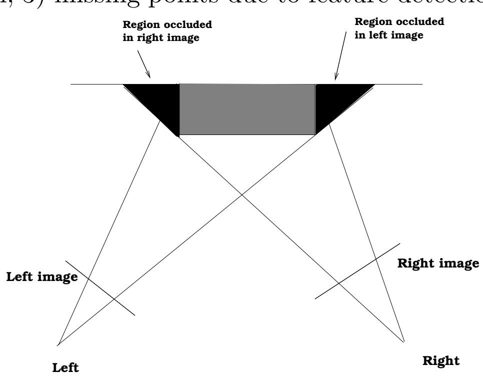
Establishing correspondences is difficult since

- the information associated with each image point is often not sufficient to uniquely establish the pointwise correspondence.



- Image points in one image may not have corresponding points in another image due to 1) two cameras have different views; 2) occlusion; 3) missing points due to feature detection techniques.

Slide 18



Slide 19

Reducing baseline distance b (narrow-angle-stereo) can alleviate the correspondence and occlusion problem, but it may lead to less accurate depth estimate.

Let $d = u_1 - u_2$ be the disparity, then we have

$$Z = \frac{bf}{d}$$

Linearizing both sides of the above equation yields

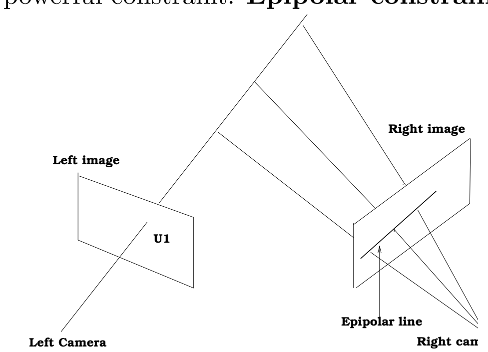
$$\Delta Z = \frac{f \Delta b}{d} - \frac{f b \Delta d}{d^2}$$

It is clear that as b decreases, ΔZ increases.

Slide 20

Epipolar Constraint

Pointwise matching can be established more efficiently using a simple yet powerful constraint: **Epipolar constraint**



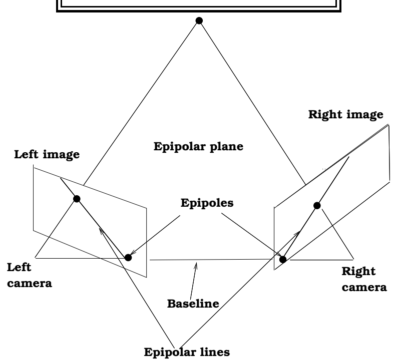
Slide 21

Epipolar Constraint (cont'd)

Epipolar constraint says that given U_1 from the left image, its corresponding point on the right image must lie on the epipolar line. Epipolar line is the projection on the right image by the 3D line going through U_1 and the left camera center. This effectively reduces the point search from 2D to 1D, substantially reducing the search time.

Slide 22

Epipolar Geometry

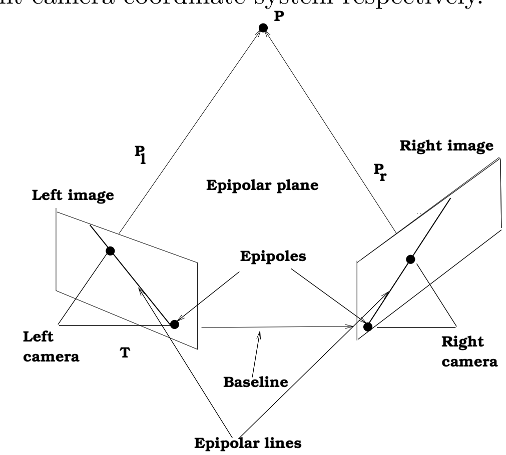


All epipolar lines go through the epipole. The epipole on the left (right) image is the image of the optical center of the right (left) camera. In rectified stereo images, epipolar lines are parallel and conjugate epipolar lines are collinear.

Slide 23

Computing Epipolar Geometry

Let P be a 3D point and P_l and P_r be the coordinates of P in the left and right camera coordinate system respectively.



Slide 24

Then we have

$$P_l = RP_r + T$$

or

$$P_r = R^t(P_l - T)$$

where R and T specify the relative orientation of the right camera frame with respect to the left one. The coplanar constraint on P_l, P_r, and T leads to

$$(T \times P_l)^t(P_l - T) = 0$$

Since $P_l - T = RP_r$, we have

$$(T \times P_l)^t RP_r = 0$$

Since $(T \times P_l)$=S P_l, where S a skew matrix

Slide 25

$$S = \begin{pmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{pmatrix}$$

As a result, we have

$$P_l^t E P_r = 0$$

where $E = S^t R$ is called the **essential matrix**. Note by the construct of S, E has a rank of 2 and two of its singular values equal and the third is 0.

Slide 26

Essential Matrix

The essential matrix E establishes a link between the epipolar constraint and the relative orientation (rotation and translation) of the two coordinate systems. It is determined up to a scale factor by 5 independent parameters (3 for rotation and 2 for translation).

Slide 27

Fundamental Matrix

P_l and P_r represent the 3D coordinates of P in the left and right camera coordinate systems respectively. Let U_l and U_r be the homogeneous pixel coordinates of point P in the left and right images. We know

$$\lambda_l U_l = W_l P_l$$

$$\lambda_r U_r = W_r P_r$$

where W_l and W_r are two 3×3 matrices involving the intrinsic left and right camera parameters.

Substituting the above equations into the essential matrix equation yields

$$U_l^t F U_r = 0$$

Slide 28

where $F = W_l^{-t} E W_r^{-1}$ is called the **fundamental matrix**.

What is the rank of F ?

The rank of F remains 2 due to E.

See table 8.1 (p226) of Hartley's book for a summary of epipolar geometry and the fundamental matrix.

Slide 29

Fundamental Matrix (cont'd)

Unlike the essential matrix E which only encodes the information about the relative orientation between the two cameras, the fundamental matrix F encodes both the relative orientation (extrinsic parameters) and the intrinsic parameters. It is determined up to a scale factor by 7 independent parameters.

Slide 30

Epipolar lines determination

Given a point U_l on the left image, the equation of the epipolar line on the right image is: $F^t U_l$.

Given a point U_r on the right image, the equation of the epipolar line on the left image is: FU_r.

Slide 31

Determination of Epipoles

Let the image coordinates of the left and right epipoles be e_l and e_r respectively. Since every epipolar line passes through the epipoles, for any image point on the left U_l, we have

$$U_l^t F e_r = 0$$

Since U_l in general is non-zero, so we have

$$F e_r = 0$$

which means that the solution to e_r is the eigen vector of F that corresponds to zero eign value (the null eigen vector).

Similarly, we will find that e_l is the null vector of F^t. Refer to page 157 of Trucco's book and page 219 of Forsyth's book on the step by step algorithm to compute epipoles.

Hence, F <===> Epipoles , in other words, given F, we can

Slide 32

compute epioles or given epipoles we can obtain F.

Slide 33

Epipolar Geometry for Lines and Circles

Let X_1 and X_2 be two points on the line segment on the left image and X_1' and X_2' be the corresponding points on the corresponding line segment on the right image. X_1 and X_2 are expressed in homogenous coordinate system.

Then from fundamental equation we have,

$$\begin{aligned} X_1^t F X_1' &= 0 \\ X_2^t F X_2' &= 0 \end{aligned}$$

Hence,

$$(X_1 - X_2)^t F (X_1' - X_2') = 0$$

Let the line on the right image be y=ax+b and the corresponding line on the left image be y'=a'x'+b'.

Slide 34

We have

$$\begin{aligned} X_1 - X_2 &= \begin{pmatrix} x_1 - x_2 \\ y_1 - y_2 \\ 0 \end{pmatrix} \\ X_1' - X_2' &= \begin{pmatrix} x_1' - x_2' \\ y_1' - y_2' \\ 0 \end{pmatrix} \end{aligned}$$

After some substitutions, the epipolar geometry for lines can be expressed as

$$(\begin{pmatrix} 1 \\ a \end{pmatrix}^t F_2 \begin{pmatrix} 1 \\ a' \end{pmatrix} = 0$$

Slide 35

where F_2 is the first 2×2 submatrix of F. This is the relationship constraining two corresponding lines.

