Affine Factorization

- Reconstructs 3D points and camera matrices, assuming affine cameras and that the same points are imaged in all cameras.

- This is the maximum likelihood estimation if the measured image points have isotropic mean-zero Gaussian noise that is independent and equal for all measurements.

Inhomogenous Affine Camera

$$x_i^c = M^c X_i + t^c$$

x_i^c inhomogeneous 2D image position of point i in camera c

X_i inhomogeneous 3D world position of point i

M^c 2x3 matrix for camera c

t^c 2D image translation for camera c

$\langle X \rangle = \frac{1}{N}\sum_{i=1}^{N} X_i$ mean 3D position of the points

$\langle x^c \rangle = \frac{1}{N}\sum_{i=1}^{N} x_i^c$ mean image position of the points in camera c

Choose coordinate systems in images and 3D world such such that :

$$\langle x^c \rangle = 0 \qquad \langle X \rangle = 0$$

For an affine camera $\langle X \rangle$ projects to $\langle x^c \rangle$:

$$M^c \langle X \rangle + t^c = M^c \frac{1}{N} \sum_{i=1}^{N} X_i + t^c = \frac{1}{N} \sum_{i=1}^{N} \left(M^c X_i + t^c \right) = \frac{1}{N} \sum_{i=1}^{N} x_i^c = \langle x^c \rangle$$

$$\Rightarrow t^c = 0$$

$$\Rightarrow x_i^c = M^c X_i$$

Do 3D reconstruction by minimizing reprojection error :

$$\min_{M^c, X_i} \sum_{i,c} \left\| x_i^c - M^c X_i \right\|^2$$

Matrix formulation

Minimize reprojection error: $$\min_{M^c, X_i} \sum_{i,c} \left\| x_i^c - M^c X_i \right\|^2$$

$$W = \begin{pmatrix} x_1^1 & x_2^1 & \cdots & x_N^1 \\ x_1^2 & x_2^2 & & x_N^2 \\ \vdots & & \ddots & \vdots \\ x_1^C & x_2^C & \dots & x_N^C \end{pmatrix} \qquad \hat{W} = \begin{pmatrix} M^1 X_1 & M^1 X_2 & \cdots & M^1 X_N \\ M^2 X_1 & M^2 X_2 & & M^2 X_N \\ \vdots & & \ddots & \vdots \\ M^C X_1 & M^C X_2 & \dots & M^C X_N \end{pmatrix} = \underbrace{\begin{pmatrix} M^1 \\ M^2 \\ \vdots \\ M^C \end{pmatrix}}_{\hat{M} \quad 2\text{C}\times 3} \underbrace{\begin{pmatrix} X_1 & X_2 & \cdots & X_N \end{pmatrix}}_{\hat{X} \quad 3\times\text{N}}$$

$\hat{\text{W}}$ should have rank 3 since $\hat{\text{M}}$ and $\hat{\text{X}}$ have rank 3

$$\sum_{i,c} \left\| x_i^c - M^c X_i \right\|^2 = \left\| W - \hat{W} \right\|_F^2 \qquad \min_{M^c, X_i} \sum_{i,c} \left\| x_i^c - M^c X_i \right\|^2 = \min_{\hat{W}} \left\| W - \hat{W} \right\|_F^2 \quad \text{s.t.} \quad rank(\hat{W}) = 3$$

Minimize reprojection error: $\min_{\hat{W}} \left\| W - \hat{W} \right\|_F^2$

s.t. $rank(\hat{W}) = 3$

The $\hat{W}$ that is closest to W while having a rank of three can be computed by doing singular value decomposition of W and taking the three biggest singular values:

$$W = UDV^T \qquad \hat{W} = U_{2C\times 3} D_{3\times 3} V_{3\times N}^T = \hat{M}\hat{X} = \begin{pmatrix} M^1 \\ M^2 \\ \vdots \\ M^C \end{pmatrix} \begin{pmatrix} X_1 & X_2 & \cdots & X_N \end{pmatrix}$$

$$\hat{M} = U_{2C\times 3} D_{3\times 3}$$

$$\hat{X} = V_{3\times N}^T$$

Affine ambiguity of the reconstruction:

$$\hat{M} = U_{2C\times 3} D_{3\times 3} A$$

$$\hat{X} = A^{-1} V_{3\times N}^T$$

Points not visible in all cameras

Assume we have performed 3D reconstruction for the points visible in all cameras. As a result we have reconstructed the camera matrices Mc. Suppose we have image measurements of an additional point in some of the cameras. We can then reconstruct the corresponding 3D point X using the computed camera matrices. Each measurement from a camera give the linear equation:

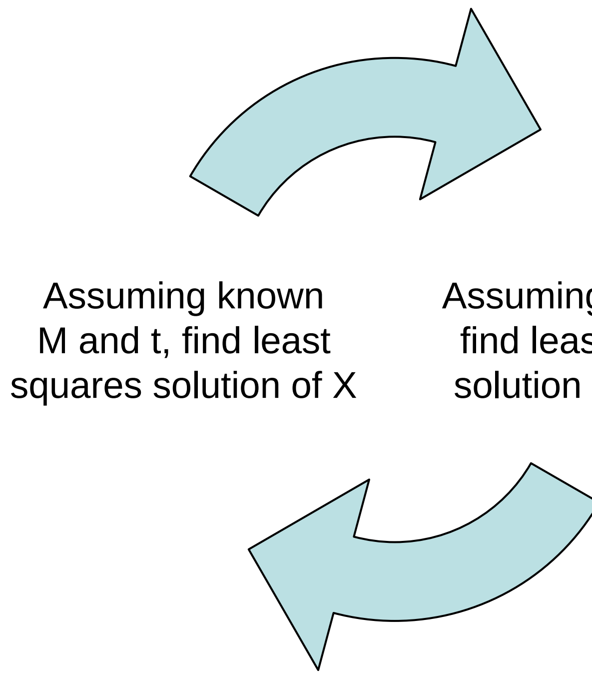
$$x^c = M^c X$$

Given sufficiently many measurements from different cameras we can find the least squares solution to X. We can also reproject X to get its image in all cameras.

$$\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^c \end{pmatrix} = \hat{M}X = \begin{pmatrix} M^1 \\ M^2 \\ \vdots \\ M^C \end{pmatrix} X$$

Affine Reconstruction by Alternation

$$x_i^c = M^c X_i + t^c$$



Can handle points not visible in all cameras and measurements being weighted differently, but in those cases global optimal convergence is not always guaranteed.

Projective Factorization

$$\lambda_i^c x_i^c = P^c X_i$$

$x_i^c = \begin{pmatrix} u_i^c \\ v_i^c \\ 1 \end{pmatrix}$ image coordinates of point i in camera c

P^c 3x4 projection matrix of camera c

X_i homogeneous world position of point i

$\lambda_\text{i}^\text{c}$ projective depth of point i in camera c

Assuming known λ_i^c and x_i^c, do reconstruction by minimizing "reprojection error":

$$\min_{P^c, X_i} \sum_{i,c} \left\| \lambda_i^c x_i^c - P^c X_i \right\|^2$$

Matrix formulation

Minimize reprojection error: $\min_{P^c, X_i} \sum_{i,c} \left\| \lambda_i^c x_i^c - P^c X_i \right\|^2$

$$W = \begin{pmatrix} \lambda_1^1 x_1^1 & \lambda_2^1 x_2^1 & \cdots & \lambda_N^1 x_N^1 \\ \lambda_1^2 x_1^2 & \lambda_2^2 x_2^2 & & \lambda_N^2 x_N^2 \\ \vdots & & \ddots & \vdots \\ \lambda_1^C x_1^C & \lambda_2^C x_2^C & \ldots & \lambda_N^C x_N^C \end{pmatrix} \qquad \hat{W} = \begin{pmatrix} P^1 X_1 & P^1 X_2 & \cdots & P^1 X_N \\ P^2 X_1 & P^2 X_2 & & P^2 X_N \\ \vdots & & \ddots & \vdots \\ P^C X_1 & P^C X_2 & \ldots & P^C X_N \end{pmatrix} = \underbrace{\begin{pmatrix} P^1 \\ P^2 \\ \vdots \\ P^C \end{pmatrix}}_{\hat{P} \quad 3\text{C}\times 4} \underbrace{\begin{pmatrix} X_1 & X_2 & \cdots & X_N \end{pmatrix}}_{\hat{X} \quad 4\times\text{N}}$$

$\hat{\text{W}}$ should have rank 4 since $\hat{\text{P}}$ and $\hat{\text{X}}$ have rank 4

$$\sum_{i,c} \left\| x_i^c - P^c X_i \right\|^2 = \left\| W - \hat{W} \right\|_F^2 \qquad \min_{P^c, X_i} \sum_{i,c} \left\| \lambda_i^c x_i^c - P^c X_i \right\|^2 = \min_{\hat{W}} \left\| W - \hat{W} \right\|_F^2 \quad \text{s.t.} \quad rank(\hat{W}) = 4$$

Minimize reprojection error: $\min_{\hat{W}} \left\| W - \hat{W} \right\|_F^2$

s.t. $rank(\hat{W}) = 4$

The $\hat{W}$ that is closest to W while having a rank of four can be computed by doing singular value decomposition of W and taking the four biggest singular values:

$$W = UDV^T \qquad \hat{W} = U_{3C\times 4} D_{4\times 4} V_{4\times N}^T = \hat{P}\hat{X} = \begin{pmatrix} P^1 \\ P^2 \\ \vdots \\ P^C \end{pmatrix} \begin{pmatrix} X_1 & X_2 & \cdots & X_N \end{pmatrix}$$

$$\hat{P} = U_{3C\times 4} D_{4\times 4}$$

$$\hat{X} = V_{4\times N}^T$$

Projective ambiguity of the reconstruction:

$$\hat{M} = U_{3C\times 4} D_{4\times 4} A$$

$$\hat{X} = A^{-1} V_{4\times N}^T$$

Issues

Need values for λ_i^c. Can set $\lambda_i^c = 1$ as initial guess,
or use other reconstruction methods as a first step.

Due to λ we are not minimizing geometric error.
We can reduce the effect of this by two normalization procedures:

1. Normalize image coordinates.

2. A heuristic to further improve the result is to take the measurement matrix W and normalize each row followed by a normalization of each column. This procedure may be iterated.