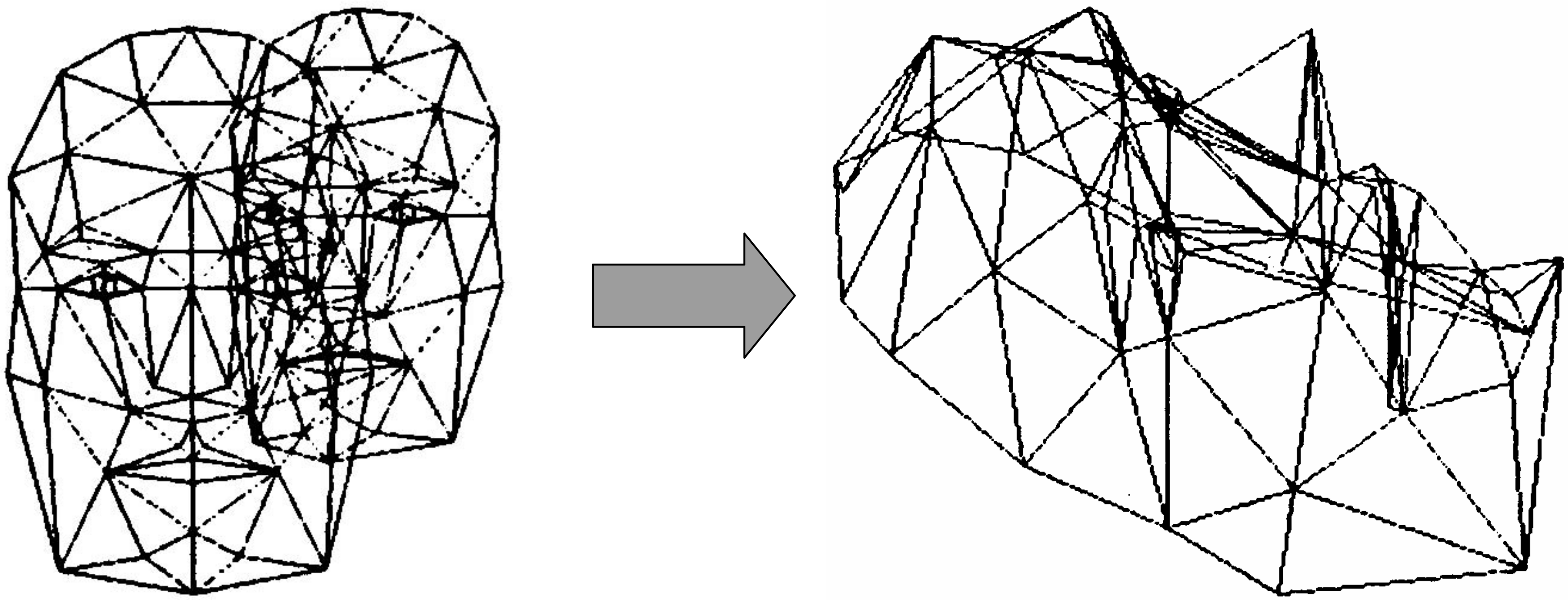
COS429: COMPUTER VISON
AFFINE STRUCTURE FROM MOTION

The Structure-from-Motion Problem
Affine Projection Models
Affine Ambiguity of Affine SFM
Affine Epipolar Geometry
Affine Reconstruction from two Images
Affine Reconstruction from Multiple Images