Slide 36

Homography Matrix

When the 3D scene are planar or when only rotation is involved between the two cameras, the two corresponding points are uniquely related via the homography. Homography matrix describes completely the relationship between corresponding points. Let H be the homography matrix,

$$\lambda \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix} = H \begin{pmatrix} c_r \\ c_r \\ 1 \end{pmatrix}$$

with F, with one to many mapping (i.e. one point maps to many points on the epipolar line). With H, we have a one-to-one mapping.

Note H and F are not mutually exclusive.

Slide 37

Stereo Calibration

Stereo calibration involves determining the parameters of a system using corresponding 2D points. The parameters of a stereo system include the **intrinsic** parameters (focal length and principle points) and the **extrinsic** parameters including the rigid transformation describing the relative position and orientation of two cameras.

Slide 38

Stereo Calibration (cont'd)

Given n pairs of corresponding points, a system of linear equations can be established involving the elements of F. F can then be solved as a linear least-squares problem.

What is the minimum value of n ?

Slide 39

Solution to F

Given each pair of matched 2D points $U_l = (c_l, r_l, 1)^t$ and $U_r = (c_r, r_r, 1)^t$ and the fundamental equation, $U_l^t F U_r = 0$, offers a linear equation of the 9 unknowns in F.

$$A^{N \times 9} f^{9 \times 1} = 0$$

where $f = (f_{11} \quad f_{12} \quad f_{13} \dots \quad f_{33})$ and

$$A = \begin{pmatrix} c_{l_1} c_{r_1} & c_{l_1} r_{r_1} & c_{l_1} & r_{l_1} c_{r_1} & r_{l_1} r_{r_1} & r_{l_1} & c_{r_1} & r_{r_1} & 1 \\ \vdots & & & & & & & & \\ c_{l_N} c_{r_N} & c_{l_N} r_{r_N} & c_{l_N} & r_{l_N} c_{r_N} & r_{l_N} r_{r_N} & r_{l_N} & c_{r_N} & r_{r_N} & 1 \end{pmatrix}$$

So, given a minimum of 8 pairs (cannot be co-planar), F can be solved up to a scale factor. see Algorithm EIGHT_POINT on page 156 of Trucco's book for details.

Slide 40

If rank(A)=8, the solution to f is the only null vector of A. If rank(A) < 8, then there many solutions equal to linear combinations of all null vectors of A.

Since F is singular, SVD can be used to find another matrix F' that is closest to the computed F but still singular.

$$F = UDV^t$$

setting the smallest value in D to zero yields D'

$$F' = UD'V^t$$

The linear solution is often not accurate and it can be improved with a non-linear optimization method proposed by Luong ea al (1993). It vastly improves the results. Their non-linear method minimizes

Slide 41

$$\sum_{i=1}^{N}[(U_{l_i}^t F U_{r_i})^2 + (U_{r_i}^t F^t U_{l_i})^2]$$

subject to rank(F)=2

To further improve the results, Hartley introduced a normalization procedure (page 156 of Trucco's book) to avoid numerical instability due to ill-conditioned matrix.

Given F, E may computed from $E = W_l^t F W_r$, if the two cameras are calibrated.

Recently Torr and Fitzgibbon (PAMI,BMVC, 2003) show that the 8-point may not be the best method since its results depend on the coordiante system used. They proposed an invariant approach to estiamte the fundemental matrix.

Slide 42

Degeneracies

The solution to F may not be unique (not even up to a scale factor) when degeneracies occur. Degeneracies occur if the rank of matrix A is less than 8. Degeneracy may occur because of a special configuration of the 3D control points such as coplanar points (or points on a special quadric surface) or because of special configuration between two cameras such as pure rotation. The former may be referred to as *critical surface* while the latter may be referred to as *critical motion.* See section 10.9 of Hartley's book for more in-depth discussion on degeneracies.

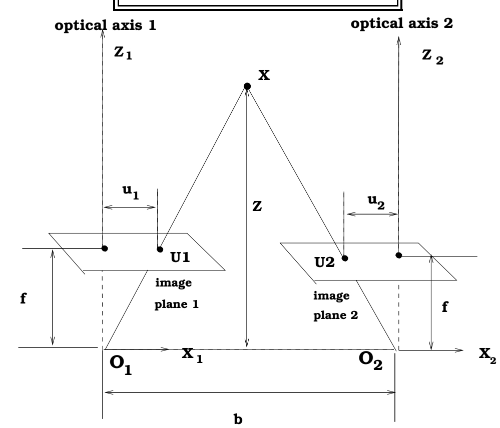
Slide 43

Rectified Stereo Geometry

For computational convenience, the two image planes are often chosen to be coplanar and parallel to their base line (this means equality in focus length). Such an arrangement can be accomplished either physically or through analytic transformation. This arrangement makes the search for correspondence points much easier. The corresponding point for any point on the left image may be found on the same row in the right image.

Slide 44

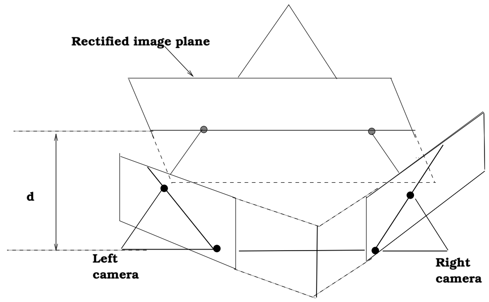
Rectified Geometry



Slide 45

Rectification

Given the extrinsic and intrinsic stereo parameters, we can compute the rotation matrices needed to rotate the cameras such that the conjugate epipolar lines are collinear and are parallel to the base line. The new image coordinates are obtained by projecting the original coordinates on the new image plane.



Also see figure 7.8 of Trucco's book.

Slide 46

Rectification Algorithm

Assuming the relative orientation between the two cameras has been obtained via a stereo calibration procedure, i.e., we have R and T, which specify the relative orientation and translation of the right camera frame with respect to that of the left camera frame.

- Identify a rotation matrix R_{rect} such that when it is applied to the left camera, the image plane of the left camera is parallel to the base line. Note R_{rect} specifies the relative orienation of camera frame before rotation to the camera frame after rotation. For example, let T be the baseline vector relative the original left camera frame and let $(0\ 0\ 1)^t$ be the x-axis after rotation, we then have $R_{rect}\frac{T}{||T||} = (0\ 0\ 1)^t$. This means the first row of R_{rect} is $\frac{T}{||T||}$. Given this, we can then construct the other two rows of R_{rect}. See equation 7.22 for details.

  let $R_l = R_{rect}$. Follow equation 7.22 of Trucco's book to

Slide 47

compute R_{rect}.

- Let $R_r = R_{rect}R$ [a]
- For each image point on the left, form a vector $U_l = (c, r, 1)^t$ and multiply it with $W_l R_l W_l^{-1}$, yielding

$$\begin{pmatrix} sc' \\ sr' \\ s \end{pmatrix} = W_l R_l W_l^{-1} \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} \tag{1}$$

[b] where s is a scale factor.

$$c' = \frac{c's}{s}$$

[a] this makes the rectified left and rigit camera coordinate systems identical in orientation.

[b] obtained from equations $\lambda(c\ r\ 1)^t = W(x_c\ y_c\ z_c)^t$ and $(x_c'\ y_c'\ z_c')^t = R_{rect}(x_c\ y_c\ z_c)^t$

Slide 48

$$r' = \frac{r's}{s}$$

- Apply $W_r R_r W_r^{-1}$ to each point on the right image to compute the new image coordinates for the rectified right image using equation 1.
- Adjust the scale (fs_x and fs_y) appropriately so that the rectified image fits to the original image size. The same scale factor should be applied to the row and column for both left and right images.

Please note W_l and W_r are the intrinsic matrix for the left and right camera and they are assumed to the same. The rectification procedure needs be modified if W_l and W_r are different. One way is to compute $W = 0.5(W_l + W_r)$. Then, replace the first W_l in eq. 1 with W. Similarly, replace the $W_r R_r W_r^{-1}$ for the right camera with $W R_r W_r^{-1}$. Note only the first W for both left and right cameras

Slide 49

are changed to W. The second Ws remain unchanged. Pay attention to the backward mapping discussed in the page 161 of Trucco's book to construct the rectified image.

