Projective Depth, Factorization & Plane + Parallax

Bill Triggs

INRIA Rhône-Alpes

Grenoble, France

http://www.inrialpes.fr/movi/people/Triggs

Contents

Projective Depth and Factorization

- Image projection of a 3D point $\mathbf{X}$ by a 3×4 camera matrix $\mathbf{P}$:

$$\boxed{\lambda \boldsymbol{x} \;=\; \mathbf{P}\,\mathbf{X}}$$

- The scale factors λ are called a **projective depths**

— they are only defined modulo normalization of $\boldsymbol{x}, \mathbf{P}, \mathbf{X}$

— but they still have an ***internal coherence*** that encodes 3D structure.

Projective SFM by Matrix Factorization

- Any method of recovering a self-consistent set of λ's for n points across m images already gives an implicit 3D reconstruction.
- We can consolidate this and read it off using **matrix factorization** :

$$\boxed{\begin{pmatrix} \lambda_{11}\boldsymbol{x}_{11} & \dots & \lambda_{1n}\boldsymbol{x}_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1}\boldsymbol{x}_{m1} & \dots & \lambda_{mn}\boldsymbol{x}_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_m \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 & \dots & \mathbf{X}_n \end{pmatrix}}$$

- The factorization is only defined up to a 4×4 transformation

$$\boxed{\mathbf{X}_p \to \mathbf{H}\mathbf{X}_p \qquad \mathbf{P}_i \to \mathbf{P}_i\mathbf{H}^{-1}}$$

— *i.e.* a change of projective 3D coordinates.

Matching Tensors

- The projection matrices depend on the 3D frame : $\mathbf{P} \to \mathbf{P}\,\mathbf{H}^{-1}$.
- Eliminating the dependence using determinants gives purely image-based **matching tensors**.
- These encode the projective camera geometry as inter-image constraints.
- Their coefficients are 4×4 determinants of projection matrix rows.
- They can be estimated quasi-linearly from the image correspondences

— but only by ignoring some algebraic consistency constraints.

- There is a hidden **scale coherence** linking different tensors

— many tensor applications need to recover this

— given the tensors and image correspondences, a consistent scaling can be recovered linearly.

Depth Recovery Relations

- Projecting a 3D point using the closure constraint projections gives inter-image **depth recovery relations** linking the image features and their projective depths :

$$\begin{array}{rcll} \boldsymbol{F}_{ij}\;(\lambda_j \boldsymbol{x}_j) + \boldsymbol{e}_{ji} \wedge (\lambda_i \boldsymbol{x}_i) & = & \boldsymbol{0} & \text{epipolar depth recovery} \\ \boldsymbol{T}_i{}^{jk}\,(\lambda_i \boldsymbol{x}_i) - (\lambda_j \boldsymbol{x}_j)\;(\boldsymbol{e}_{ik})^\top + \boldsymbol{e}_{ij}\;(\lambda_k \boldsymbol{x}_k)^\top & = & \boldsymbol{0} & \text{trifocal depth recovery} \end{array}$$

- To recover depths for projective factorization, use a tree of these equations to link the feature depths across all images
- The result is a very stable multi-image projective reconstruction algorithm.

Closure Constraints & Projection Reconstruction

- The matching tensors' dependence on the projections implies consistency conditions called **closure constraints**. The simplest two are :

$$\begin{array}{rcll} \boldsymbol{F}_{ij}\,\mathbf{P}_j + [\boldsymbol{e}_{ji}]_\times\,\mathbf{P}_i & = & \boldsymbol{0} & \text{epipolar closure} \\ \boldsymbol{T}_i{}^{jk} \cdot \mathbf{P}_i - \mathbf{P}_j \otimes (\boldsymbol{e}_{ik}) + \boldsymbol{e}_{ij} \otimes \mathbf{P}_k & = & \boldsymbol{0} & \text{trifocal closure} \end{array}$$

- These can be used to estimate projection matrices linearly from the matching tensors [Triggs BMVC'96,IVC'97]
- Coherent scaling of the tensors is essential — use the below *depth recovery relations* around a cycle of images.

Avoiding Depth Recovery

- Sometimes the depths come for free, *e.g.* for **affine cameras**.
- A camera is **affine** iff its optical plane is the plane at infinity

— *i.e.* the third row of its camera matrix is $(0\;0\;0\;1)$.

