Stereo (Part 2)

Introduction to Computer Vision

CSE 152

Lecture 9

Announcements

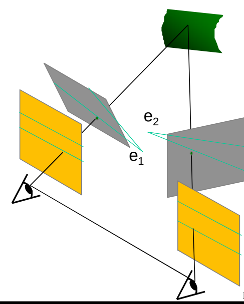
- Homework 3 is due May 10, 11:59 PM
- Reading:
  - Chapter 7: Stereopsis

Stereo Vision Outline

- Offline: Calibrate cameras & determine
B "epipolar geometry"
- Online
  1. Acquire stereo images
  C 2. Rectify images to convenient epipolar geometry
  D 3. Establish correspondence
  A 4. Estimate depth

Rectification

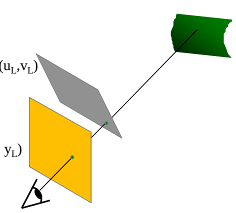
Given a pair of images, transform both images so that epipolar lines are scan lines.



Rectification

Under perspective projection, the mapping from a plane to a plane is given by a 2D projective transformation (homography)

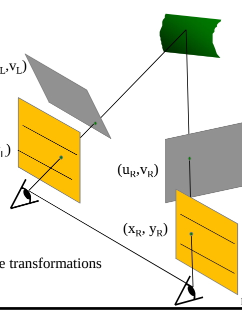
$$\begin{bmatrix} x_L \\ y_L \\ w_L \end{bmatrix} = H_L \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix}$$



Rectification

Under perspective projection, the mapping from a plane to a plane is given by a 2D projective transformation (homography)

$$\begin{bmatrix} x_L \\ y_L \\ w_L \end{bmatrix} = H_L \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix}$$



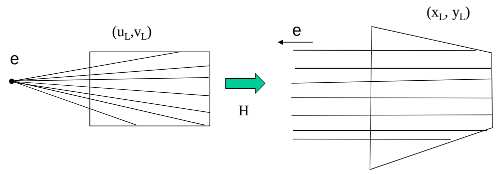
$$\begin{bmatrix} x_R \\ y_R \\ w_R \end{bmatrix} = H_R \begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix}$$

Two images – Two projective transformations

Image pair rectification

Simplify stereo matching by warping the images

Apply projective transformation so that epipolar lines correspond to horizontal scanlines



H should map epipole e to (1,0,0), a point at infinity on the x-axis

H should minimize image distortion

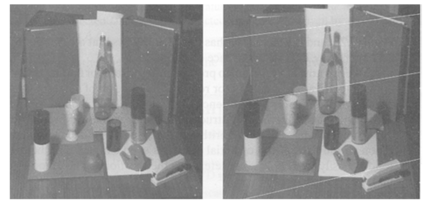
$$\begin{bmatrix}1\\0\\0\end{bmatrix} = He$$

Note that rectified images are usually not rectangular

See textbook for complete method

Rectification

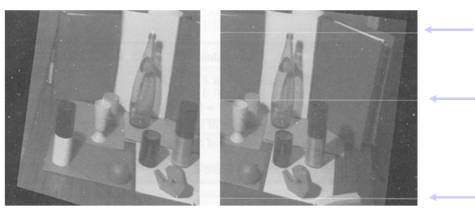
Given a pair of images, transform both images so that epipolar lines are scan lines.



Input Images

Rectification

Given a pair of images, transform both images so that epipolar lines are scan lines.

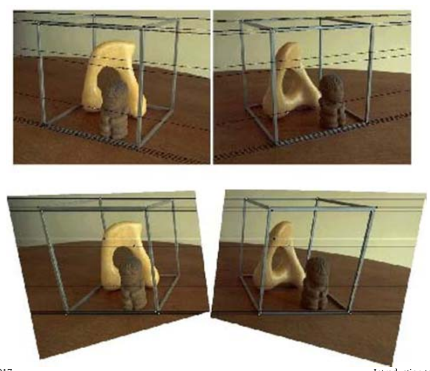


Rectified Images

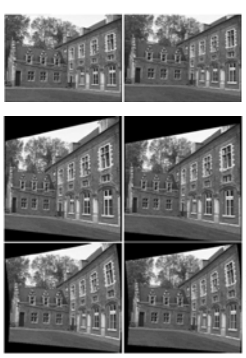
epipolar lines run parallel with the *x*-axis and are aligned between two views (no *y* disparity)