Slide 50

Examples of Rectification

Slide 51

Figure 2. The first row shows the rectified images of calibration pattern, the second row shows the original images of human face, and the last row shows the rectified images of human face. We can easily verify that the corresponding points are all on the same row.

Figure 2: Rectification results

Slide 52

Techniques for establishing correspondence

- Correlation methods
- Feature-based methods
- Matching constraints
- Hypothesis generation and verification

For all above methods, we assume we are dealing with rectified images and that match takes place along the same row with a disparity (δd) range. The disparity range can be determined by the max and min z distance to the camera, using the fact

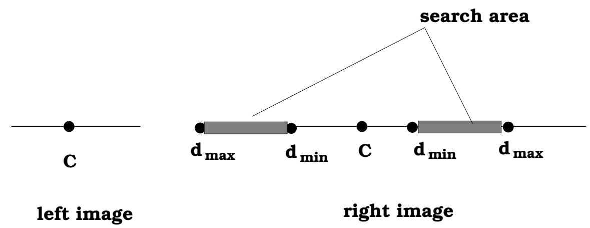
$$\delta d = \frac{fb}{z}$$

Hence, the disparity range for a point (c, r) is

Slide 53

$[c \pm \delta d_{min}, c \pm \delta d_{max}]$, where

$$\delta d_{min} = \frac{fb}{z_{max}}$$
$$\delta d_{max} = \frac{fb}{z_{min}}$$



Slide 54

Correlation methods

The principle of the correlation-based methods is to find two corresponding points based on the intensity distributions of their neighborhoods. The similarity between two neighborhoods is measured by their cross-correlation.

The underlying assumptions include: 1) corresponding image regions look similar, 2) pointed and distant or single light source, 3) corresponding points are visible from both viewpoints.

Slide 55

Correlation methods (cont'd)

Given an image point on the left image, the task is to locate another point in a specified region on the right image that is maximally correlated with the one on the left.

Let W_1 and W_2 be the vectors representing elements in window 1 and window 2, then the cross-correlation $C_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$. The window size varies, typically 5×5 or 7×7.

See the correlation algorithm on pages 146 and 147 of Trucco's book.

Slide 56

Correlation methods (cot'd)

Issues in using correlation based method

- Different methods may be used to compute the correlations. Two different ways (cross correlation and sum of squared differences (ssd)) to compute correlation is introduced in Trucco's book.
- The correlation window size is usually chosen differently, depending on images. Some algorithms determine the window size automatically and adaptive to different parts of the image.
- The search region also varies from algorithms. If the object is far from the camera, small disparity is expected and we can then search in a small neighborhood around the point with the same as the left image point. If we are dealing with a rectified image, we can only search along the same row on the right image. Many techniques assume $\delta d_{min} = 0$ and the window

Slide 57

size is set to $[c - \delta d_{max}, c + \delta d_{max}]$. But in practice, due to errors with image, it is often necessary to search the **neighboring rows**.

- The identified corresponding points should basically satisfy the epipolar constraint.
- To reduce the impact of noise, this technique is often preceded by a smoothing operation.
- Correlation methods need textured images to work well.
- To account for different illuminations, image intensity is usually normalized by subtracting the mean (remove the brightness) and divided by the standard deviation (removes contrast).
- To impose the local smoothness constraint on the disparity map, Markov Random Field is often imposed on the disparity map. Alternatively, we can also perform a post smoothing processing such as median fitering to smooth the disparity map.

Slide 58

Feature-based methods

Feature-based methods restrict the search for correspondences to a sparse set of features. Typical examples of features include edge points, corner points, and lines. Each feature is described by some feature descriptors. For example, for edges, feature descriptors may include edge direction and strength. Features are then matched based on the closeness between their feature descriptors. The assumption is that feature values remain unchanged across images for the corresponding points.

Slide 59

Feature-based correlation method

This technique combines the feature-based and correlation methods by computing the correlation coefficients based on the distributions of feature descriptors (rather than intensity).

Slide 60

Matching Constraints

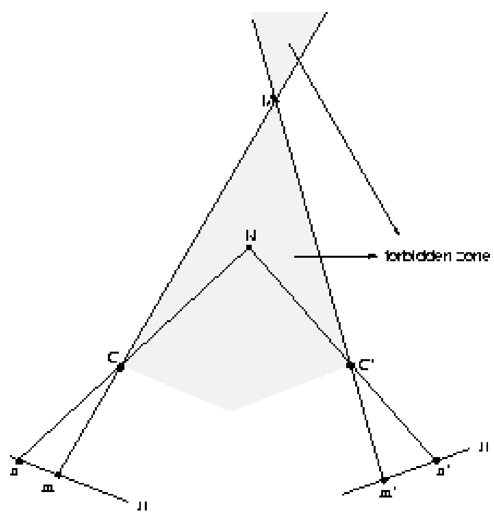
- Smooth constraint (surface locally smooth z=f(d))
  - 0 order disparity constraint: two neighboring image points should have close disparity.
  - first order disparity constraint (disparity gradient): disparity gradient is upper-bounded.

Slide 61

Matching Constraints (cont'd)

- Physical and geometric constraints:
  - Physical constraints (monotonic ordering): the 3D point generated by two corresponding image points should be: a) located on the surface of the same object; b) visible from both views; c) not located inside the object; d) unique (one to one matching); d) ordering is preserved. The ordering constraint fails at regions known as *forbidden zone*.

Slide 62

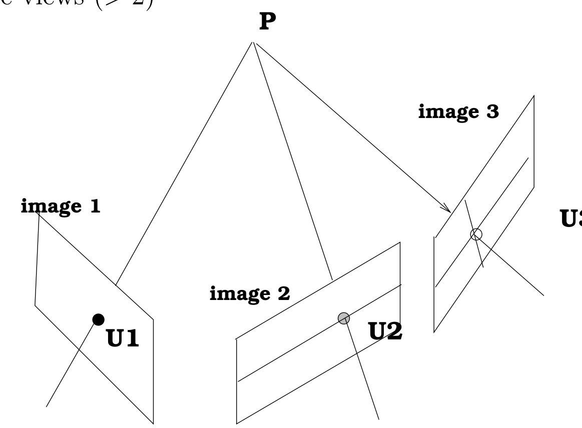


where N is located in the forbidden zone of point M.

Slide 63

Matching Constraints (cont'd)

- Multiple views (> 2)



  - predict point in third image (0 order constraint).

Slide 64

  - predict tangent in third image (first order constraint)
  - predict curvature in third image (second order constraint)

  Check the point in the third image to ensure its geometric properties match the predicted ones.

  Trifocal tensor for three views like fundamental matrix for two views can be used to constrain point in the third view based on the points in two views. See chapters 14 and 15 of Hartley's book.

Slide 65

Matching Constraints (cont'd)

- Locally known surface form (e.g. locally planar or sphere). Given the plane parameters, we can then uniquely determine the corresponding points.

Slide 66

Prediction and verification

- extract primitive features from the images
- assign initial correspondences between the extracted primitives and the features of a possible object model
- infer the 3D pose of the assigned feature
- Backproject the 3D feature onto the image using the computed pose and the known 3D model. Measure the discrepancies between the projected images and the original images.

  This technique requires 3D model of the object.

Slide 67

Other Techniques

- Relaxation techniques
- Dynamic programming

Slide 68

Reconstruction

- Known camera parameters
  - Linear technique
  - Non-linear technique
- Unknown camera parameters
  - Simultaneous estimation of 3D points and camera parameters

Slide 69

Reconstruction: known camera parameters

This means we know R , T, W_l and W_r, where R and T specify the relative orientation between two cameras. Assume the object frame coincide with the left camera frame and let the two image points be (c_l, r_l) and (c_r, r_r), algebraically, we can solve the 3D coordinates (x, y, z) relative to the camera frame as follows:

$$\lambda_l \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix} = W_l \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Slide 70

$$\lambda_r \begin{pmatrix} c_r \\ r_r \\ 1 \end{pmatrix} = W_r[R\ T] \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

given a total of 5 unknowns $(x, y, z, \lambda_l, \lambda_r)$ and 6 linear equations of the unknown, we can setup a least-squares linear system to solve for the five unknowns.