- Finite 3D points $(x\;y\;z\;1)^\top$ project affinely to image points $(u\;v\;1)$ ***with unit scale factor*** $\lambda = 1$.
- For affine cameras, there is no need for depth recovery :

— simply stacking corresponding image points into a vector already gives an implicit linear reconstruction

— factorization can be used to consolidate and concretize this as usual.

“Pseudo-Affine” Cameras

• For any set of projective cameras with a ***common optical plane***, any affine reconstruction method gives ***projectively exact results***

— the cameras *become* affine in any projective frame that places their common optical plane at infinity.

• A camera geometry is pseudo-affine if and only if all of its epipoles lie on a common line in each image.

Pseudo-Affine Normalization

• For reconstruction by pseudo-affine factorization, homographies are not actually needed — just normalize the input points affinely against the line $\boldsymbol{l}$ of epipoles:

$$\boxed{\boldsymbol{l} \cdot \boldsymbol{x} \;=\; 1 \qquad \text{for all } \boldsymbol{x}}$$

• Even if the epipoles are not quite aligned, normalizing points against a line passing close to them can significantly improve the accuracy of the pseudo-affine approximation.

Rectification to Pseudo-Affinity

• For any set of projective cameras with ***coplanar optical centres***, the images can be rectified so that the projections share a common optical plane

— this allows pseudo-affine methods to be used.

1. For each image, choose any homography that moves its line of epipoles to the line at infinity
2. Normalize the output image points in the usual affine way $(u\ v\ 1)^\top$
3. Any affine reconstruction method now gives projective structure.

• This applies to *any* two or three image geometry.

• The affinizing normalization is singular for points on the common plane, so pseudo-affinization fails if the plane cuts through the scene

— *e.g.* a horizontal camera in a moving car.

Measuring Non-Pseudo-Affinity

• With affine image normalization, the depth of a 3D point $\mathbf{X}$ in a camera with optical plane $\mathbf{p}$ is $\lambda \;=\mathbf{p}\cdot\mathbf{X}$.

• The projective depth cross ratio of two points in two images is thus :

$$\boxed{\frac{\lambda_{11}\,\lambda_{22}}{\lambda_{21}\,\lambda_{12}} \;=\; \frac{(\mathbf{p}_1\cdot\mathbf{X}_1)\,(\mathbf{p}_2\cdot\mathbf{X}_2)}{(\mathbf{p}_2\cdot\mathbf{X}_1)\,(\mathbf{p}_1\cdot\mathbf{X}_2)}}$$

• Pseudo-affine geometries have all cross-ratios 1, so this is a projective measure of the degree of non-pseudo-affinity, independent of the point and plane normalizations and 3D coordinates.

• Expanding to first order with $\mathbf{p}_i = (\boldsymbol{n}_i \;\; d_i)$ and $\mathbf{X} = (\boldsymbol{x} \; 1)^\top$ gives

$$\frac{(\boldsymbol{n}_1 \cdot \boldsymbol{x}_1 + d_1)(\boldsymbol{n}_2 \cdot \boldsymbol{x}_2 + d_2)}{(\boldsymbol{n}_2 \cdot \boldsymbol{x}_1 + d_2)(\boldsymbol{n}_1 \boldsymbol{x}_2 + d_1)} \approx 1 + \left(\frac{\boldsymbol{n}_1}{d_1} - \frac{\boldsymbol{n}_2}{d_2}\right) \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2) + \mathcal{O}\left(\left(\frac{\boldsymbol{n} \cdot \boldsymbol{x}}{d}\right)^2\right)$$

• So the pseudo-affine approximation is good for :

1 Small objects in distant cameras.

2 Cameras with similar orientations and distances, even for large scenes.

Plane + Parallax

• The key to *3D* camera geometry is the centres of projection.

• Camera rotations & calibration changes produce only "trivial" *image* deformations — *2D* homographies.

• To projectively cancel these out, "fixate" on a given 3D plane by homographically warping all images into the chosen planar coordinate system.

• The ***residual parallax*** (disparity) gives the "projective distance" of points from the plane.

Advantages of plane + parallax

• Displacements relative to the plane are often small, so feature detection and image processing are easier.

• The 3D and matching tensor geometry becomes *much* simpler.

