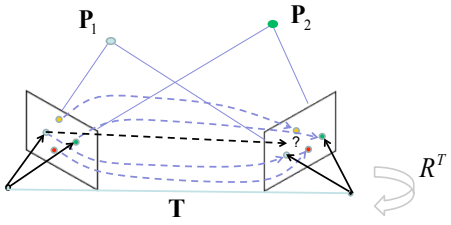
Correspondence Matching

- **Principles and challenges**
- **Region based matching**
- **Feature based matching**

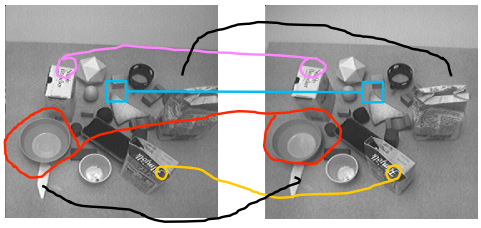
Stereo Correspondence



Key Questions

- **What to match ?**
- **Where to search for matches ?**
- **How to compare elements?**

What to Match ?



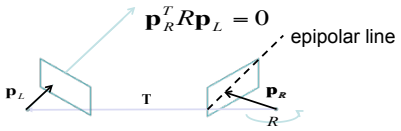
Elements to match : pixels, regions, feature points, edges and boundaries, objects....

What to Match ?

- **Pixels** : dense but not distinctive
- **Regions** : dense, more distinctive but maybe > one depth – disparity variation
- **Feature points** : sparse but distinctive
- **Edges** : sparse and ambiguous
- **Objects** : sparse, distinctive but hard to identify and match

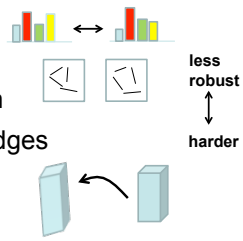
Where to Search?

- **Uncalibrated views** : correct matches could lie anywhere in other view(s)
- **Known fundamental matrix** : matches constrained to lie along epipolar lines



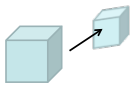
How to Compare Elements ?

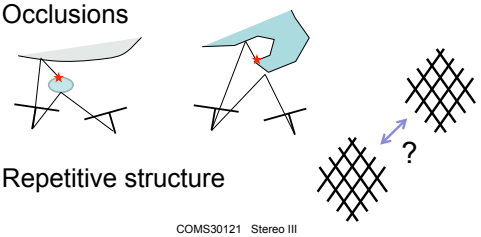
- *Depends on elements being matched*
- Pixel intensities
- Local colour variation
- Local structure variation
- Position and angle of edges
- Shape and pose



Challenges

- Noise and lighting effects
- Perspective transformations
- Occlusions

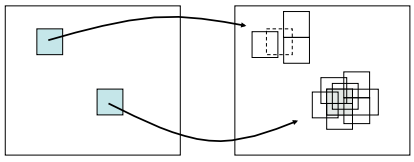




- Repetitive structure

Region-Based Methods

- Compare pixel values in image regions.
- For region in left image, compute similarity with regions of same size in right image.
- Corresponding point - centre of most similar region

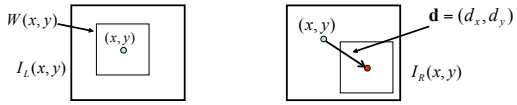


Region Matching

- Stereo image pair: $I_L(x,y)$ and $I_R(x,y)$
- For each pixel, find disparity $\mathbf{d} = (d_x, d_y)$ which minimises (or maximises) cost function

$$c(x,y,\mathbf{d}) = \sum_{(i,j)\in W(x,y)} sim(I_L(x+i,y+j), I_R(x+i+d_x, y+j+d_y))$$

where $W(x,y)$ is a window of pixels around (x,y)