Slide 71

Reconstruction: known camera parameters (cont'd)

Alternatively, we can solve (x, y, z) geometrically. The solution in Trucco's book is a geometric solution. Due to image noise, the two rays (representing the two image points) may not intersect. The goal is to identify the line segment that interests with and orthogonal to the two rays. The center of the line segment is the 3D coordinates we need to compute.

Slide 72

Reconstruction: known camera parameters (cont'd)

The linear solutions, while simple, often do not produce accurate solution. An non-linear solution can be used to improve the linear estimates. The non-linear solution can be formulated as minimizing

$$(P_l(X) - \begin{pmatrix} c_l \\ r_l \end{pmatrix})^t (P_l(X) - \begin{pmatrix} c_l \\ r_l \end{pmatrix}) +$$
$$(P_r(X) - \begin{pmatrix} c_r \\ r_r \end{pmatrix})^t (P_r(X) - \begin{pmatrix} c_r \\ r_r \end{pmatrix})$$

where $P_l(X)$ and $P_r(X)$ are the projected image points on the left and the right image for 3D point X, respectively. (c_l, r_l) and (c_r, r_r) are the corresponding observed image points.

Slide 73

Reconstruction: unknown extrinsic parameters

The 3D geometry can only be reconstructed with respective to either left and right camera coordinate system. Compared with the full reconstruction in the object frame, the reconstruction is up to a scale factor.

If we have 8 or more point correspondences, then we can obtain the fundamental matrix F. Given F and the intrinsic camera parameters, we can then obtain the essential matrix E. From E, we can then recover R and T. We can then have a full reconstruction.

Slide 74

Computing R and T

The above solution assumes the knowledge of R and T. If we do not know R and T, we can solve for them given the fundamental matrix F and the intrinsic camera parameters.

Using the 8-point algorithm, we obtain F and from W_l and W_r, we can then solve for the essential matrix E. Given E, we can solve for (R, T) up to four solutions. See pages 164 and 165 of Trucco's book for details or the paper by Horn [?]. The four (R, T) solutions lead to four reconstructions, only one of them is geometrically consistent.

Slide 75

Computing R and T (cont'd)

Here we introduce Horn's method [?], Let S by the skew matrix computed from T. Given $E = S^t R$, the four solutions for T can be obtained as follows

$$TT^t = \frac{1}{2} Trace(EE^t)I - EE^t$$

where I is the 3×3 identity matrix. The two solutions for the orientation can be found using

$$(T^t T)R = Cofactors(E)^T - SE$$

where $Cofactor(E)_{ij} = (-1)^{i+j} D_{ij}$. D_{ij} is the determinant of the submatrix of E, formed by omitting ith row and jth column.

Slide 76

Reconstruction: unknown camera parameters

Here we are interested in reconstruction with only matched image points and in the absence of any information on both the intrinsic and extrinsic parameters. The reconstruction is unique up to an unknown projective transformation.

If we are given a minimum of three images from the same camera, then we can use self-calibration techniques to recover the intrinsic camera parameters. We can then obtain the relative orientation from the essential matrix, and eventually the complete 3D reconstruction. Alternatively, we can setup a system of non-linear equations and solve for the camera parameters and 3D coordinates simultaneously.

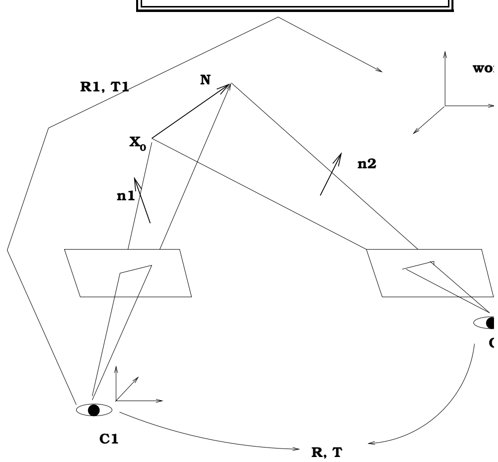
Slide 77

Stereo from Lines and Conics

- Stereo from lines
- Stereo from conic curves

Slide 78

Stereo from Lines



Slide 79

Line Matching

If two line segments match, then any two points from the first line segment must match two points on the second lines. The two points on the second line can be found from the intersections between the second line segment and the epipolar lines .

Slide 80

Stereo from Lines (cont'd)

Assume the object frame coincides with the left camera frame and the relative orientation between two cameras R and T are given, then

$$\begin{aligned} n_1^t N &= 0 \\ n_2^t R^t N &= 0 \\ n_1^t X_0 &= 0 \\ n_2^t (R X_0 + T) &= 0 \end{aligned}$$

the first two equations allow the solve for N. and the last two equations allow to solve for X_0 up to a scale factor.

Slide 81

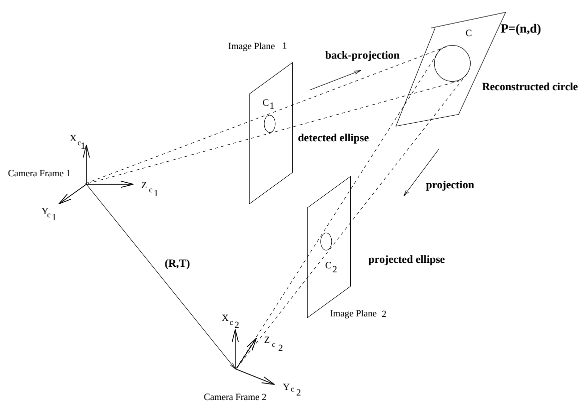
Stereo from Lines (cont'd)

This problem can be made over-determined if 3D lines share the same 3D points X_0. We We need a minimum of 2 3D pencil lines.

Slide 82

Stereo from Conic Curves [?]

- Correspondences between ellipses can be established using epipolar constraint. The corresponding points on the right image are the intersections between the epipolar lines and the ellipse.



Slide 83

- Reconstruction

  In general, to reconstruct 3D conic, we need two image conics. However, when the conic becomes a circle, one image is sufficient to recover the 3D circle if given the radius of the 3D circle.

Slide 84

Limitations with Passive Stereo

A key limitation with the passive stereo approach is the accuracy with the reconstructed 3D points. This lack of accuracy may be due to

- Positional errors (e.g., quantization error) with image coordinates.
- Inaccuracy in point matching (e.g. mismatch),
- Errors with camera parameters (like focal length, baseline distance, etc..)

Slide 85

Bayesian Triangulation

Given observed image points $\hat{U}_1$ and $\hat{U}_2$ and their covariance matrices Σ_{U_1} and Σ_{U_2}, the camera parameters $\hat{\Theta}_1$, and $\hat{\Theta}_2$, and their covariance matrices Σ_{Θ_1}, and Σ_{Θ_2}, estimate the corresponding 3D point X to maximize

$$p(X|\hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$$

Slide 86

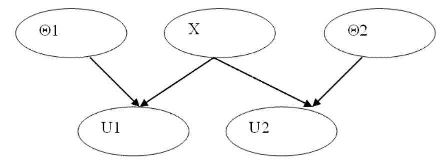
Bayesian Triangulation (cont'd)

Since $p(\hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$ is merely a normalizing constant, maximizing $p(X|\hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$ is the same as maximizing $p(X, \hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$.

Slide 87

Bayesian Triangulation (cont'd)

Assuming the relations between U_i, X, and Θ_i can be modelled by the following Bayesian Network



Slide 88

We then have

$$\begin{aligned}
& p(X|U_1, U_2, \Theta_1, \Theta_2) = \frac{P(X, U_1, U_2, \Theta_1, \Theta_2)}{P(U_1, U_2, \Theta_1, \Theta_2)} \\
& = \alpha P(X, U_1, U_2, \Theta_1, \Theta_2) \\
= \;& P(X)P(\Theta_1)P(\Theta_2)P(U_1|X, \Theta_1)P(U_2|X, \Theta_2) \\
= \;& \alpha P(X)P(\Theta_1)P(U_1|X, \Theta_1)P(\Theta_2)P(U_2|X, \Theta_2) \\
= \;& \alpha P(X) \prod_{i=1}^{2} P(\Theta_i)P(U_i|X, \Theta_i) \qquad (2)
\end{aligned}$$

Slide 89

Bayesian Triangulation (cont'd)

Since $\hat{\Theta}_i \sim N(\Theta_i, \Sigma_{\Theta_i})$,

$$P(\hat{\Theta}_i) = \frac{\exp(-\frac{1}{2}(\hat{\Theta}_i - \Theta_i)^t \Sigma_{\Theta_i}^{-1}(\hat{\Theta}_i - \Theta_i))}{2\pi|\Sigma_{\Theta_i}|^{\frac{1}{2}}}$$

Slide 90

Bayesian Triangulation (cont'd)

Assume given X and $\hat{\Theta}_i$, $\hat{U}_i \sim N(U_i, \Sigma_i)$.