• The ***Carlsson duality*** between 3D points and camera centres becomes very explicit and easy to use.

Standard Coordinates for Plane + Parallax

• Place the coordinatized reference plane at projective infinity and a chosen reference point (not on the plane) at the origin

• This fixes the projective coordinates up to a single global scale factor.

• Plane-aligned cameras are parametrized by their projection centres alone:

$$\mathbf{P}_i \simeq (\boldsymbol{I}_{3\times3} \,|\, -\boldsymbol{c}_i) \quad \text{with centre of projection} \quad \begin{pmatrix} \boldsymbol{c}_i \\ 1 \end{pmatrix}$$

• The geometry is projectively equivalent to ***purely translating cameras***.

• We can view the origin as "camera 0" with $\boldsymbol{c}_0 = \boldsymbol{0}$

— moving the origin just translates all centres of projection : $\boldsymbol{c}_i \to \boldsymbol{c}_i + \boldsymbol{t}$.

3D Points in Plane + Parallax

• 3D points are linear combinations of points on the plane and the reference point/origin :

$$\mathbf{X} \equiv \binom{\boldsymbol{x}}{w} = \binom{\boldsymbol{x}}{0} + w \binom{\boldsymbol{0}}{1}$$

• Projected point vectors are displaced towards their centres of projection proportionally to w :

$$\mathbf{P}_i \mathbf{X} = \boldsymbol{x} - w\,\boldsymbol{c}_i$$

• w is called **projective height** off the plane — $w = 0$ on the plane, $w = \infty$ at the origin.

• w depends on the normalization convention for planar points $\boldsymbol{x}$ and the choice of global scale (which affects $\boldsymbol{c}_i$)

3D Lines in Plane + Parallax

- Any 3D line $\mathbf{L}$ can be parametrized by $(\boldsymbol{l}, z)$ with $\boldsymbol{l} \cdot z = 0$, where :

— $\boldsymbol{l}$ is $\mathbf{L}$'s the projection from the origin onto the reference plane

— z is $\mathbf{L}$'s intersection with the reference plane (this lies on $\boldsymbol{l}$)

- The relative scale of $\boldsymbol{l}$ and z gives $\mathbf{L}$'s 'steepness' — lines on the plane have $z = \boldsymbol{0}$, the ray from the origin to z has $\boldsymbol{l} = \boldsymbol{0}$

- In image i, $\mathbf{L}$ projects to $\boxed{\boldsymbol{l}_i = z \wedge \boldsymbol{c}_i + \boldsymbol{l}}$

— this vanishes for $\boldsymbol{c}_i$ on $\mathbf{L}$

3D Planes in Plane + Parallax

- The 3D plane $(\boldsymbol{n}^\top, d)$ (*i.e.* $\boldsymbol{n} \cdot \boldsymbol{x} + d\,w = 0$) intersects the reference plane in the line $\boldsymbol{n} \cdot \boldsymbol{x} = 0$ and induces an image i to j homography

$$\boxed{\boldsymbol{H}_i^j = \boldsymbol{I} + \frac{\boldsymbol{e}_{ij}\boldsymbol{n}^\top}{\boldsymbol{n} \cdot \boldsymbol{c}_i + d}}$$

- For any i, j, this fixes each point on the intersection line $\boldsymbol{n} \cdot \boldsymbol{x} = 0$, and also the epipole $\boldsymbol{e}_{ij}$.

- The relative size of $\boldsymbol{n}$ and d give the 'steepness' of the 3D plane : $\boldsymbol{n} = \boldsymbol{0}$ for the reference plane, $d = 0$ for planes through the origin.

Relation to Plücker Coordinates

- In 3D coordinates, $\mathbf{L}$ is represented by the Plücker matrices

$$\boxed{\mathbf{L} = \begin{pmatrix} [\boldsymbol{l}]_\times & z \\ -z & 0 \end{pmatrix} \qquad \text{or dually} \qquad \mathbf{L}_* = \begin{pmatrix} [z]_\times & \boldsymbol{l} \\ -\boldsymbol{l} & 0 \end{pmatrix}}$$

- A 3D point $\mathbf{X} = \binom{\boldsymbol{x}}{w}$ lies on $\mathbf{L}$ iff

$$\boxed{\mathbf{L}_* \mathbf{X} = \begin{pmatrix} z \wedge \boldsymbol{x} + w\,\boldsymbol{l} \\ \boldsymbol{l} \cdot \boldsymbol{x} \end{pmatrix} = \mathbf{0}}$$

