

Perceiving 3D from 2D Images

How can we derive 3D information from one or more 2D images?

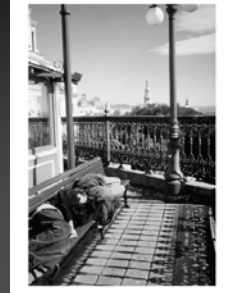
There have been 2 approaches:

1. **intrinsic images**: a 2D representation that stores some 3D properties of the scene
2. **3D shape from X**: methods of inferring 3D depth information from various sources

1

What can you determine about

1. the sizes of objects
2. the distances of objects from the camera?



What knowledge do you use to analyze this image?

2

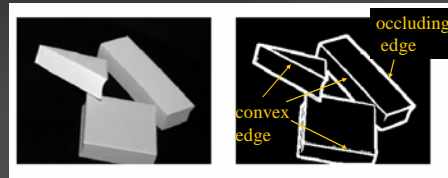
What objects are shown in this image?
How can you estimate distance from the camera?
What feature changes with distance?



3

Intrinsic Images: 2.5 D

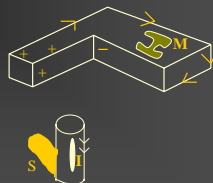
The idea of intrinsic images is to label features of a 2D image with information that tells us something about the 3D structure of the scene.



4

Contour Labels for Intrinsic Images

- convex crease (+)
- concave crease (-)
- blade (>)
- limb (>>)
- shadow (S)
- illumination boundary (I)
- reflectance boundary (M)



5

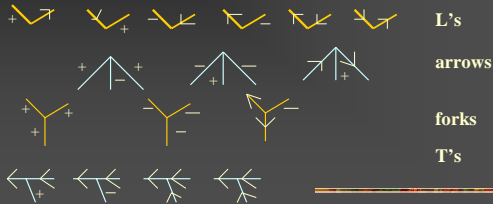
Labeling Simple Line Drawings

- Huffman and Clowes showed that **blocks world drawings** could be labeled (with +, -, >) based on real world constraints.
- Labeling a simple blocks world image is a **consistent labeling problem!**
- Waltz extended the work to **cracks and shadows** and developed one of the first discrete relaxation algorithms, known as **Waltz filtering**.

6

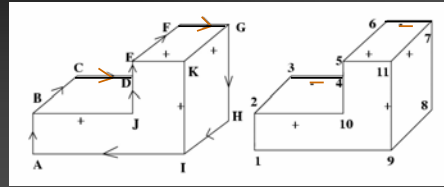
Simple Blocks World Constraints for Objects with Trihedral Junctions

There are only 16 topologically possible junctions for this class of images. Huffman/Clowes categorized these.



7

2 Interpretations



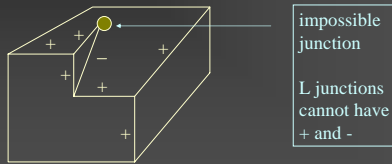
floating

glued to the wall

8

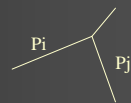
Line Drawing Labeling

Given a line drawing extracted from an image, find the correct labeling(s).



9

Automatic Labeling



Finding a legal labeling can be done by:

1. **tree search** with backtracking when a node is inconsistent
2. **Waltz filtering** or discrete relaxation

Initialize the label set for each line segment to $\{+, >, <\}$

At each iteration, remove inconsistent labels as follows

If L is a label for edge P_i and there is another edge P_j connected to P_i that has no label consistent with L , then remove L from the label set of P_i .

10

Problems with this Approach

- Research on how to do these labelings was confined to **perfect blocks world images**
- There was **no way to extend** it to real images with missing segments, broken segments, nonconnected junctions, etc.
- It led some groups down the **wrong path for a while**.

11

3D Shape from X

- shading
 - silhouette
 - texture
- } mainly research
- stereo
 - light striping
 - motion
- } used in practice

12

Perspective Imaging Model: 1D

This is the axis of the real image plane.

O is the center of projection.