Hence,

$$P(\hat{U}_i|X, \hat{\Theta}_i) = \frac{\exp(-\frac{1}{2}(\hat{U}_i - U_i)^t \Sigma_i^{-1}(\hat{U}_i - U_i))}{2\pi|\Sigma_i|^{\frac{1}{2}}}$$

We need Σ_i and U_i.

Slide 91

Bayesian Triangulation (cont'd)

By perspective projection, we have $\hat{U}_i = M(\hat{\Theta}_i, X)$, where M is the projection matrix. Assume the noise with camera parameters are small, by a first order Taylor series approximation of $M(\hat{\Theta}_i, X)$, we have

$$\begin{aligned} \hat{U}_i &= M(\hat{\Theta}_i, X) \\ &= M(\Theta, X) + \frac{\partial M}{\partial \Theta_i}\Delta\Theta_i \end{aligned}$$

Slide 92

Bayesian Triangulation (cont'd)

Since $M(\Theta_i, X) \sim N(U_i, \Sigma_{U_i})$ and $\Delta\Theta_i \sim N(0, \Sigma_{\Theta_i})$, we have

$$\Sigma_i = \Sigma_{\Theta_i} + (\frac{\partial M}{\partial \Theta_i})\Sigma_{\Theta_i}(\frac{\partial M}{\partial \Theta_i})^t$$

Since U_i represents the ideal yet unobserved image projection given Θ_i and X. Assume small perturbations with image and camera, U_i may be approximated by $M(\hat{\Theta}_i, X)$.

Slide 93

Bayesian Triangulation (cont'd)

To solve for $p(X, \hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$, we still need to compute $p(X)$. Assume X is uniformally distributed, then maximizing $p(X, \hat{U}_1, \hat{U}_2, \hat{\Theta}_1, \hat{\Theta}_2)$ is equivalent to maximizing $E = \prod_{i=1}^{2} p(\hat{U}_i|\hat{\Theta}_i, X)p(\Theta_i)$.

Slide 94

Bayesian Triangulation (cont'd)

$$\begin{aligned}\log E &= \sum_{i=1}^{2}\{\log p(\hat{U}_i|\hat{\Theta}_i, X) + \log p(\hat{\Theta}_i)\} \\ &= \sum_{i=1}^{2}\{-\frac{1}{2}(\hat{U}_i - U_i)^t\Sigma_i^{-1}(\hat{U}_i - U_i) - \\ &\quad \log(2\pi) - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(\hat{\Theta}_i - \Theta_i)^t\Sigma_{\Theta_i}^{-1}(\hat{\Theta}_i - \Theta_i) - \\ &\quad \log(2\pi) - \frac{1}{2}\log|\Sigma_{\Theta_i}|\}\end{aligned}$$

Slide 95

Bayesian Triangulation (cont'd)

Minimizing $\log E$ is equivalent to minimizing $-\log E$. Removing the constant terms in $-\log E$, minimizing $-\log E$ is equivalent to minimizing

$$\begin{aligned}&\sum_{i=1}^{2}\{(\hat{U}_i - U_i)^t\Sigma_i^{-1}(\hat{U}_i - U_i) + \log|\Sigma_i| \\ &+(\hat{\Theta}_i - \Theta_i)^t\Sigma_{\Theta_i}^{-1}(\hat{\Theta}_i - \Theta_i) + \log|\Sigma_{\Theta_i}|\}\end{aligned}$$

Slide 96

Bayesian Triangulation (cont'd)

Minimize

$$\epsilon = \sum_{i=1}^{2}(\hat{U}_i - M(X,\Theta_i))^T\Sigma_i^{-1}(\hat{U}_i - M(X,\Theta_i)) + \log|\Sigma_i|$$

where

$$\Sigma_i = \Sigma_{U_i} + [\frac{\partial M}{\partial \Theta_i}(X,\Theta_i)]\Sigma_{\Theta_i}[\frac{\partial M}{\partial \Theta_i}(X,\Theta_i)]^t$$

Slide 97

Monocular Approach

Problem Statement : *Given a single 2D image of an object, reconstruct the shape of the visible surface of the object.*

A monocular image alone does not contain sufficient information to uniquely retrieve range information. Range, however, can be recovered from monocular images in conjunction with certain visual cues or prior knowledge of certain geometric properties of the object.

Slide 98

Visual Cues

What are the visual cues used by human to infer depth or shape?

- Geometric properties
- Shade
- Distortion
- Vanishing points
- Blurriness

Slide 99

Geometric Properties

- Euclidean distance relationships.
- Angular relationships.

Slide 100

Monocular 3D Techniques

- Shape from shading.
- Shape from texture.
- Shape from geometry.
- Photometric stereo.
- Shape from focus.

Slide 101

Shape from Shading

Shape-from-shading technique infers the depth of an image pixel based on its shade (or intensity). The underlying theory of the technique is that the intensity of a pixel is determined in part by, among many other factors, the angle between surface normal (slope) and the illumination direction. See fig. 9.1.

Slide 102

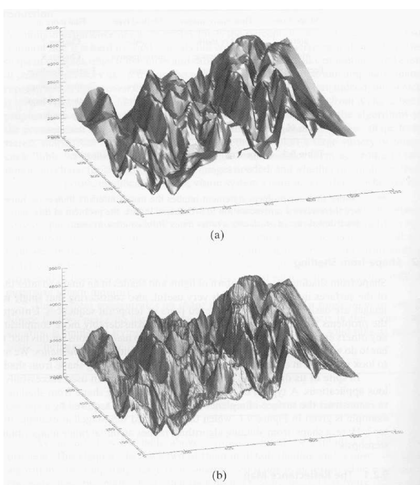
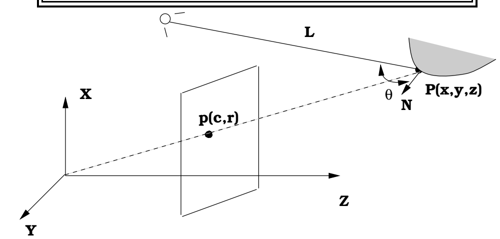


Figure 9.1 (a): 3-D rendering of the surface reconstructed by stereopsis from the stereo pair on page 142. (b): the refined surface obtained by the shape-from-shading algorithm; notice the increased level of detail. Courtesy of Alois Goller, Institute for Computer Graphics and Vision, Technical University of Graz.

The top is constructed using stereo, while the bottom is from SFS.

Slide 103

Shape from Shading Geometry



Slide 104

Surface radiance and Lambertian Model

Let $S(x, y, z)$ be the surface radiance at point (x, y, z), $\mathbf{n}$(x,y,z) be the surface normal at this point, and $\mathbf{L}$ be a vector representing the incident light,

$$S(x, y, z) = \rho \mathbf{L}^t \mathbf{n} = \rho ||L|| \cos\theta \tag{3}$$

where ρ is the surface *albedo* and θ is the angle between L and n. A surface reflectance model that follows the radiance equation 3 is called *Lambertian Reflectance Model*.

The above equation assumes distant and single illumination source.

Slide 105

The Image irradiance Equation

Let $E(c,r)$ and $I(c,r)$ be the image irradiance and intensity at pixel (c,r) respectively. Then surface radiance $S(x,y,z)$ relates to E(c,r) via the fundamental radiometric equation, i.e.,

$$E(c,r) = S(x,y,z)\frac{\pi}{4}(\frac{d}{f})^2\cos^4\alpha \tag{4}$$

where d, f, and α represent the lens diameter, focal length, and the angle formed by the principal ray through the point at (x,y,z) and the optical axis. In the case of small angular aperture, this effect can be ignored; therefore the image irradiance can be regarded as proportional to the scene radiance, i.e.,

$$E(c,r) = \alpha S(x,y) = \alpha\rho L^t n$$

Since $E(c,r)$ is proportional to $I(c,r)$, so for Lambertian model

Slide 106

and a single distant light source, we have

$$I(c,r) = \alpha\beta\rho L^t n$$

where β is the coefficient that relates image irradiance to intensity. Assuming the optical system has been calibrated, the constant terms α and β can be dropped from the above equation, leading to

$$I(c,r) = \rho L^t n \tag{5}$$

Equation 5 is the fundamental equation for shape from shading, assuming Lambertian illumination model, a single distant, and point light source.