- The line through $\binom{\boldsymbol{x}}{w}$ and $\binom{\boldsymbol{y}}{v}$ is

$$\boxed{(\boldsymbol{l}, z) = (\boldsymbol{x} \wedge \boldsymbol{y}, w\boldsymbol{y} - v\boldsymbol{x})}$$

Carlsson's "Duality"

- As aligned cameras are parametrized by their optical centres, we can swap the roles of 3D points and cameras :

$$\boxed{\mathbf{X} = \begin{pmatrix} \boldsymbol{x} \\ w \end{pmatrix} \quad \longleftrightarrow \quad \mathbf{P} = (w\boldsymbol{I} \,|\, -\boldsymbol{x})}$$

— Points on the reference plane have singular dual projection matrices.

- The below matching tensor formulae can be dualized trivially — just exchange points $\boldsymbol{x}$ and projection centres $\boldsymbol{c}$

— *e.g.* the **dual epipole** of two 3D points $\mathbf{X} = \binom{\boldsymbol{x}}{w}$ and $\mathbf{Y} = \binom{\boldsymbol{y}}{v}$ is $v\,\boldsymbol{x} - w\,\boldsymbol{y}$, the intersection of the line $\mathbf{X}$-$\mathbf{Y}$ with the reference plane.

- In fact, the projection centres are just another type of 3D point

— points and projection centres can be mixed arbitrarily

— the "duality" is really an *equivalence*

Matching Tensors in Plane + Parallax

• The matching tensors for aligned projections are *very simple* functions of the scaled epipoles or the centres of projection :

$\mathbf{e}_{ij} = \mathbf{c}_i - \mathbf{c}_j$	epipole of c.o.p. i in image j
$\mathbf{F}_{ij} = [\,\mathbf{e}_{ij}\,]_\times = [\,\mathbf{c}_i - \mathbf{c}_j\,]_\times$	image i-j fundamental matrix
$\mathbf{T}_i^{\,jk} = \mathbf{I}_i^j \otimes \mathbf{e}_{ik} - \mathbf{e}_{ij} \otimes \mathbf{I}_i^k$	image i-j-k trifocal tensor
$\mathbf{Q}^{A_1A_2A_3A_4} = \sum_{i=1}^{3}(-1)^{i-1}\boldsymbol{\epsilon}^{A_1\ldots\hat{A}_i\ldots A_4}\cdot \mathbf{e}_{i4}^{A_i}$	1-2-3-4 quadrifocal tensor

• The intersection of the ray from $\mathbf{c}_j$ to $\mathbf{c}_i$ with the reference plane is

$$\begin{pmatrix}\mathbf{e}_{ij}\\0\end{pmatrix} = \begin{pmatrix}\mathbf{c}_i\\1\end{pmatrix} - \begin{pmatrix}\mathbf{c}_j\\1\end{pmatrix}$$

Tensor Brightness Constraints

• Replacing the lines with corresponding tangents to iso-intensity contours gives ***tensor brightness constraints*** on the dense normal flow

• The point-line-line case is the Hanna-Okamoto-Stein-Shashua brightness constraint — this is the main constraint for small, mostly-translational motions.

• For non-negligible rotation, the 3 line constraint gives additional information.

Matching Constraints in Plane + Parallax

• In P+P, the first few matching relations become :

$[\,\mathbf{x}_i, \mathbf{e}_{ij}, \mathbf{x}_j\,] = 0$	epipolar point
$(\mathbf{x}_i \wedge \mathbf{x}_j)\,(\mathbf{e}_{ik} \wedge \mathbf{x}_k)^\top - (\mathbf{e}_{ij} \wedge \mathbf{x}_j)\,(\mathbf{x}_i \wedge \mathbf{x}_k)^\top = \mathbf{0}$	trifocal point
$(\mathbf{l}_i \wedge \mathbf{l}_j)\,(\mathbf{l}_k \cdot \mathbf{e}_{ik}) - (\mathbf{l}_j \cdot \mathbf{e}_{ij})\,(\mathbf{l}_i \wedge \mathbf{l}_k) = \mathbf{0}$	trifocal line
$(\mathbf{l}_j \cdot \mathbf{x}_i)\,(\mathbf{l}_k \cdot \mathbf{e}_{ik}) - (\mathbf{l}_j \cdot \mathbf{e}_{ij})\,(\mathbf{l}_k \cdot \mathbf{x}_i) = 0$	trifocal point-line-line

• The trifocal point constraint contains two epipolar constraints, and also a proportionality-of-parallax relation.

