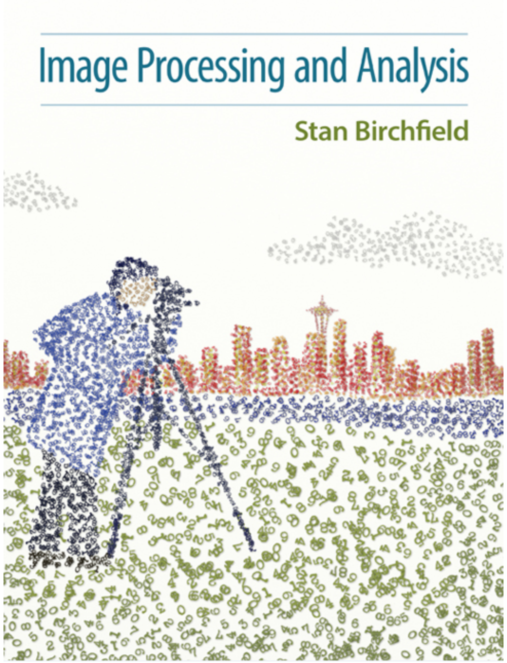
University of Stavanger

Prof. Kjersti Engan

ELE510 Image processing and computer vision

Stereopsis and correspondence problem, (chap 13.1, 13.2 Birchfield) 2020



Parts in course presentations is material from Cengage learning. It can be used for teaching, and it can be share it with students on access controlled web-sites (Canvas) for use in THIS course. **But it should not be copied or distributed further in any way** (by you or anybody).

Correspondence problem

- Given (xL,yL) where can the corresponding (xR,yR) be?
- It is constrained to be along a line called epipolar line (the epipolar constrain)
- For rectified cameras orrectified images, the scanlines are the epipolar lines.
- Need to find the corresponding points from the **matching space** (possible matches)
- Frontoparallell: an object parallell to both image planes, and thus of constant depth and disparity.
- If d=0 -> zw (depth) goes to infinity, like stars in the sky

$$d = \frac{f \cdot b}{z_w}$$

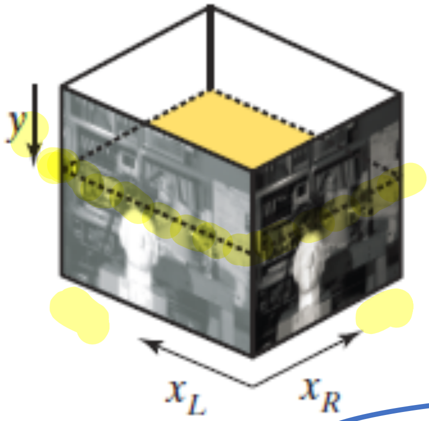
- If d is large, zw (depth) becomes small. This means object close to camera.
- usually we have d<dmax

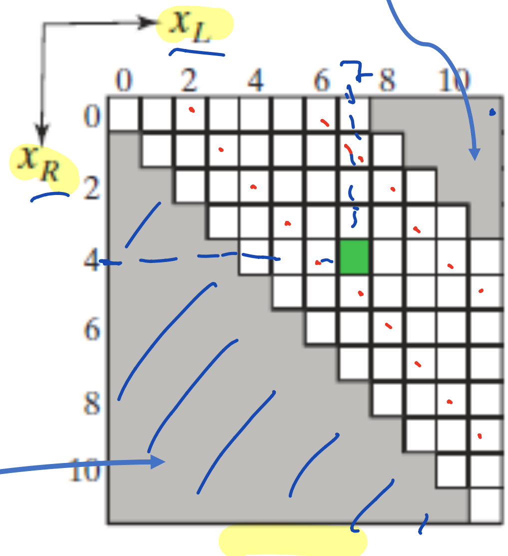
Correspondence - matching space

Maximum allowed disparity constricts this corner (here 7)

Example, matching space for rectified images.
Green represent a match between pixel xl=7 and xR=4, so that d=3.
The possible disparity is bounded by the shaded region.