Slide 107

Surface Normal Computation

Assume the surface depth z may be thought as a function of (x,y), i.e., z=f(x,y). A 3D point P can therefore be represented as $P = (x,y,f(x,y))$. Hence the normal vector at (x,y,z) is

$$n(x,y,z) = \frac{\frac{\partial P}{\partial x} \times \frac{\partial P}{\partial y}}{||\frac{\partial P}{\partial x} \times \frac{\partial P}{\partial y}||} = \frac{[-p(x,y), -q(x,y), 1]^t}{\sqrt{1+p^2(x,y)+q^2(x,y)}}$$

where $p = \frac{\partial f}{\partial x}$ and $q = \frac{\partial f}{\partial y}$ Note (x,y,z) is the coordinates of 3D point relative to the camera frame. As a result, p and q are specified with respect to the camera coordinate frame.

Slide 108

Therefore for surface that follows the Lambertian Model, ideally we have

$$I(x,y) = \rho L^t \frac{[-p(x,y), -q(x,y), 1]^t}{\sqrt{1+p^2(x,y)+q^2(x,y)}} \tag{6}$$

Slide 109

Weak Perspective Projection Assumption

Since the object is assumed to be far away from the viewer (camera), we can assume weak perspective projection. As a result, we have

$$\begin{aligned} c &= \frac{fs_x}{\bar{z}}x + c_0 \\ r &= \frac{fs_y}{\bar{z}}y + r_0 \end{aligned}$$

$\bar{z}$ is the average object z distance to the camera. Given $P(x, y, f(x, y))$ and P is also a function of c and r, hence

$$\begin{aligned} \frac{\partial P}{\partial c} &= (\frac{\partial x}{\partial c}, 0, \frac{\partial f}{\partial x}\frac{\partial x}{\partial c}) = (\frac{\bar{z}}{fs_x}, 0, p(x,y)\frac{\bar{z}}{fs_x}) \\ \frac{\partial P}{\partial r} &= (0, \frac{\partial y}{\partial r}, \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}) = (0, \frac{\bar{z}}{fs_y}, q(x,y)\frac{\bar{z}}{fs_y}) \end{aligned}$$

Slide 110

Hence, we can easily prove that

$$\frac{\frac{\partial P}{\partial c} \times \frac{\partial P}{\partial r}}{||\frac{\partial P}{\partial c} \times \frac{\partial P}{\partial r}||} = \frac{\frac{\partial P}{\partial x} \times \frac{\partial P}{\partial y}}{||\frac{\partial P}{\partial x} \times \frac{\partial P}{\partial y}||}$$

Hence,

$$\begin{aligned} n(x,y,z) &= \frac{[-p(x,y), -q(x,y), 1]^t}{\sqrt{1 + p^2(x,y) + q^2(x,y)}} \\ &= \frac{[-p(c,r), -q(c,r), 1]^t}{\sqrt{1 + p^2(c,r) + q^2(c,r)}} \qquad (7) \end{aligned}$$

Slide 111

Problem Statement

Given ρ and L, the problem is to reconstruct surface slopes p and q and then surface height z for which

$$I(c,r) = \rho R(L, p(c,r), q(c,r))$$

where $R(L, p(c,r), q(c,r))$ is the reflectance map of a surface. For Lambertian Model, we have

$$R = L^t \frac{[-p(c,r), -q(c,r), 1]^t}{\sqrt{1 + p^2(c,r) + q^2(c,r)}}$$

In general, R is complicate or is estimated numerically via experiments.

Slide 112

A Calculus of Variational Approach

Find p and q for each point by minimizing

$$\epsilon = \int \{[(I(c,r) - \rho R(p,q)]^2 + \lambda(p_c^2(c,r) + p_r^2(c,r) + q_c^2(c,r) + q_r^2(c,r))\} dcd$$

where $R(p,q)$ is the reflectance map and λ is the Lagrange multiplier for imposing the *smoothness constraint*. $p_c = \frac{\partial p}{\partial c}$, $p_r = \frac{\partial p}{\partial r}$, $q_c = \frac{\partial q}{\partial c}$, $q_r = \frac{\partial q}{\partial r}$. λ is always positive and determines the relative importance between two terms (the fundamental equation and smoothness).

Slide 113

The Euler-Lagrange Equations

Taking partial derivatives with respect to p and q and set them to zeros yield

$$\frac{\partial \epsilon}{\partial p} - \frac{\partial}{\partial c}\frac{\partial \epsilon}{\partial p_c} - \frac{\partial}{\partial r}\frac{\partial \epsilon}{\partial p_r} = 0$$
$$\frac{\partial \epsilon}{\partial q} - \frac{\partial}{\partial c}\frac{\partial \epsilon}{\partial q_c} - \frac{\partial}{\partial r}\frac{\partial \epsilon}{\partial q_r} = 0$$

Slide 114

Discrete Case

Find p and q for each point by minimizing

$$\epsilon = \sum_c \sum_r [(I(c,r) - \rho R(p,q)]^2 + \lambda(p_c^2 + p_r^2 + q_c^2 + q_r^2)$$

Slide 115

Discrete Case (cont'd)

In digital domain, p_c, p_r, q_c, q_r may be numerically approximated as

$$\begin{aligned} p_c &= p(c+1,r) - p(c,r) \\ p_r &= p(c,r+1) - p(c,r) \\ q_c &= q(c+1,r) - q(c,r) \\ q_r &= q(c,r+1) - q(c,r) \end{aligned}$$

Slide 116

Discrete Case (cont'd)

As a result, find p and q for each point by minimizing

$$\begin{aligned} \epsilon = & \sum_c \sum_r [(I(c,r) - R(p,q)]^2 + \lambda[(p(c+1,r) - p(c,r))^2 \\ + & (p(c,r+1) - p(c,r))^2 + (q(c+1,r) - q(c,r))^2 + \\ & (q(c,r+1) - q(c,r))^2] \end{aligned}$$

Slide 117

Discrete Case (cont'd)

Taking partial derivatives with respect to p and q and set them to zeros yield

$$
\begin{aligned}
p_{c,r} &= \bar{p}_{c,r} + \frac{1}{4\lambda}(I-R)\frac{\partial R}{\partial p} \\
q_{c,r} &= \bar{q}_{c,r} + \frac{1}{4\lambda}(I-R)\frac{\partial R}{\partial q}
\end{aligned}
$$

$\bar{p}_{c,r}$ and $\bar{q}_{c,r}$ are the average of p and q over the four nearest neighbors.

Slide 118

$$
\begin{aligned}
p^{k+1}_{c,r} &= \bar{p}^{k}_{c,r} + \frac{1}{4\lambda}(I-R)\frac{\partial R}{\partial p}|^{k} \\
q^{k+1}_{c,r} &= \bar{q}^{k}_{c,r} + \frac{1}{4\lambda}(I-R)\frac{\partial R}{\partial q}|^{k}
\end{aligned}
$$

Slide 119

Convergence Properties

- Convergence-depending on the initial p and q.
- The optimal λ-normally around 1000.
- Stopping condition-residual error may not be always accurate

Slide 120

Other issues

- Estimate illumination direction.
- Estimate the reflectance map R(p,q).
  - A more general reflectance model is the **3-lobe model** or the **BRDF** (bidirection reflectance distribution function) model [**?**].
- Derive depth z from p and q.
  - Given p and q, z may be computed by minimizing $\epsilon = \sum_c \sum_r (z_c - p)^2 + (z_r - q)^2$, where the partial directives p_c and p_r may be approximated numerically.
- Estimate surface albedo
- Handle specular surface, see Phone's model in [**?**]
- Deal with inter-reflections.

Slide 121

Limitations with Shape from Shading Theory

- Assume single point illumination source.
- Assume all visible surface points receive direct illumination
- Assume Lambertian surface.
- Assume weak perspective projection (new paper in ICCV03 to extend it to perspective projection)
- Assume Z=f(x,y)

Slide 122

Photometric Stereo

At each point (x,y), we have

$$I_k = \rho n^t L_k,$$

where $k = 1, 2, \ldots N$ and $N > 3$, representing N light sources.