Tensor Consistency in Plane + Parallax

• Modulo overall scaling, the number of linearly independent components of $\mathbf{F}, \mathbf{T}, \mathbf{Q}$ is reduced from $9, 27, 81$ to $3, 6, 9$.

• These are the $3m-3$ *d.o.f.* of the aligned camera geometry — individual tensors have *no* internal constraints and can be estimated linearly.

• P+P reduces the *inter*-tensor consistency constraints to coincidence, alignment and Desargues-like relations between epipoles derived from different tensors:

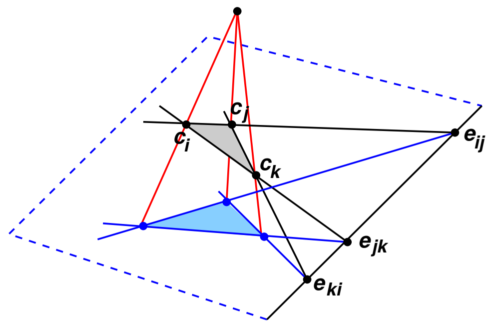
$$\begin{aligned}\mathbf{e}_{ji} \wedge \mathbf{e}_{ij} &= 0\\ \mathbf{e}_{ij} \wedge (\mathbf{e}_{ij})' &= 0\\ [\,\mathbf{e}_{ij}, \mathbf{e}_{jk}, \mathbf{e}_{ki}\,] &= 0\end{aligned}$$

Epipole Alignment and Desargues Theorem

• All circular triplets of epipoles are aligned: $[\boldsymbol{e}_{ij}, \boldsymbol{e}_{jk}, \boldsymbol{e}_{ki}] = 0$

— the epipoles lie on the intersection line of the reference plane with the plane defined by the three camera centres

— projection of this configuration into other images gives Desargues theorem configurations.



Closure based Reconstruction in P+P

• Given the epipole scalings, P+P reduces closure based projection matrix reconstruction to a trivial linear system

— just linearly read off $\boldsymbol{c}_i$'s consistent with the $\boldsymbol{e}_{ij}$'s.

• However, alignment-based epipolar scale recovery fails for :

1. Coincident epipoles / aligned camera centres

— but the trifocal tensor or matching relations give the missing scale factors here.

2. Untriangulated loops of more than 3 epipoles

— reconstruction is impossible without additional cross-links.

Coherent Scaling for P+P Tensors

• Depth recovery and closure based reconstruction requires coherent scalings for the tensors.

• P+P reduces all such relations to *the vanishing of cyclic sums of (correctly scaled) epipoles* :

$$\boldsymbol{e}_{ij} + \boldsymbol{e}_{jk} + \boldsymbol{e}_{kl} \ldots + \boldsymbol{e}_{mi} = \boldsymbol{0} \qquad \text{because} \qquad \boldsymbol{e}_{ij} = \boldsymbol{c}_i - \boldsymbol{c}_j$$

• Each aligned triplet of epipoles has a ***unique consistent relative scaling***:

$$\boldsymbol{e}_{ij} + \boldsymbol{e}_{jk} + \boldsymbol{e}_{ki} = \boldsymbol{0}$$

• If the $\boldsymbol{e}_{ij}$ scale factors are unknown, they can be recovered by assembling and solving these linear relations.

P+P Depth Recovery Relations

• For P+P registered points and lines with unknown relative scales, the first few depth recovery relations simplify to :

$$\begin{array}{rcll} \boldsymbol{e}_{ij} \wedge (\lambda_i \boldsymbol{x}_i - \lambda_j \boldsymbol{x}_j) & = & \boldsymbol{0} & \text{epipolar} \\ \boldsymbol{e}_{ij}\, (\lambda_k \boldsymbol{x}_k - \lambda_i \boldsymbol{x}_i)^\top - (\lambda_j \boldsymbol{x}_j - \lambda_i \boldsymbol{x}_i)\, (\boldsymbol{e}_{ik})^\top & = & \boldsymbol{0} & \text{trifocal} \\ (\mu_i \boldsymbol{l}_i - \mu_j \boldsymbol{l}_j) \cdot \boldsymbol{e}_{ij} & = & 0 & \text{line} \end{array}$$

• The trifocal point relations contain two epipolar ones, plus an additional relative scale constraint.