See Section 7.2.1 for specific method

Rectification

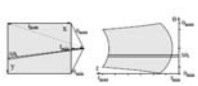


Polar Rectification



Homography-based Rectification

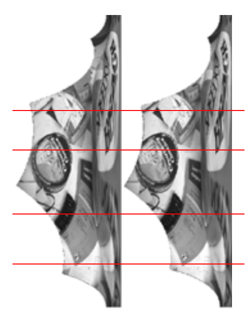
Polar Rectification



Polar Rectification

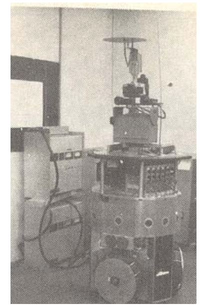
Epipoles are in images (white dot on ball)

Homography-based rectification is not possible

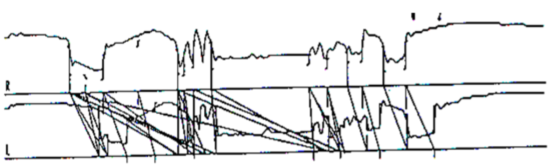


Features on same epipolar line

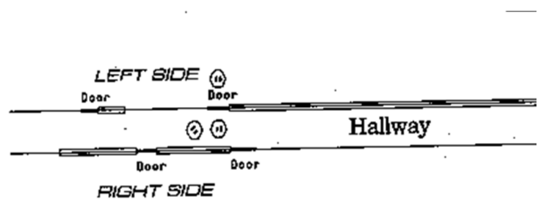
Mobi: Stereo-based navigation



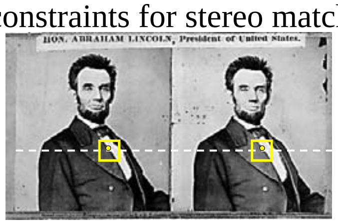
Epipolar correspondence



Symbolic Map



Using epipolar & constant Brightness constraints for stereo matching



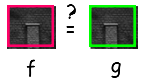
For each epipolar line
For each pixel in the left image
- compare with every pixel on same epipolar line in right image
  - pick pixel with minimum match cost
    - This will never work, so:
      match ***windows***

(Seitz)

Finding Correspondences

$W(\mathbf{p}_l)$ $W(\mathbf{p}_r)$

Comparing Windows:



$$SSD = \sum_{[i,j]\in R} (f(i,j) - g(i,j))^2$$

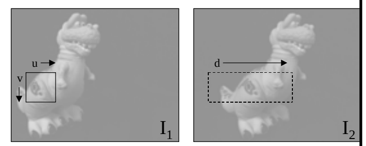
$$C_{fg} = \sum_{[i,j]\in R} f(i,j)g(i,j)$$

Most popular

For each window, match to closest window on epipolar line in other image.

(Camps)

Correspondence Search Algorithm




```
For  i = 1:nrows
  for j=1:ncols
        best(i,j) = -1
        for k = mindisparity:maxdisparity
          c = Match_Metric(I1(i,j),I2(i,j+k),winsize)
          if (c > best(i,j))
                best(i,j) = c
                disparities(i,j) = k
          end
        end
  end
end
```

O(nrows * ncols * disparities * winx * winy)

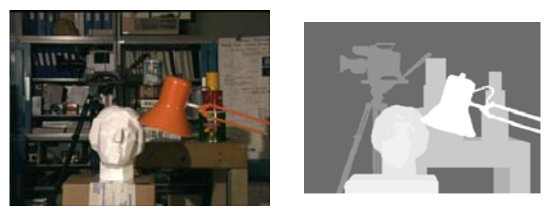
Match Metric Summary

MATCH METRIC	DEFINITION
Normalized Cross-Correlation (NCC)	$\dfrac{\sum_{u,v}(I_1(u,v)-\bar{I}_1)\cdot(I_2(u+d,v)-\bar{I}_2)}{\sqrt{\sum_{u,v}(I_1(u,v)-\bar{I}_1)^2\cdot\sum_{u,v}(I_2(u+d,v)-\bar{I}_2)^2}}$
Sum of Squared Differences (SSD)	$\sum_{u,v}(I_1(u,v)-I_2(u+d,v))^2$
Normalized SSD	$\sum_{u,v}\left(\dfrac{(I_1(u,v)-\bar{I}_1)}{\sqrt{\sum_{u,v}(I_1(u,v)-\bar{I}_1)^2}}-\dfrac{(I_2(u+d,v)-\bar{I}_2)}{\sqrt{\sum_{u,v}(I_2(u+d,v)-\bar{I}_2)^2}}\right)^2$
Sum of Absolute Differences (SAD)	$\sum_{u,v}\lvert I_1(u,v)-I_2(u+d,v)\rvert$
Zero Mean SAD	$\sum_{u,v}\lvert(I_1(u,v)-\bar{I}_1)-(I_2(u+d,v)-\bar{I}_2)\rvert$
Rank	$I_k'(u,v)=\sum_{m,n}I_k(m,n)<I_k(u,v)$ $\sum_{u,v}(I_1'(u,v)-I_2'(u+d,v))$
Census	$I_k'(u,v)=BITSTRING_{m,n}(I_k(m,n)<I_k(u,v))$ $\sum_{u,v}HAMMING(I_1'(u,v),I_2'(u+d,v))$