Similarity Measures

- **Sum of squared differences:** $sim(u,v) = (u-v)^2$
- **Similar pixel count:**

$$sim(u,v) = \begin{cases} 1 & if\ |u-v| < T \\ 0 & else \end{cases}$$

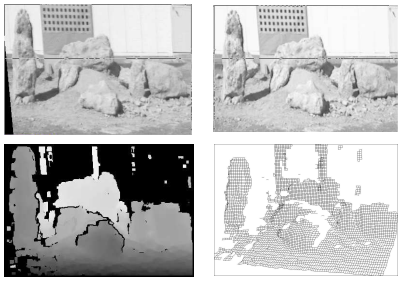
- **Normalised cross correlation**:

$$sim(u,v) = (u-\bar{u})(v-\bar{v}) / N_u N_v$$

$$\bar{u} = \frac{1}{|W(i,j)|} \sum_{k,l \in W(i,j)} u(k,l) \qquad N_u = \sqrt{\sum_{k,l \in W(i,j)} (u-\bar{u})^2}$$

mean ∝ variance

Example



Real time correlation-based stereo, Faugeras et al, 1993

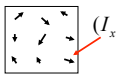
Feature-Based Methods

- Restrict search to sparse set of features
- Find salient (distinct) points in each view and match points by:
  - comparing pixels or image descriptors in local regions about each point
- Examples
  - Salient points : Harris corner detector
  - Using RANSAC for uncalibrated matching
  - Matching using the Scale-Invariant Feature Transform

Harris Corner Detector

- Covariance of spatial gradient vectors within region W

$$A = \sum_{x,y \in W} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

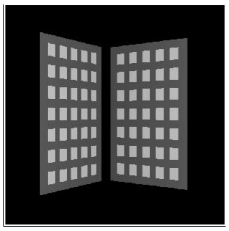


spatial gradient

$I_x \equiv I_x(x,y)$

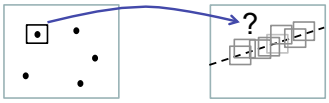
- Eigenvalues λ_1 and λ_2 of A indicate 'spread' of gradients in region, e.g. 2 high values → 'busy' region.
- Example saliency metric: $sal = \lambda_1 \lambda_2 / (\lambda_1 + \lambda_2)$
- Properties:
  - if eigenvalues both large → sal large
  - if either eigenvalue small → sal small

Examples



Calibrated Feature Matching

- Determine salient points in each image
- For each salient point in one image, find best matching point in other image :
  - region based matching about salient points
  - searching bands defined by epipolar lines

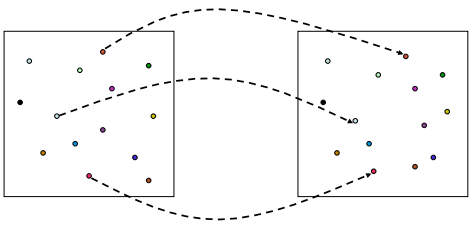


Uncalibrated Matching

- Known geometry helps matching via epipolar lines
- Need different approach when geometry unknown
- Use **RANSAC** – **RAN**dom **SA**mple **C**onsensus:
  - find potential correspondences, e.g. by comparing regions around salient points
  - select subset of matches at random (minimum 7)
  - compute fundamental matrix F from subset (lecture 9, slide 24)
  - assess support for F matrix amongst other potential correspondences using
  - repeat until best F found

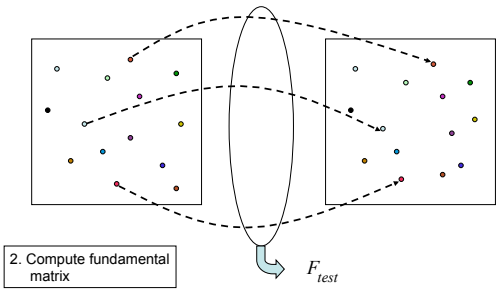
$$\mathbf{p}_R^T F \mathbf{p}_L = 0$$

RANSAC Matching



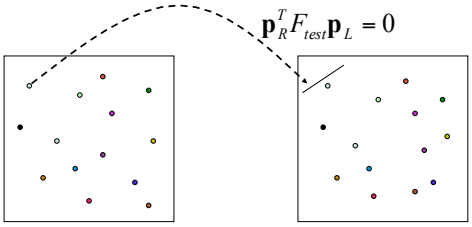
1. Random selection of potential matches

RANSAC Matching



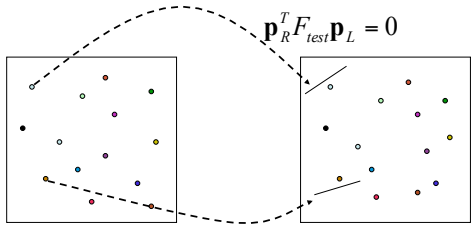
2. Compute fundamental matrix

RANSAC Matching



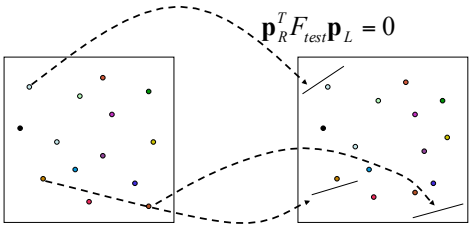
3. Compute support for fundamental matrix

