Multi View Geometry (Spring '08)

Computing 3–View Geometry

Prof. Kyoung Mu Lee
SoEECS, Seoul National University

2008-1

Computing 3 View Geometry

- **Computation of trifocal tensor**
  - ✓ Linear method – 7 point algorithm
  - ✓ Minimal method – 6 point algorithm
  - ✓ Geometric error minimize method
  - ✓ RANSAC method

Basic equations

- Given 3-point correspondence

$$\{x_i \leftrightarrow x'_i \leftrightarrow x''_i\}$$

- The trifocal tensor relationship is

$$x^i(x'^j \epsilon_{jpr})(x''^k \epsilon_{kqs})\mathcal{T}_i^{pq} = 0_{rs}$$

  - ✓ Linear in the entries of *T*
  - ✓ Each correspondence gives 9 equations, 4 linear independent
  - ✓ *T* has 27 entries- defined up to scale
  - ✓ 7 point correspondences give 28 equations
  - ✓ Linear or least-square solution for the entries of *T*

Basic equations

- **Basic trilinear relations**
  - ✓ Constraints on T

Correspondence	Relation	#lin. indep.Eq.
Three point	$x^i(x'^j \epsilon_{jpr})(x''^k \epsilon_{kqs})\mathcal{T}_i^{pq} = 0_{rs}$	4
Two point, one line	$x^i(x'^j \epsilon_{jpr})l''_k \mathcal{T}_i^{pk} = 0_r$	2
One point, two line	$x^i l'_j l''_k \mathcal{T}_i^{jk} = 0$	1
Three lines	$(l_r \epsilon^{ris}) l'_j l''_k \mathcal{T}_i^{jk} = 0^s$	2

Normalized linear algorithm

- Given 26 equations we can solve for the 27 entries of T
  - ✓ 7 point correspondences
  - ✓ Or 13 line correspondences
  - ✓ Or other mixture of point & lines

- The set of equations

$$\boxed{\mathbf{At} = \mathbf{0}}$$

  - ✓ With 26 equations $\longrightarrow$ an exact solution
  - ✓ More equations $\longrightarrow$ least square solution

Normalized linear algorithm

- Solution:
  - ✓ SVD : $\mathbf{A} = \mathbf{UDV}^T$
  - ✓ Solution is the last column of **V** corresponding to the smallest e-value
  - ✓ Minimizes $\|\mathbf{At}\|$ subject to $\|\mathbf{t}\| = 1$
  - ✓ Data normalization is essential

Normalized linear algorithm

- **Normalized Linear algorithm**

Objective
Given $n \geq 7$ image point correspondences across 3 images, or a least 13 lines, or a mixture of point and line corresp., compute the trifocal tensor.

Algorithm
(i) Find transformation matrices **H**,**H'**,**H"** to normalize 3 images
(ii) Transform points with **H** and lines with $\mathbf{H}^{-1}$
(iii) Compute trifocal tensor **T** from **At**=**0** (using SVD)
(iv) Denormalize trifocal tensor $T_i^{jk} = \mathbf{H}_i^r \left(\mathbf{H}'^{-1}\right)_s^j \left(\mathbf{H}''^{-1}\right)_t^k \hat{T}_r^{st}$

Internal constraints on T

- T has 27 coefficients up to scale
- Geometry has only 18 (33-15) DOF
  - ✓ 3 camera matrices, 3 x 11 = 33 DOF
  - ✓ invariant under projective transform, 15 DOF
- Note that not every 3x3x3 tensor is a valid trifocal tensor. T must satisfy several intrinsic constraints (like fundamental matrix case)
- The 8 constraints are not easy to represent explicitly
- Trifocal tensor satisfies all intrinsic constraints if it corresponds to three cameras {**P**, **P'**, **P"**}
- Constraints can be imposed through parameterization

Constraints through parameterization

- Define T in terms of a set of parameters
- Only valid Ts may be generated from parameters

$$T_i^{jk} = a_i^j b_4^k - a_4^j b_i^k$$

- Only valid trifocal tensors are generated by this formula
- Parameters are the entries a_j^i and b_i^k
- Over-parameterized: 24 parameters in all

Algebraic minimization method

- Minimization knowing the epipoles

$$T_i^{jk} = a_i^j b_4^k - a_4^j b_i^k$$

- a_4 and b_4 are the epipoles of the first camera
- If a_4 and b_4 are known, then T is linear in terms of the other parameters

$$\mathbf{t} = \mathbf{E}\mathbf{a}$$

✓ $\mathbf{a}$ is the vector of the remaining entries a_j^i and b_i^k
✓ $\mathbf{t}$ is the 27-vector of entries of T
✓ $\mathbf{E}$ is a 27x18 matrix