Slide 123

Photometric Stereo (cont'd)

Given N illumination directions, we can setup a system of linear equations involving

$$A = \rho \begin{pmatrix} L_{1_x} & L_{1_y} & L_{1_z} \\ L_{2_x} & L_{2_y} & L_{2_z} \\ \vdots & & \\ L_{N_x} & L_{N_y} & L_{N_z} \end{pmatrix} \qquad b = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{pmatrix}$$

n can be solved for by minimizing $||An - b||^2$, which leads to $n = \frac{(A^t A)^{-1} A^t b}{\rho}$, where $\rho = ||(A^t A)^{-1} A^t b||$.

Slide 124

Photometric Stereo (cont'd)

The assumptions of the above solution is that surface points can not be in shadows for any of the light sources and illumination directions are known.

Slide 125

Photometric calibration

For SFS and photometric stereo, the image intensity I needs to be photometrically calibrated to remove the camera gain and offset.

Let I be the observed intensity and I' be the intensity without camera gain and offset. We have

$$I = a\rho L^t n + b$$

where a and b are the camera gain and offset.

Camera photometric calibration involves computing I' from I, i.e, $I' = \frac{I-b}{a}$. This may be accomplished via a geometric setup [?] page 16 of Trucco's book.

Slide 126

Shape from Texture

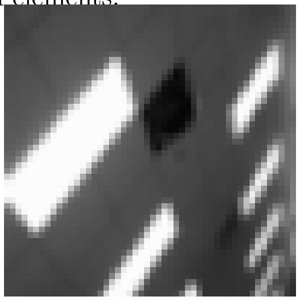
Given a single image of a surface consisting of texture patterns, estimate the shape of the surface from the distortions of the observed image texture patterns.

Slide 127

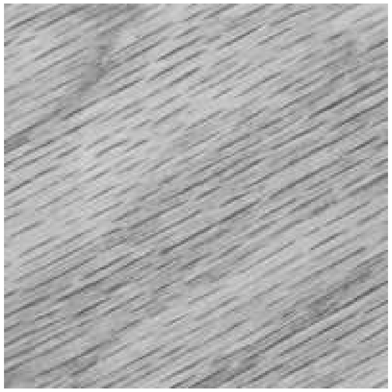
What is Texture

A texture surface is created by regular repetitions of a basic texture element or regular repetitions of certain statistical properties of surface color. The former may be referred to as the *deterministic* texture while the latter may be referred to as the *statistic* texture.

For deterministic textural elements, they can be characterized by the shape parameters such as parameters for lines or ellipses if they are the basic textural elements.



Slide 128



For statistic texture, they are often characterized by their statistical properties such as entropy, randomness, and spatial frequency over a region.

Slide 129

Slide 130

Sources of Texture Distortions

Texture distortions refer to texture properties change due to imaging process.

A uniformly-textured surfaces undergo two types of projective distortions

- Perspective distortion
- Foreshortening

Slide 131

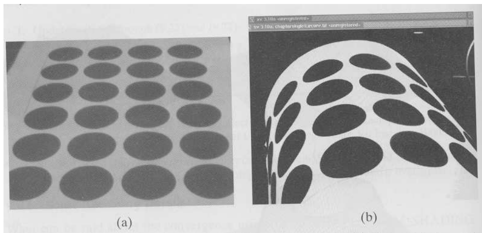
Sources of Texture Distortions (cont'd)

Perspective distortion causes a uniform compression in the area of a texture element as the distance between surface and the camera increases. *Foreshortening* causes an anisotropic compression of a texture element. The texture distortions provide information about the relative distances and orientations of the textured surface in an image.

Slide 132

Perspective distortion can be quantified by the area variations, while foreshortening distortion can be characterized by the ratio of the ellipse's semi-axes.

Slide 133



Slide 134

Measures of Texture Distortions [?]

- Distortion measures (e.g., area, shape ratio)
- Distortion measures gradients (change of distortion measures)

While the first measure may be used to quantify distortion for a single texture element, the latter may be applicable for a region containing several textural elements.

Slide 135

Steps for Shape Estimation from Texture

- identify a region in the image containing a texture element.
- determine the textural properties to use
- compute texture distortions w.r.t the selected texture properties such as including area, aspect ratio, and density gradients.
- estimate surface from the computed texture distortions.

Slide 136

Shape from Geometry [?, ?]

Also referred to as *shape from inverse perspective projection*, *shape from geometry* reconstructs a 3D geometric entity from a single image in conjunction with geometric constraints on the 3D geometric entity.

The geometric constraints may come from the world model of the object being viewed in the perspective projection.

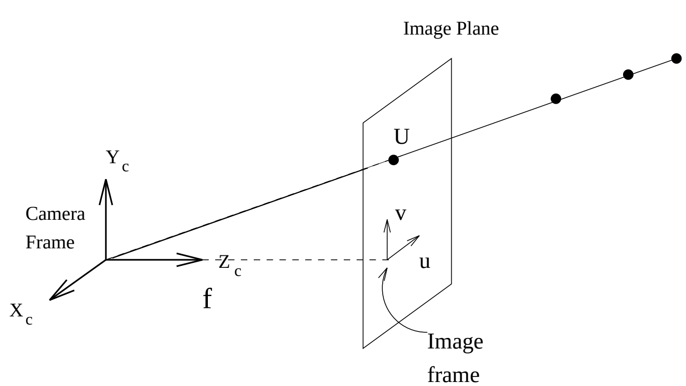
Slide 137

3D Geometric Entities

The 3D geometric entities may include points, lines, planes, and 3D curves like circles or ellipses

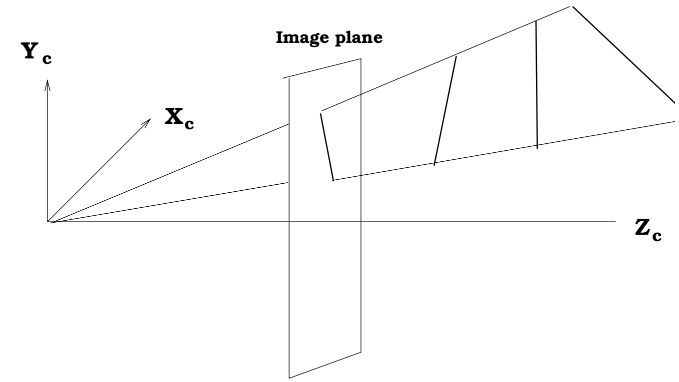
Slide 138

An Ill-posed Problem: Point



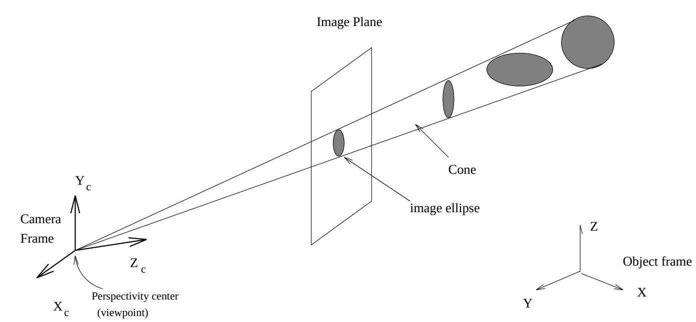
Slide 139

An Ill-posed Problem: Line



Slide 140

An Ill-posed Problem: Curves



Slide 141

Geometric Constraints

The geometric constraints employed include

- Euclidean distance constraints.
- Angular constraints

Slide 142

Euclidean Distance Constraints

The distance relationships may include distance among points, lines, and planes

- distance between points
  - two unknown 3D points
  - a known 3D point and an unknown 3D point
- distance between a point and a line
  - an unknown 3D point and a known line
  - an known 3D point and an unknown line
- distance between a point and a plane
  - an unknown 3D point and an unknown plane
  - a known 3D point and an unknown plane

Slide 143

Euclidean Distance Constraints (cont'd)

Distance relationship may also generalize to collinear and coplanar. They include

- three unknown 3D points are collinear
- an unknown 3D point is on a known line
- an unknown 3D point is on an unknown plane
- a known 3D point is on an unknown plane

Slide 144

Euclidean Angular Constraints

In addition to distance relationships, there also exist angular relationships. They may include angle formed by

- two unknown planes.
- three unknown non-collinear points.
- a known line and an unknown plane.

