Two view odometry

D.A. Forsyth

University of Illinois at Urbana Champaign

Monocular visual odometry

- A calibrated camera
  - views a static scene,
  - moves,
  - views again

- Q: how did it move?
  - We care, because this allows us to recover movement from single cameras
  - We should be able to tell
    - Recall epipoles etc are quite informative about movement

Visual odometry

- Use eight point algorithm, recover fundamental matrix
- Recall:

$$\mathcal{F} = k\mathcal{K}_L^{-T}\left[\mathcal{T}_x\mathcal{R}\right]\mathcal{K}_R^{-1}$$

- But we know calibration, which yields essential matrix

$$\mathcal{E} = k\mathcal{T}_x\mathcal{R}$$

- Q: what can we get out of essential matrix?

Visual odometry, II

Antisymmetric

$$\mathcal{E} = k\mathcal{T}_x\mathcal{R}$$

Unknown constant

Rotation

Translation

The essential matrix you estimate $\hat{\mathcal{E}}$ is

$$\hat{\mathcal{E}} = s[\mathbf{t}]_X \mathcal{R}$$

(for some unknown $s \neq 0$). Further $\mathbf{t}^T[\mathbf{t}]_X = \mathbf{0}$, so you can immediately recover the translation $\mathbf{t}$ *up to scale* by finding the unit vector $\mathbf{u}$ such that

$$\mathbf{u}^T \hat{\mathcal{E}} = \mathbf{0}^T.$$

This vector is occasionally referred to as the *left null vector* of the essential matrix. It is an estimate up to scale of the translation $\mathbf{t}$. Notice there are two unit left null vectors because $-\mathbf{u}$ is also a unit vector and is also a left null vector. Notice there is also a unit vector $\mathbf{s}$ such that

$$\hat{\mathcal{E}}\mathbf{s} = \mathbf{0}.$$

This is the *right null vector* of the essential matrix; again, there are two.

Rotation

Choose one of the left null vectors, and call it $\mathbf{u}$. Then

$$\hat{\mathcal{E}} = [\mathbf{u}]_X \mathcal{R}$$

for an unknown $\mathcal{R}$. Take the singular value decomposition of $[\mathbf{u}]_X$ to get

$$[\mathbf{u}]_X = \mathcal{U}_u \Sigma_e \mathcal{V}_u^T$$

where $\Sigma_e = \text{diag}(1, 1, 0)$ **exercises** . Take the singular value decomposition of $\hat{\mathcal{E}}$ to get

$$\hat{\mathcal{E}} = \mathcal{U}_e \Sigma_e \mathcal{V}_e^T.$$

Now

$$\hat{\mathcal{E}} = [\mathbf{u}]_X \mathcal{R} = \mathcal{U}_u \Sigma_e \mathcal{V}_u^T \mathcal{R} = \mathcal{U}_e \Sigma_e \mathcal{V}_e^T.$$

Conclude that

$$\mathcal{R} = \mathcal{V}_u \mathcal{V}_e^T.$$

However, there are important ambiguities, because Σ_e does not have full rank.

Ambiguities - I

Because $[\mathbf{u}]_X \mathbf{u} = \mathbf{0}$

$$\mathcal{V}_u^T \mathbf{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(**exercises**) so you can write

$$\mathcal{V}_u = [\mathbf{a}_1 \mathbf{a}_2 \mathbf{u}] \, .$$

Because $\hat{\mathcal{E}} \mathbf{s} = \mathbf{0}$

$$\mathcal{V}_e^T \mathbf{s} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(**exercises**) so you can write

$$\mathcal{V}_e = [\mathbf{b}_1 \mathbf{b}_2 \mathbf{s}] \, .$$

Ambiguities - II

$$\mathcal{V}_u = [\mathbf{a}_1 \mathbf{a}_2 \mathbf{u}]$$

$$\mathcal{V}_e = [\mathbf{b}_1 \mathbf{b}_2 \mathbf{s}]$$

Now multiply out these two expressions to obtain

$$\mathcal{R} = \mathbf{a}_1 \mathbf{b}_1^T + \mathbf{a}_2 \mathbf{b}_2^T + \mathbf{u}\mathbf{s}^T$$

The sign ambiguity means there are two possible versions of $\mathbf{u}$ and two possible versions of $\mathbf{s}$. However, there are only two possible versions of $\mathbf{u}\mathbf{s}^T$ (two negatives are the same as two positives, **exercises**). Now construct the two matrices

$$\mathcal{W}_+ = \mathbf{a}_1 \mathbf{b}_1^T + \mathbf{a}_2 \mathbf{b}_2^T + \hat{\mathbf{u}}\hat{\mathbf{s}}^T$$

and

$$\mathcal{W}_- = \mathbf{a}_1 \mathbf{b}_1^T + \mathbf{a}_2 \mathbf{b}_2^T - \hat{\mathbf{u}}\hat{\mathbf{s}}^T.$$