Algebraic Minimization method

- Minimization

$$\text{Minimize } \|\mathbf{At}\| \text{ subject to } \|\mathbf{t}\| = 1$$
$$\Updownarrow$$
$$\text{Minimize } \|\mathbf{AEa}\| \text{ subject to } \|\mathbf{Ea}\| = 1$$

- Same as with the Fundamental matrix
- Solution using SVD

Algebraic Minimization method

- Algebraic estimation algorithm

(i) Find a solution for T using the normalized linear (7-point) method

(ii) Compute the epipoles a_4 and b_4
   - (i) Find the left (respectively right) null spaces of T^k
   - (ii) Epipole is the common perpendicular to the null spaces

(iii) Re-estimate T by algebraic method assuming values for epipoles

- Iterative algebraic minimization method (Algorithm 16.2)
  - Vary epipoles to minimize the algebraic error
  - Produces excellent results

Geometric distance

- Maximum Likelihood Estimation
  - ✓ Given correspondences

$$\mathbf{x}_i \leftrightarrow \mathbf{x}'_i \leftrightarrow \mathbf{x}''_i$$

  - ✓ Cost function (reprojection error)

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 + d(\mathbf{x}''_i, \hat{\mathbf{x}}''_i)^2$$

  - ✓ Parameterization

$$\hat{\mathbf{x}}_i = [\mathbf{I} \mid \mathbf{0}]\mathbf{X}_i, \hat{\mathbf{x}}'_i = \mathbf{P}'\mathbf{X}_i, \hat{\mathbf{x}}''_i = \mathbf{P}''\mathbf{X}_i \quad \text{(24 parameters+3}N\text{)}$$

- Also possible to Sampson error (24 parameters)

Automatic Computation of T

<u>Objective</u>

Compute the trifocal tensor between three images

<u>Algorithm</u>

(i) **Interest points:** Compute interest points in each image

(ii) **Putative correspondences:** Compute interest correspondences (and **F**) between 1&2 and 2&3

(iii) **RANSAC robust estimation:** Repeat for N samples

(a) Select at random 6 (or 7) correspondences and compute $\boldsymbol{T}$

(b) Calculate the distance $d_\perp$ for each putative match

(c) Compute the number of inliers consistent with $\boldsymbol{T}$ ($d_\perp < t$)

Choose $\boldsymbol{T}$ with most inliers

(iv) **Optimal estimation:** re-estimate $\boldsymbol{T}$ from all inliers by minimizing ML cost function with Levenberg-Marquardt

(v) **Guided matching**: Determine more matches using prediction by computed $\boldsymbol{T}$

Optionally iterate last two steps until convergence

Automatic Computation of T

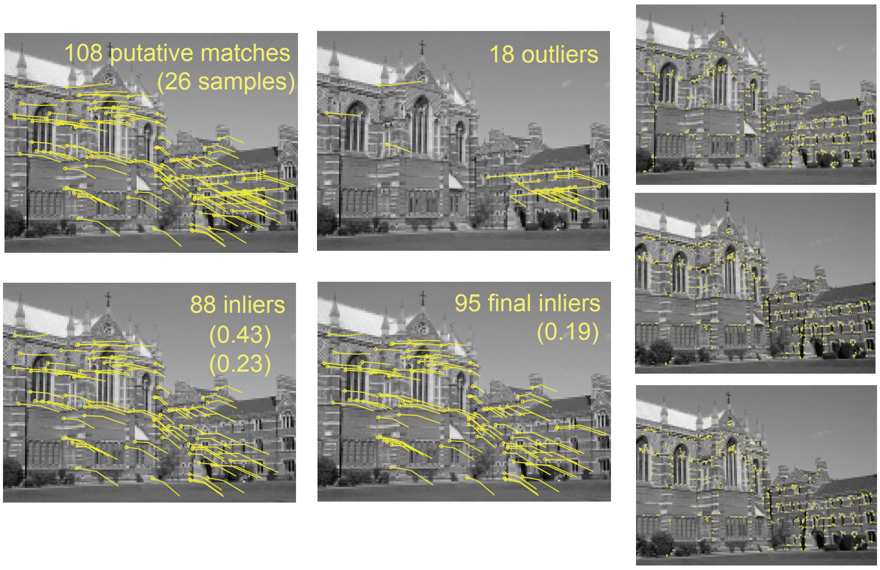
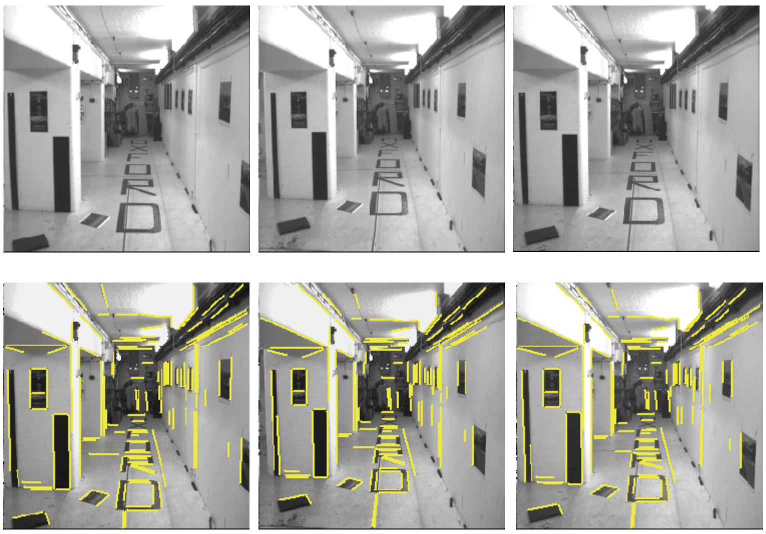