These two are actually the same

Stereo results

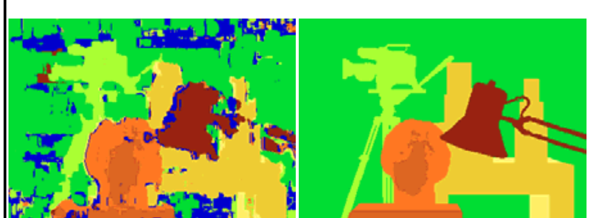
– Data from University of Tsukuba



Scene Ground truth

(Seitz)

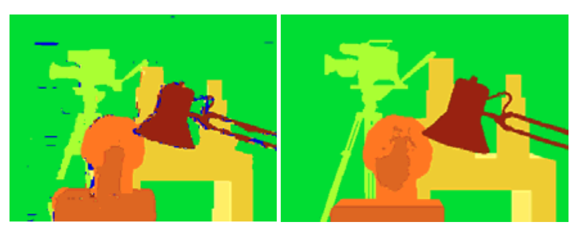
Results with window correlation



Window-based matching (best window size) Ground truth

(Seitz)

Results with better method



Using global optimization Ground truth

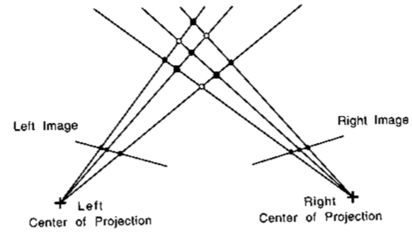
Boykov et al., Fast Approximate Energy Minimization via Graph Cuts, International Conference on Computer Vision, September 1999.

(Seitz)

Some Issues

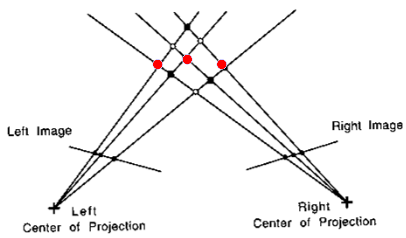
- Epipolar ordering
- Ambiguity
- Window size
- Window shape
- Lighting
- Half occluded regions

A challenge: Multiple Interpretations



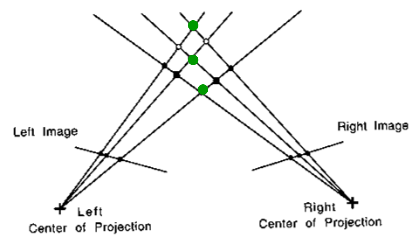
Each feature on left epipolar line match one and only one feature on right epipolar line.

Multiple Interpretations



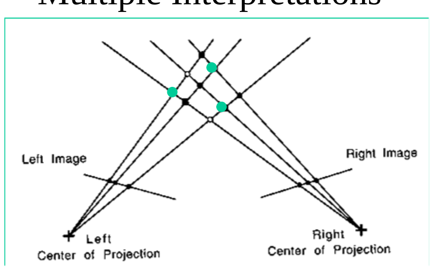
Each feature on left epipolar line match one and only one feature on right epipolar line.

Multiple Interpretations



Each feature on left epipolar line match one and only one feature on right epipolar line.

Multiple Interpretations

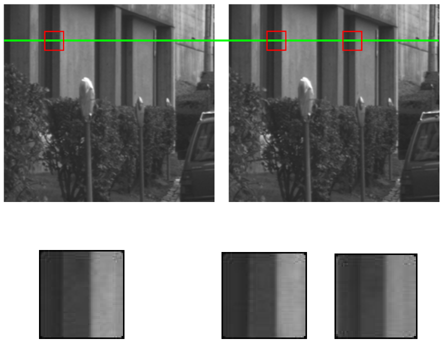


Each feature on left epipolar line match one and only one feature on right epipolar line.