RANSAC Matching



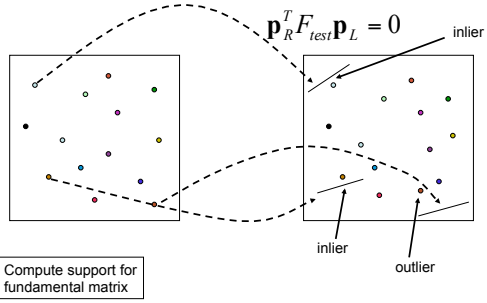
3. Compute support for fundamental matrix

RANSAC Matching



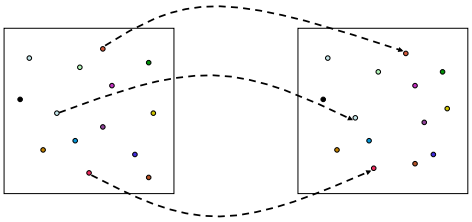
3. Compute support for fundamental matrix

RANSAC Matching



3. Compute support for fundamental matrix

RANSAC Matching



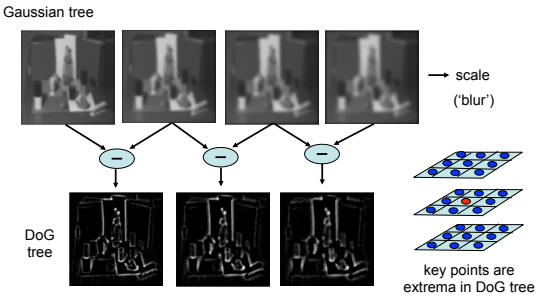
1. Random selection of potential matches

…… and so on

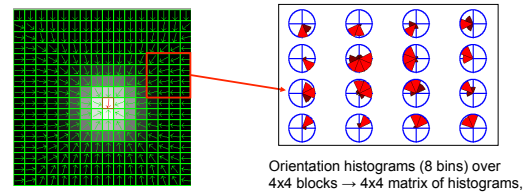
Scale-Invariant Feature Transform

- Two main elements:
  - scale invariant detection of salient (key) points
  - matching by highly distinct local descriptors
- Key point detection:
  - extrema (max or min) in difference of Gaussian blurred versions of image → DoG tree
  - points imaged at different resolutions appear at different levels of DoG tree ⟶ scale invariance
- Spatial gradient descriptors
  - built from histograms of spatial gradients in local neighbourhood
  - good rotation and perspective warp invariance properties

Difference of Gaussians



Spatial Gradient Descriptor



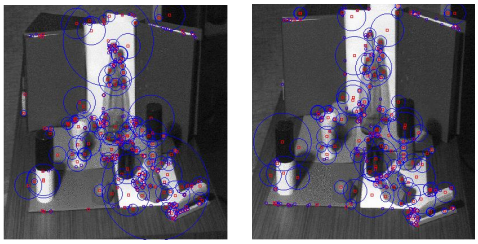
16x16 gradient field

Orientation histograms (8 bins) over 4x4 blocks → 4x4 matrix of histograms, normalised → 128-D feature vector.

Histograms built w.r.t dominant orientation(s) (potentially multiple feature vectors per point)

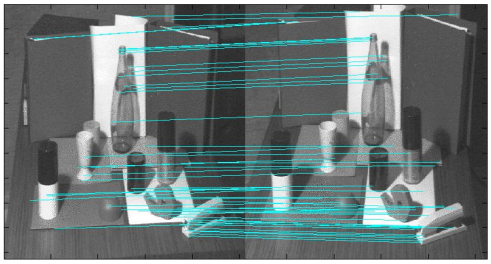
Descriptor matching based on Euclidean distance between 128-D feature vectors

SIFT Example – Selected Key Points



Each key point shown with circle indicating scale

SIFT Example – Matched Key Points



Example SIFT Result

Another SIFT Result

Yet Another SIFT result