Image triplet matching

- With T , line features can be matched [Schmid-97]



Automatic estimation of a projective reconstruction for a sequence

Outline of the procedure

- Obtain putative correspondences and **F** and *T* over consecutive 2/3-views
- Compute initial reconstruction (still black art)
  - ✓ sequential reconstruction method
  - ✓ hierarchical reconstruction method
- Refine structure and motion
  - ✓ bundle adjustment

$$\min_{\hat{\mathbf{P}}^i, \hat{\mathbf{X}}_j} \sum_{i \in \text{points}} \sum_{j = \text{frames}} d\left(\mathbf{x}_j^i, \hat{\mathbf{P}}^i \hat{\mathbf{X}}_j\right)^2$$

- Auto-calibration (optional)
- Stratification (optional)

Sequential reconstruction

- Initialize structure and motion from 2 views
- For each additional view
  - ✓ Determine pose
  - ✓ Refine and extend structure

Sequential reconstruction

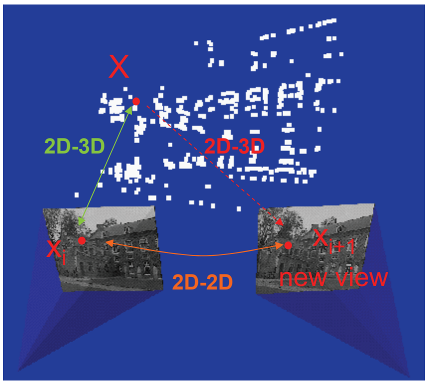
- Initial structure and camera matrices

$$\mathbf{F} \Leftrightarrow \begin{array}{l} \mathbf{P}_1 = [\mathbf{I} \quad \mathbf{0}] \\ \mathbf{P}_2 = [[\mathbf{e}]_x \mathbf{F} + \mathbf{e}\mathbf{a}^T \quad \mathbf{e}] \end{array}$$

- Compute structure through triangulation
- Determine pose for additional view

$$\mathbf{x}_{i+1} = \mathbf{P}_{i+1}\mathbf{X}$$

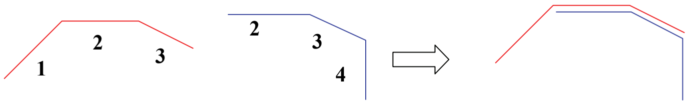
- Compute $\mathbf{P}_{i+1}$ using robust approach
- Extend and refine reconstruction



Ref: M. Pollefey, "tutorial on 3D modeling from images" (http://www.cs.unc.edu/%7Emarc/tutorial/index.htm

Hierarchical reconstruction

- Compute all 2 view reconstruction for consecutive views
- Compute all 3 view reconstruction for consecutive views
- Stitch 3-views reconstructions by hierarchical merging



- Refine by bundle-adjustment (minimizing reprojection error)
- Automatic method:
  - ✓ A. Fitzgibbon and A. Zisserman, “Automatic camera recovery for closed or open image sequences”, in proc. of ECCV, pp. 311-326, June, 1998