• In fact, with fully coherent scalings for the epipoles and points $\mathbf{X} = \binom{\boldsymbol{x}}{w}$:

$$\begin{array}{rcl} \lambda_i \boldsymbol{x}_i - \lambda_j \boldsymbol{x}_j & = & -w\, \boldsymbol{e}_{ij} \\ \mu_i \boldsymbol{l}_i - \mu_j \boldsymbol{l}_j & = & \boldsymbol{z} \wedge \boldsymbol{e}_{ij} \end{array} \quad \text{because} \quad \begin{array}{rcl} \lambda_i \boldsymbol{x}_i & = & \mathbf{P}_i \mathbf{X} = \boldsymbol{x} - w\, \boldsymbol{c}_i \\ \mu_i \boldsymbol{l}_i & = & \left([\boldsymbol{z}]_\times \,|\, \boldsymbol{l} \right) \binom{\boldsymbol{c}_i}{1} = \boldsymbol{z} \wedge \boldsymbol{c}_i + \boldsymbol{l} \end{array}$$

Depth / Projection Centre Factorization for P + P

- Given the correct scales, the $(\lambda_{ip}\boldsymbol{x}_{ip})$ matrix reduces to rank 4 as usual.
- Standard SFM factorization applies, but can be simplified as follows:

1 Averaging the projection $\boxed{\lambda_{ip}\boldsymbol{x}_{ip} \approx \mathbf{P}_i\mathbf{X}_p = \boldsymbol{x}_p - w_p\boldsymbol{c}_i}$ of a 3D point over all m images gives :

$$\begin{aligned} \bar{\boldsymbol{x}}_p &\equiv \tfrac{1}{m}\textstyle\sum_{i=1}^m \lambda_i \boldsymbol{x}_{ip} \approx \boldsymbol{x}_p - w_p\bar{\boldsymbol{c}} \\ \delta\boldsymbol{x}_{ip} &\equiv \lambda_{ip}\boldsymbol{x}_{ip} - \bar{\boldsymbol{x}}_p \approx -\delta\boldsymbol{c}_i\, w_p \end{aligned} \qquad \text{where} \qquad \begin{aligned} \bar{\boldsymbol{c}} &\equiv \tfrac{1}{m}\textstyle\sum_{i=1}^m \boldsymbol{c}_i \\ \delta\boldsymbol{c}_i &\equiv \boldsymbol{c}_i - \bar{\boldsymbol{c}} \end{aligned}$$

- $\bar{\boldsymbol{c}}$ is the same for all points and can be seen as a change of 3D origin.
- We implicitly fix the origin by taking $\bar{\boldsymbol{c}} = \boldsymbol{0}$, $\bar{\boldsymbol{x}}_p = \boldsymbol{x}_p$ and $\delta\boldsymbol{c}_i = \boldsymbol{c}_i$.

2 Rank 1 factorization of the combined difference matrix gives the projection centres $\delta\boldsymbol{c}_i$ and point depths w_p ***with their correct relative scales*** :