This is the axis of the front image plane, which we use.

$$\frac{x_i}{f} = \frac{x_c}{z_c}$$

13

Perspective in 2D (Simplified)

3D object point $P=(x_c, y_c, z_w)$

ray

optical axis

camera X_c

$P'=(x_i, y_i, f)$

F

f

x_i

y_i

z_c

x_c

$z_w = z_c$

y_c

Here camera coordinates equal world coordinates.

$$\frac{x_i}{f} = \frac{x_c}{z_c}$$

$$\frac{y_i}{f} = \frac{y_c}{z_c}$$

$$x_i = (f/z_c)x_c$$

$$y_i = (f/z_c)y_c$$

14

3D from Stereo

3D point

left image

right image

disparity: the difference in image location of the same 3D point when projected under perspective to two different cameras.

$$d = x_{left} - x_{right}$$

15

Depth Perception from Stereo Simple Model: Parallel Optic Axes

image plane

camera L

baseline

camera R

f

x_l

f

x_r

$P=(x, z)$

$x-b$

X

Z

$$\frac{z}{f} = \frac{x}{x_l}$$

$$\frac{z}{f} = \frac{x-b}{x_r}$$

$$\frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}$$

y-axis is perpendicular to the page.

16

Resultant Depth Calculation

For stereo cameras with parallel optical axes, focal length f , baseline b , corresponding image points (x_l, y_l) and (x_r, y_r) with disparity d :

$$z = f*b / (x_l - x_r) = f*b/d$$

$$x = x_l*z/f \quad \text{or} \quad b + x_r*z/f$$

$$y = y_l*z/f \quad \text{or} \quad y_r*z/f$$

This method of determining depth from disparity is called **triangulation**.

17

Finding Correspondences

- If the correspondence is correct, triangulation works **VERY** well.
- But correspondence finding is **not perfectly solved** for the general stereo problem.
- For some **very specific applications**, it can be solved for those specific kind of images, e.g. windshield of a car.

18

2 Main Matching Methods

1. Cross correlation using small windows.



dense

2. Symbolic feature matching, usually using segments/corners.

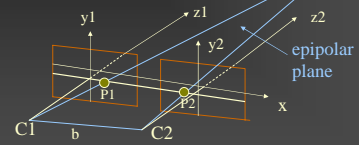


sparse

19

Epipolar Geometry Constraint: 1. Normal Pair of Images

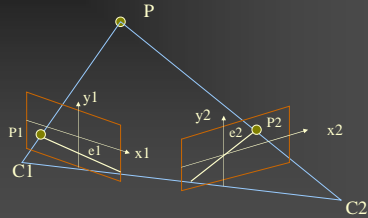
The epipolar plane cuts through the image plane(s) forming 2 epipolar lines.



The match for P1 (or P2) in the other image, must lie on the same epipolar line.

20

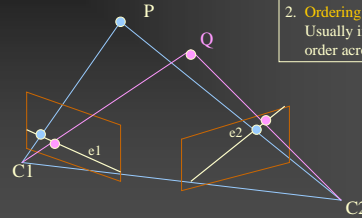
Epipolar Geometry: General Case



21

Constraints

1. **Epipolar Constraint:** Matching points lie on corresponding epipolar lines.
2. **Ordering Constraint:** Usually in the same order across the lines.

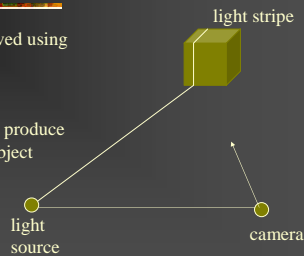


22

Structured Light

3D data can also be derived using

- a single camera
- a light source that can produce stripe(s) on the 3D object

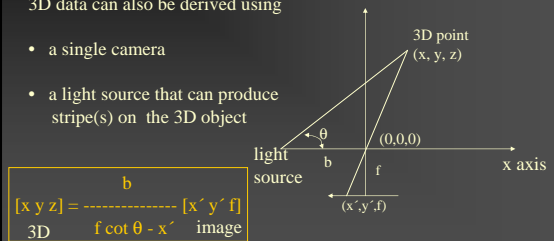


23

Structured Light 3D Computation

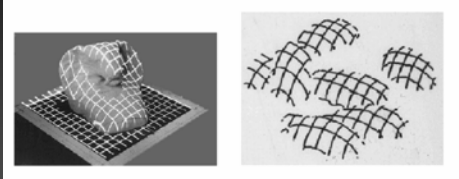
3D data can also be derived using

- a single camera
- a light source that can produce stripe(s) on the 3D object



24

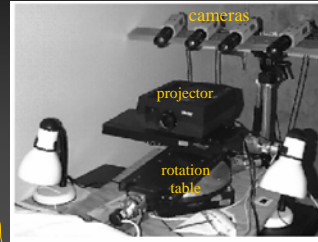
Depth from Multiple Light Stripes



What are these objects?

25

Our (former) System 4-camera light-stripping stereo



3D
object

26