Some Issues

- Epipolar ordering
- **Ambiguity**
- Window size
- Window shape
- Lighting
- Half occluded regions

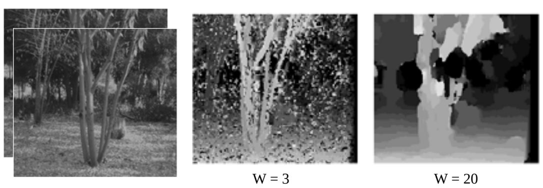
Ambiguity



Some Issues

- Epipolar ordering
- Ambiguity
- **Window size**
- Window shape
- Lighting
- Half occluded regions

Window size



- Effect of window size

Better results with *adaptive window*

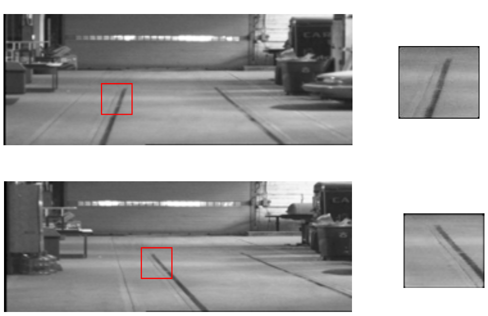
- T. Kanade and M. Okutomi, A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment,, Proc. International Conference on Robotics and Automation, 1991.
- D. Scharstein and R. Szeliski. Stereo matching with nonlinear diffusion. International Journal of Computer Vision, 28(2):155-174, July 1998

(Seitz)

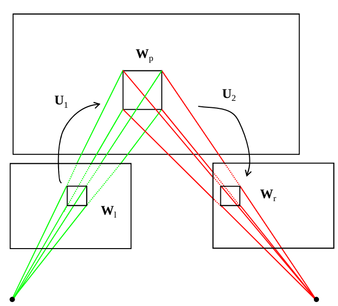
Some Issues

- Epipolar ordering
- Ambiguity
- Window size
- **Window shape**
- Lighting
- Half occluded regions

Window Shape and Forshortening



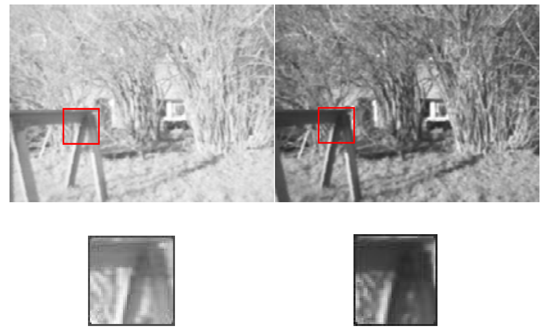
Window Shape: Fronto-parallel Configuration



Some Issues

- Epipolar ordering
- Window size
- Ambiguity
- Window shape
- **Lighting**
- Half occluded regions

Lighting Conditions (Photometric Variations)

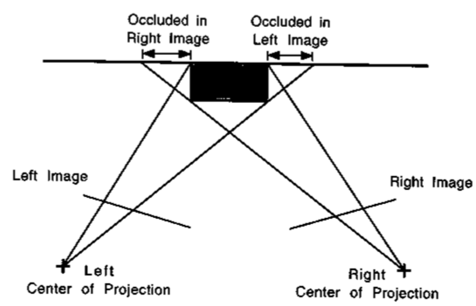


W($\mathbf{P}_l$) *W*($\mathbf{P}_r$)

Some Issues

- Epipolar ordering
- Ambiguity
- Window size
- Window shape
- Lighting
- **Half occluded regions**

Half occluded regions



Summary of Stereo Constraints

CONSTRAINT	BRIEF DESCRIPTION
1-D Epipolar Search	Arbitrary images of the same scene may be rectified based on epipolar geometry such that stereo matches lie along one-dimensional scanlines. This reduces the computational complexity and also reduces the likelihood of false matches.
Monotonic Ordering	Points along an epipolar scanline appear in the same order in both stereo images, assuming that all objects in the scene are approximately the same distance from the cameras.
Image Brightness Constancy	Assuming Lambertian surfaces, the brightness of corresponding points in stereo images are the same.
Match Uniqueness	For every point in one stereo image, there is at most one corresponding point in the other image.
Disparity Continuity	Disparities vary smoothly (i.e. disparity gradient is small) over most of the image. This assumption is violated at object boundaries.
Disparity Limit	The search space may be reduced significantly by limiting the disparity range, reducing both computational complexity and the likelihood of false matches.
Fronto-Parallel Surfaces	The implicit assumption made by area-based matching is that objects have fronto-parallel surfaces (i.e. depth is constant within the region of local support). This assumption is violated by sloping and creased surfaces.
Feature Similarity	Corresponding features must be similar (e.g. edges must have roughly the same length and orientation).
Structural Grouping	Corresponding feature groupings and their connectivity must be consistent.

(From G. Hager)

Next Lecture

- Early vision: multiple images
  - Structure from motion
- Reading:
  - Chapter 8: Structure from Motion