- Reading: Chapter 12

Many of the slides in this lecture are courtesy to Prof. J. Ponce

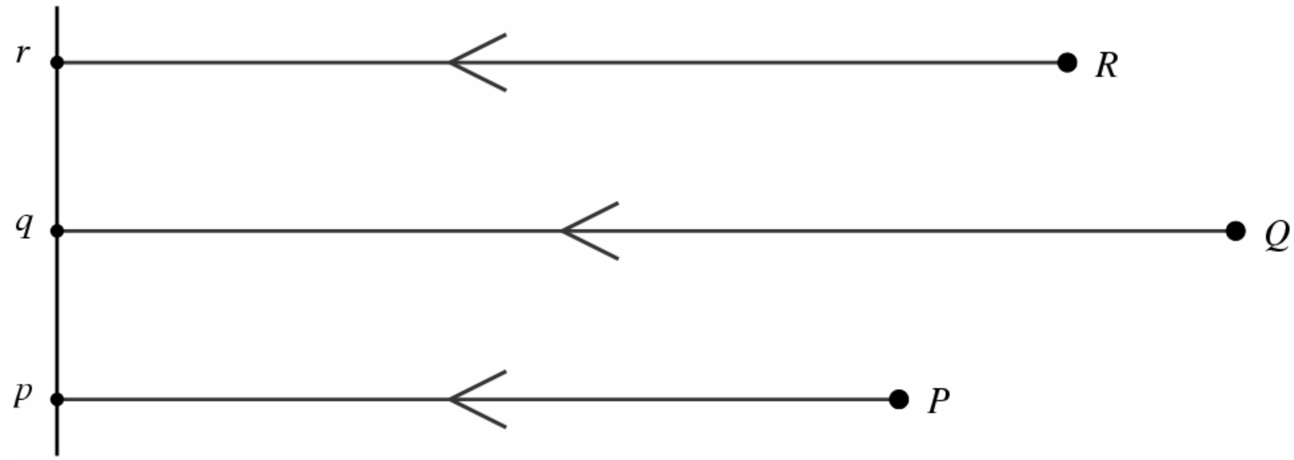
Affine Structure from Motion



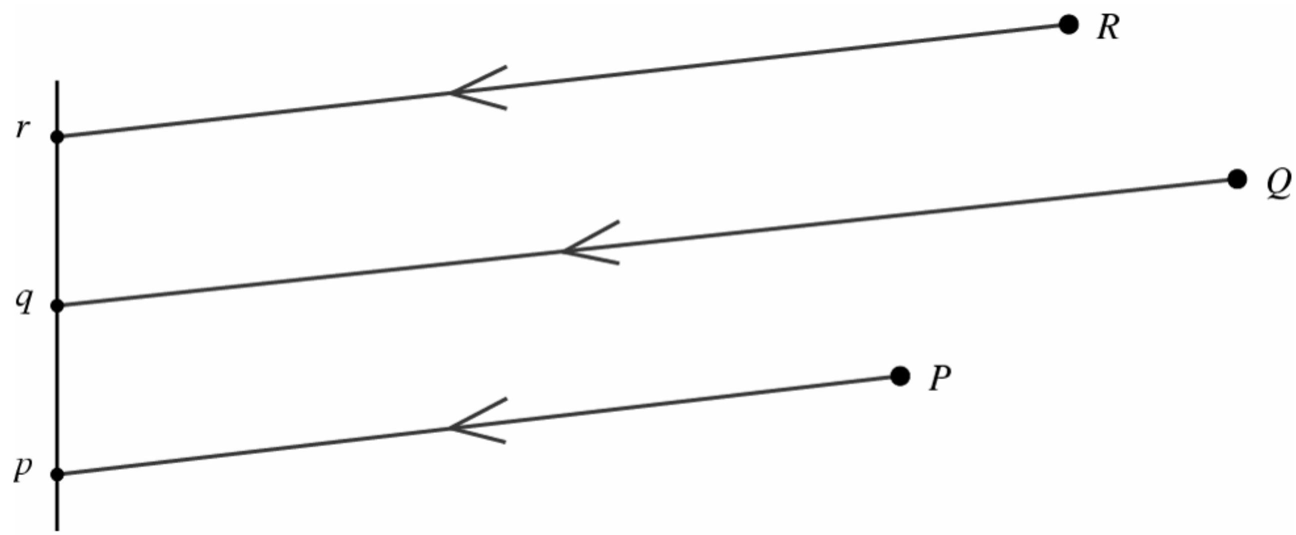
Reprinted with permission from “Affine Structure from Motion,” by J.J. (Koenderink and A.J.Van Doorn, Journal of the Optical Society of America A, 8:377-385 (1990). © 1990 Optical Society of America.

Given *m* pictures of *n* points, can we recover
- the three-dimensional configuration of these points? (structure)
- the camera configurations? (motion)