Slide 145

Reconstruct 3D Points

A 3D point and its image offers two equations. The 3D point can be reconstructed if a constraint is available about the 3D point.

Slide 146

Reconstruct 3D Points

If an unknown 3D point X is located on a known plane, then X can be solved for.

Slide 147

Reconstruct 3D Points

If the distance between X and a known 3D point Y is known, then X can be solved for.

Slide 148

Reconstruct 3D Points (cont'd)

Given the images of three 3D points and the distances between the 3D points. Then the 3D points can be reconstructed. This is the famous 3-point problem. The solution is not unique however. If we know one of the 3 points, then the other two points can be determined uniquely.

Slide 149

Reconstruct 3D Points (cont'd)

Similarly, for three non-collinear points, they may be reconstructed if the three angles they form are known. Likewise, the solution is not unique.

Slide 150

Reconstruct 3D Lines

A 3D line may be represented by a point on the line P_0 and its direction cosine N. Hence,

$$P_0 + \lambda N$$

is the equation of a line, where λ is a scaler.

A line therefore has a total of 5 parameters (3 for P_0 and 2 for N). The number of parameters may reduce to 4 if P_0 is chosen such that $P_0^t N = 0$.

Slide 151

The relationships between a 3D line and 2D line

$$\begin{aligned} n^t N &= 0 \\ n^t P_0 &= 0 \end{aligned}$$

where n is the normal of the plane formed by the 3D line and the perspectivity center.

Each line offers two equations. But to solve for the four line parameters, additional constraints are needed.

Slide 152

Reconstruct 3D Lines: M Parallel Lines

Given $M \geq 2$ parallel lines, we can solve for the direction cosine N using the first constraint equation. When M=2, N can be derived as a special case.

For parallel lines, we can obtain their vanishing point. The coodinates of the vanishing points can be used to recover the oritentation of the lines. In fact, the line that connecting the lens center and the vanishing point is parallel to the 3D line.

Slide 153

Reconstruct 3D Lines: N Pencil Lines

Given $M \geq 3$ lines intersecting at point P_0, we can solve for P_0 (up to a scale factor) using the second constraint equation. The potential application of this is we can compute the 3D coordinates of a vertex.

Slide 154

Reconstruct 3D Lines: N Pencil Lines

Given $M \geq 3$ lines intersecting at a point with known angles between each pair of 3D lines, we can then solve for the orientation of each 3D line. The solution, however, is nonlinear and requires initial estimates.

Slide 155

Collinear Points with Known Interpoint Distances

Suppose we observe the perspective projection of a 3D line whose position and orientation are unknown. On this line there are $N(> 2)$ distinguished points with known interpoint distances. This constitutes enough information to determine all the parameters of the line as well as the 3D positions of the points.

Slide 156

Collinear Points with Known Interpoint Distances (cont'd)

Let $P_0 = (P_{0_x}, P_{0_y}, P_{0_z})$ and $N = (N_x, N_y, N_z)$ be the first point on the 3D line and the orientation of the line. Let P_1 be the second point on the line and (u_1, v_1) be the image projection of P_1, then we have

$$P_1 = P_0 + d_1 N$$

where d_1 is the distance between P_0 and P_1.

From the perspective projection equation, we have

$$\begin{aligned} u_1 &= f\frac{P_{0_x} + d_1 N_x}{P_{0_z} + d_1 N_z} \\ v_1 &= f\frac{P_{0_y} + d_1 N_y}{P_{0_z} + d_1 N_z} \end{aligned}$$

Slide 157

As a result, we can setup a system of linear equations involving the unknown parameters P_0 and N, which can be solved by a constrained linear least-squares method.

Slide 158

Lines with known length and orientation

Given the length of a line segment as d and its orientation N, the coordinates of two endpoints can be reconstructed.

Let the two endpoints be $X_1 = (x_1, y_1, z_1)$ and $X_2 = (x_2, y_2, z_2)$ and their corresponding image points be $U_1 = (u_1, v_1)$ and $U_2 = (u_2, v_2)$. Since

$$X_2 = X_1 + \lambda N$$

$$\begin{aligned} u_1 &= \frac{f x_1}{z_1} \\ v_1 &= \frac{f y_1}{z_1} \\ u_2 &= \frac{f(x_1 + dN_x)}{z_1 + dN_z} \\ v_2 &= \frac{f(y_1 + dN_y)}{z_1 + dN_z} \end{aligned}$$

Slide 159

Hence the above four equations yield an over-determined system of linear equations involving 3 unknowns x_1, y_1, and z_1. Solving X_1, we can then obtain X_2.

Slide 160

Lines in a known plane

Let the plane parameters be d and V and the line parameters be P_0 and N, we then have

$$\begin{aligned} P_0^t V &= d \\ N^t V &= 0 \end{aligned}$$

We can then solve for P_0 and N using the above two equations along with the two constraint equations.

Slide 161

Parallel Lines with known vanishing points

Given the vanishing point of a set of parallel lines, the orientation of the lines can be determined as follows

Let the vanishing point be (u, v) and $N = (N_x, N_y, N_z)$, then we have

$$
\begin{aligned}
u &= f\frac{N_x}{N_z} \\
v &= f\frac{N_y}{N_z}
\end{aligned}
$$

Using the above equation and the fact that $N_x^2 + N_y^2 + N_z^2 = 1$, we can solve for N.

Slide 162

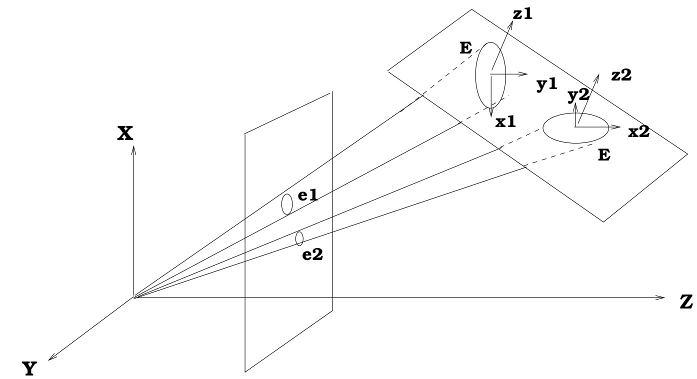
3D Circle Reconstruction

Given the radius of the 3D circle, we can then reconstruct the 3D circle from its image.

Slide 163

3D Conic Reconstruction

If we are given the images of two 3D coplanar conic curves of identical size but different orientations, we can then reconstruct the two 3D conic curves.



Slide 164

Let $X = (x, y, z)$ be a point in the camera coordinate frame X-Y-Z, $X_1 = (x_1, y_1, 1)$, and $X_2 = (x_2, y_2, 1)$ be the homogeneous coordinates of the corresponding points in the frames X1-Y1-Z1 and X2-Y2-Z2 respectively.

X, X_1, and X_2 are related as follows

$$
\begin{aligned}
X &= R_1 X_1 \\
X &= R_2 X_2 \\
X_1 &= R X_2
\end{aligned}
$$

where $R_1 = (r_1\ r_2\ t_1)$, $R_2 = (r_1'\ r_2'\ t_2)$, R represents a 2D transformation that maps points in frame X2-Y2-Z2 to frame X1-Y1-Z1.

The two 3D conics can be represented in matrix format as follows

$$
\begin{aligned}
X_1^t E X_1 &= 0 \\
X_2^t E X_2 &= 0
\end{aligned}
$$

References

Slide 165

Similarly, the two image conics can be written as

$$X^t e_1 X = 0$$
$$X^t e_2 X = 0$$

Using the relationships between X, X_1, and X_2, we have

$$R_1^t e_1 R_1 = kE$$
$$R_2^t e_2 R_2 = kE$$

where k is a scalar.

Using the relationships between X_1 and X_2, we have $R_2 = R_1 R$. Substituting it into the above equation leads to

$$R_1^t e_1 R_1 = kE$$
$$R^t R_1^t e_2 R_1 R = kE$$

Slide 166

The above two equations offer (due to symmetry) a total of 12 equations with 12 unknowns (6 from R_1, 3 from R, 1 from k, and 2 from E). We shall be able to solve for all unknowns.