Ambiguities - III

$$\mathcal{W}_+ = \mathbf{a}_1\mathbf{b}_1^T + \mathbf{a}_2\mathbf{b}_2^T + \hat{\mathbf{u}}\hat{\mathbf{s}}^T$$

and

$$\mathcal{W}_- = \mathbf{a}_1\mathbf{b}_1^T + \mathbf{a}_2\mathbf{b}_2^T - \hat{\mathbf{u}}\hat{\mathbf{s}}^T.$$

and write

$$\hat{\mathcal{R}}_+ = \mathcal{W}_+(\det(\mathcal{W}_+))$$

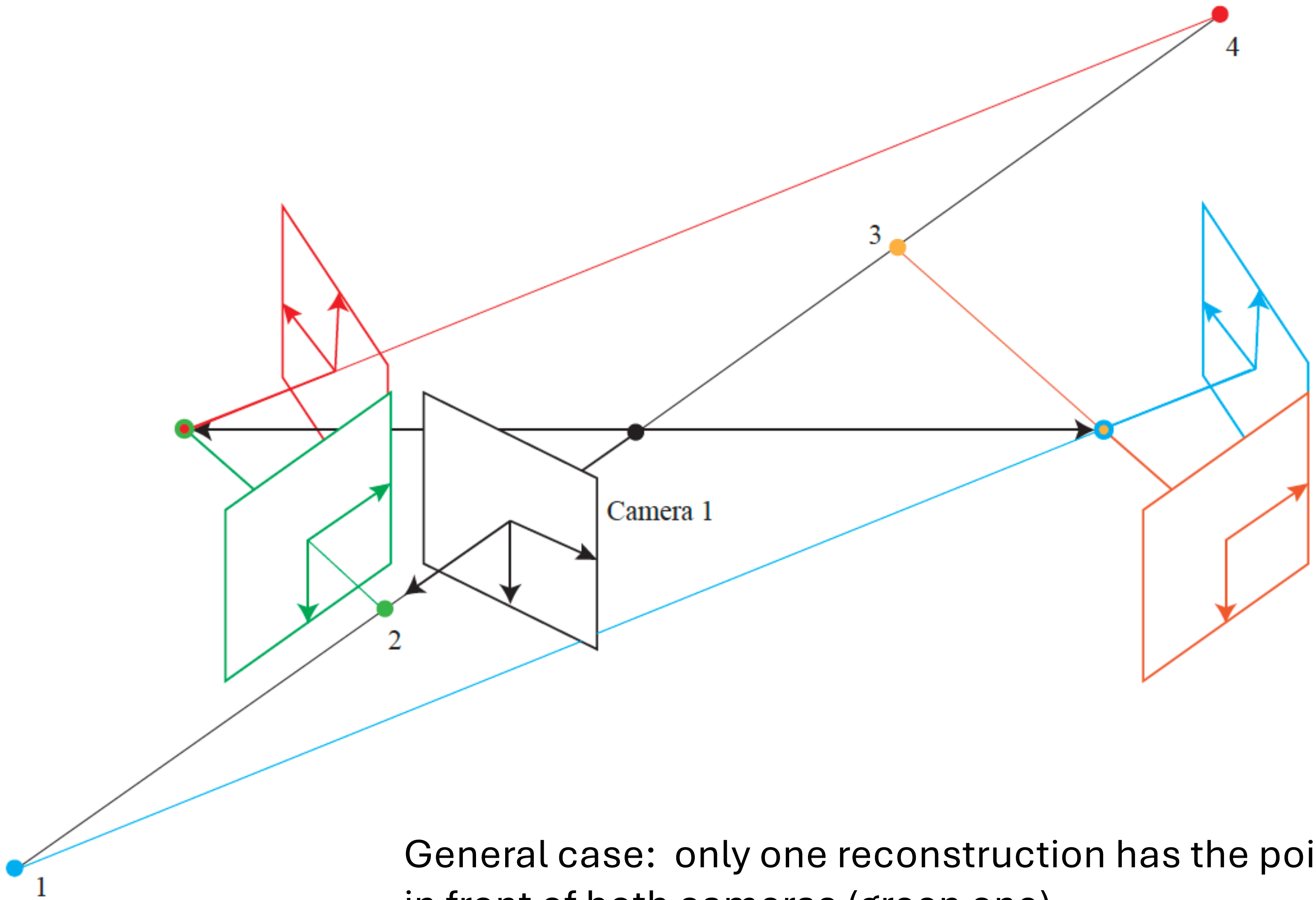
$$\hat{\mathcal{R}}_- = \mathcal{W}_-(\det(\mathcal{W}_-)).$$

These are (a) true rotations, because their determinants are positive and (b) estimates of the rotation. Recall that there are two possible estimates of the translation, $\hat{\mathbf{u}}$ and $-\hat{\mathbf{u}}$. Then the essential matrix yields four possible *distinct* camera configurations. One of

$$(\hat{\mathbf{u}}, \hat{\mathcal{R}}_+), (-\hat{\mathbf{u}}, \hat{\mathcal{R}}_+), (\hat{\mathbf{u}}, \hat{\mathcal{R}}_-), (-\hat{\mathbf{u}}, \hat{\mathcal{R}}_-)$$

is the correct pair of (scaled translation, rotation).

Ambiguities - IV



General case: only one reconstruction has the point in front of both cameras (green one)

Visual odometry, VII

- So we can recover
  - Rotation exactly
  - Translation up to scale

- From an essential matrix
  - which we can recover from correspondences
    - obtained using RANSAC+eight point algorithm
  - and camera calibration matrices