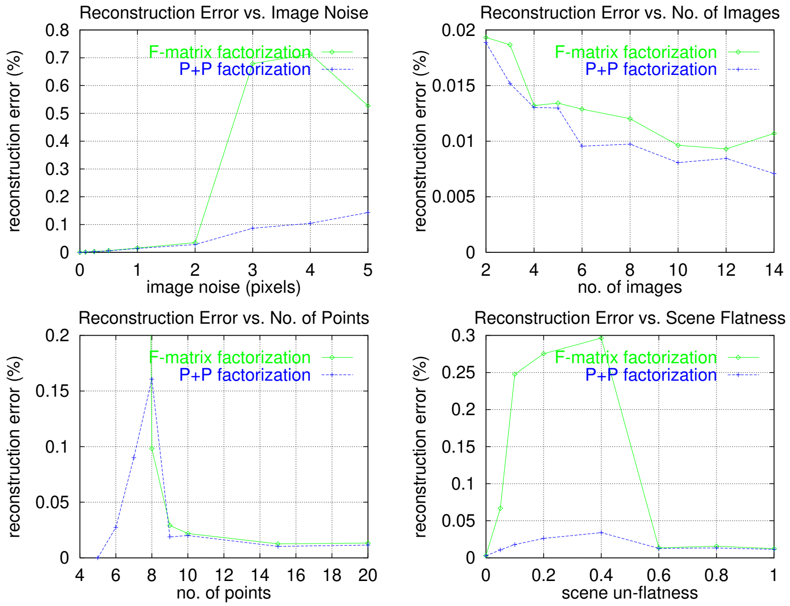
$$\begin{pmatrix} \delta\boldsymbol{x}_{11} & \dots & \delta\boldsymbol{x}_{1n} \\ \vdots & \ddots & \vdots \\ \delta\boldsymbol{x}_{m1} & \dots & \delta\boldsymbol{x}_{mn} \end{pmatrix} \approx \begin{pmatrix} -\delta\boldsymbol{c}_1 \\ \vdots \\ -\delta\boldsymbol{c}_m \end{pmatrix} \begin{pmatrix} w_1 & \dots & w_n \end{pmatrix}$$

3 The final reconstruction is :

$$\mathbf{P}_i = (\boldsymbol{I} \,|\, -\delta\boldsymbol{c}_i) \qquad \text{and} \qquad \mathbf{X}_p = \begin{pmatrix} \bar{\boldsymbol{x}}_p \\ w_p \end{pmatrix}$$

- This process requires (*i*) the initial P+P alignment and (*ii*) estimates of the epipoles for scale recovery.
- It returns the full 3D structure and camera centres in a heuristic 'average of camera centres' frame.

Preliminary Results



Line Factorizations

- One way to handle lines is :

1 Parametrize the line in one image by two via points

2 Find the corresponding points in the other images by epipolar or trifocal point transfer along the line.

3 Reconstruct the via points using, *e.g.* factorization.

- Transfer automatically gives the correct depths for the via points :

$$\begin{aligned} \boldsymbol{x}_i &= -\frac{\boldsymbol{l}_i \wedge (\boldsymbol{F}_{ij}\,\boldsymbol{x}_j)}{\boldsymbol{l}_i \cdot \boldsymbol{e}_{ji}} && \text{general case} \\ \boldsymbol{x}_j &\equiv \boldsymbol{x}_i + \frac{\boldsymbol{l}_j \cdot \boldsymbol{x}_i}{\boldsymbol{l}_j \cdot \boldsymbol{e}_{ij}}\,\boldsymbol{e}_{ij} && \text{P+P case} \end{aligned}$$

- In the P+P case, all of the image lines $\boldsymbol{z} \wedge \boldsymbol{c}_i + \boldsymbol{l}$ intersect in a common point $\boldsymbol{z}$, so we only need one other via point

— *e.g.* choose $\boldsymbol{x}_i \simeq \boldsymbol{l}_i \wedge \boldsymbol{z}$ in some image.

Another Line Factorization

- Correctly rescaled image lines can be factorized directly, using the line projection $\boldsymbol{l}_{il} = \left([\, z_l \,]_\times \,|\, \boldsymbol{l}_l \right) \binom{\boldsymbol{c}_i}{1}$
- The line parameters behave like 'projections', the centres of projection like 'points'.
- This enforces neither the rank-2-ness of $\left([\, z \,]_\times \,|\, \boldsymbol{l} \right)$, nor the skewness of its leading submatrix $[\, z \,]_\times$.
- Rank-2-ness can be enforced by estimating the intersection point z_l of the images $\boldsymbol{l}_{il}$ of line l, choosing 2 vectors $\boldsymbol{u}_l$ and $\boldsymbol{v}_l$ orthogonal to z_l, projecting the equations along these, factorizing, then un-projecting.
- Skewness can only be enforced by a Tomasi-Kanade-like post-factorization step.

Summary

- A coherent picture of multi-image projective geometry in terms of matching tensors and projective depths.
- There's more to projective scale factors than meets the eye.
- Plane + parallax greatly simplifies matching tensor geometry and reconstruction.
- P+P factorization looks promising.