The cells at the diagonal represent the matches that would occur for a frontoparallell object

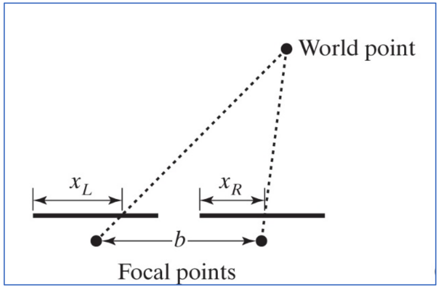




XR > XL – impossible, thus forbidden region.

Stereo constraints

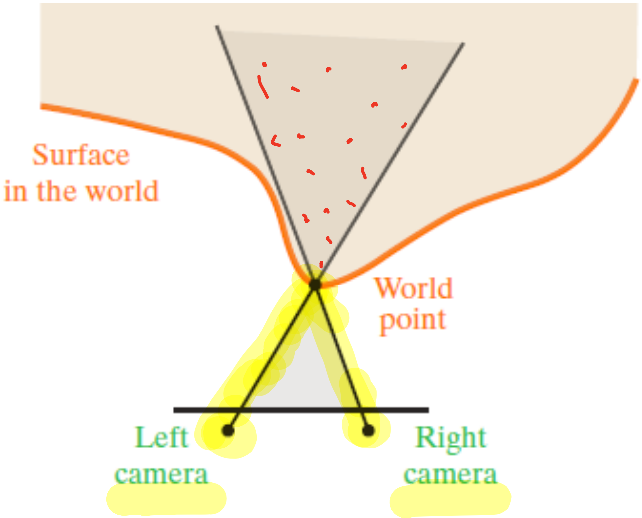
- The **cheirality constraint** requires $x_L \geq x_R$ for matching pixels since only objects in front of the camera can be visible.



- The **maximum disparity constraint** forbids matches whose disparity exceeds a certain amount, which effectively enforces a minimum distance from the camera to the surface being viewed.

- The **uniqueness constraint** says that if $\boxed{x_L \leftrightarrow x_R}$ is a match, then there is no other match $x_L \leftrightarrow x$ where $x \neq x_R$, and there is no other match $x \leftrightarrow x_R$ where $x \neq x_L$.

- **Forbidden zone:** when a point on a continuous surface is viewed by both cameras, it is not physically possible for another point on the same surface to also be visible in both cameras if it lies within the region defined by two lines passing through the centers of projection and the point.
- The forbidden zone is taken care of by the **ordering constraint**.



Stereo constraints

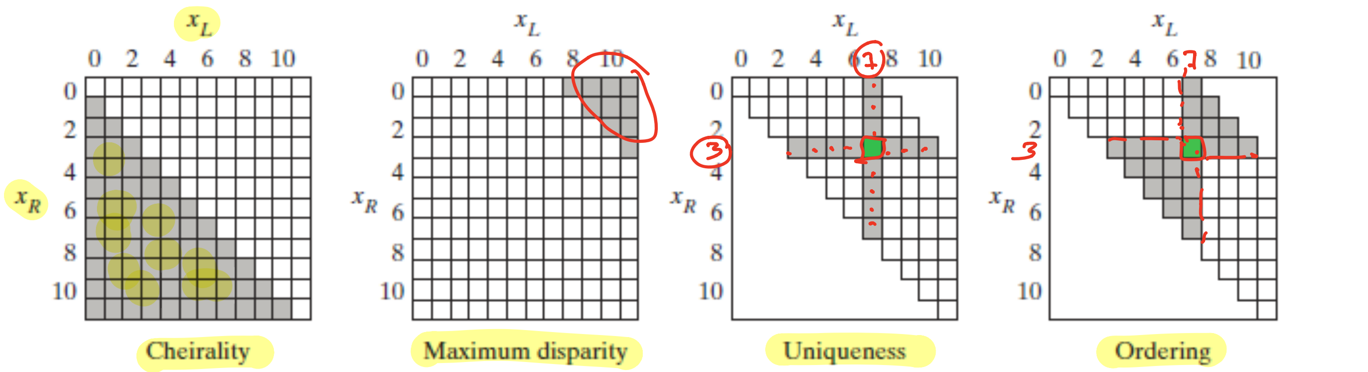


Figure 13.6 Stereo constraints. The gray cells indicate matches that are simply not allowed (left two grids) or that are illegal when the green cell indicates a match (right two grids). Cheirality precludes matches with $x_L < x_R$, which would refer to points behind the camera. Maximum disparity precludes matches whose disparity exceeds a threshold. Uniqueness prevents a pixel in either scanline from matching more than one pixel. Ordering ensures that the pixel coordinates of the matches are monotonically increasing as the pixels along either scanline are traversed. Note that the gray cells in the right grid are the forbidden zone.

Block Matching

- **Block matching** is an *area-based* approach that relies upon a statistical correlation between local intensity regions.
- For each pixel (x,y) in the left image, the right image is searched for the best match among all possible disparities $0 \leq d \leq d_{max}$ (in the matching space)
- A similarity measure is used.
- A window of possible matches is searched.

- dL: left disparity map, i.e. with respect to the left image. Can also find dR, and do left-right disparity check.

  Agree? -> OK!

  Disagree? -> unrealiable

$$d_L(x, y) = \arg \min_{0 \le d \le d_{\max}} dissim(I_L(x, y), I_R(x - d, y))$$

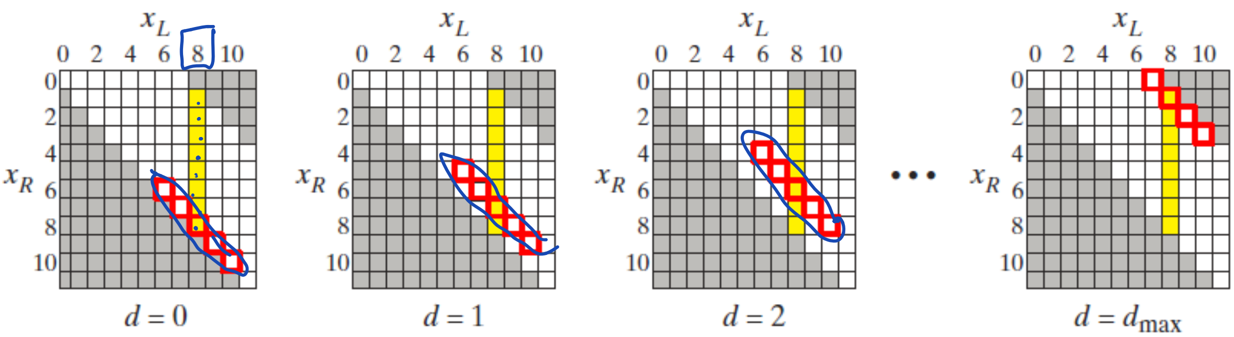


Figure 13.10 Block matching algorithm. For every pixel in the left image, a search is performed to find the disparity yielding the lowest cost. The red cells indicate the dissimilarities that are aggregated in Lines 5–6 of BLOCKMATCH1, while the yellow cells indicate the matches considered during the search. Shown is the pixel $x_L = 8$ with a window size of $w = 5$.

Dissimilarity Measures

- **Sum of absolute differences (SAD):**

$$dissim(I_L(\mathbf{x}_L), I_R(\mathbf{x}_R)) = |I_L(\mathbf{x}_L) - I_R(\mathbf{x}_R)| \qquad \text{(SAD)}$$

- **Sum of squared differences (SSD)**:

$$dissim(I_L(\mathbf{x}_L), I_R(\mathbf{x}_R)) = (I_L(\mathbf{x}_L) - I_R(\mathbf{x}_R))^2 \qquad \text{(SSD)}$$

- **Crosscorrelation, the product of their intensities:**

$$dissim(I_L(\mathbf{x}_L), I_R(\mathbf{x}_R)) = -I_L(\mathbf{x}_L)I_R(\mathbf{x}_R) \qquad \text{(cross correlation)}$$

Passive and Active Stereo

- Passive:
  - Mimic what humans do, like two cameras to estimte depth by disparity
  - Photometric stereo: two images taken by same camera, same location, but different lighting conditions
  - Depth from defocus; multiple images taken with different focal lengts

- Active
  - Light is projected onto the scene, and the light is sensed in some way
  - Laser range finder (single point)
  - Laser scanner
  - Time of flight (TOF) camera (light from LED or diode. Entire scene captured simultaneously
  - Structured light (example Kinect)