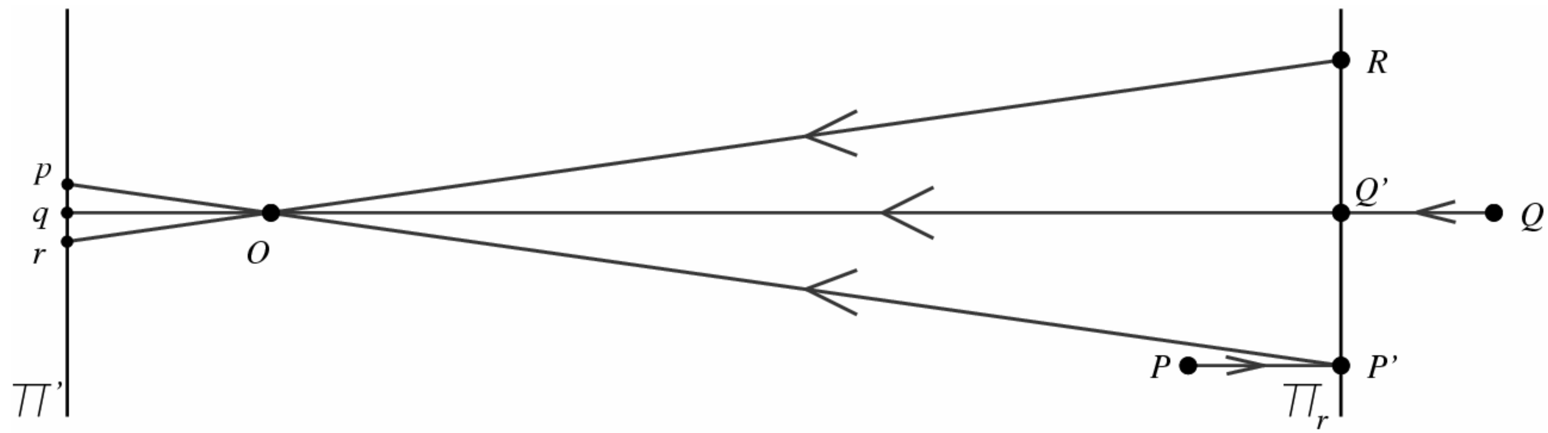
Orthographic Projection



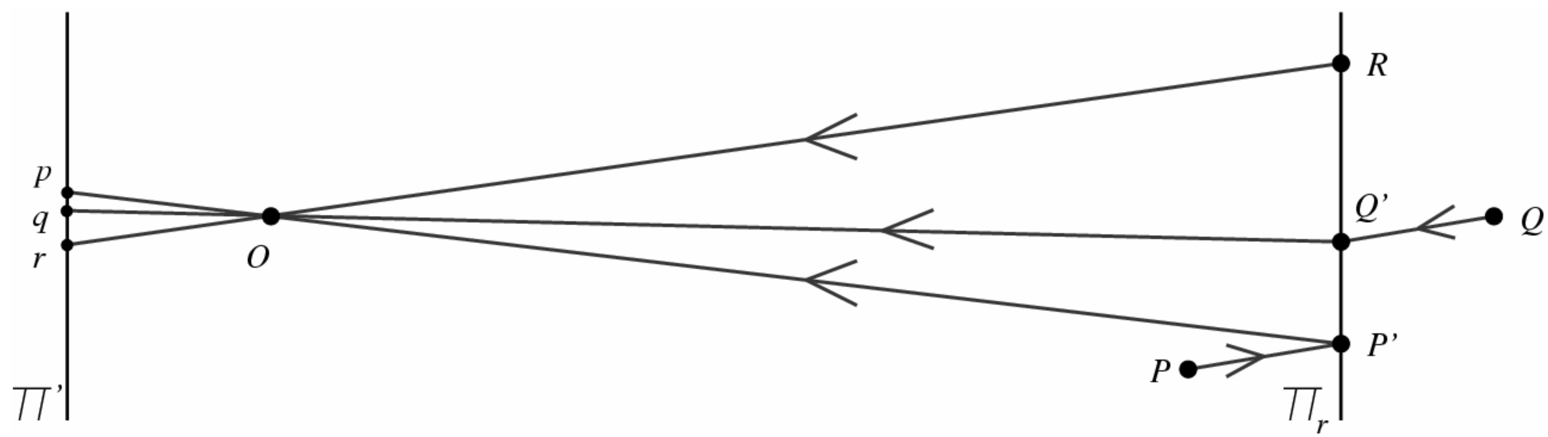
Parallel Projection



Weak-Perspective Projection



Paraperspective Projection



The Affine Structure-from-Motion Problem

Given m images of n fixed points P_j we can write

$$\boldsymbol{p}_{ij} = \mathcal{M}_i \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix} = \mathcal{A}_i \boldsymbol{P}_j + \boldsymbol{b}_i \quad \text{for} \quad i = 1, \ldots, m \quad \text{and} \quad j = 1, \ldots, n.$$

Problem: estimate the m 2x4 matrices $\mathcal{M}_i$ and
the n positions P_j from the mn correspondences $\boldsymbol{p}_{ij}$.

$2mn$ equations in $8m+3n$ unknowns

Overconstrained problem, that can be solved
using (non-linear) least squares!

The Affine Epipolar Constraint

$$\begin{cases} \boldsymbol{p} = \mathcal{A}\boldsymbol{P} + \boldsymbol{b} \\ \boldsymbol{p}' = \mathcal{A}'\boldsymbol{P} + \boldsymbol{b}' \end{cases} \Rightarrow \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} \begin{pmatrix} \boldsymbol{P} \\ -1 \end{pmatrix} = \boldsymbol{0}$$

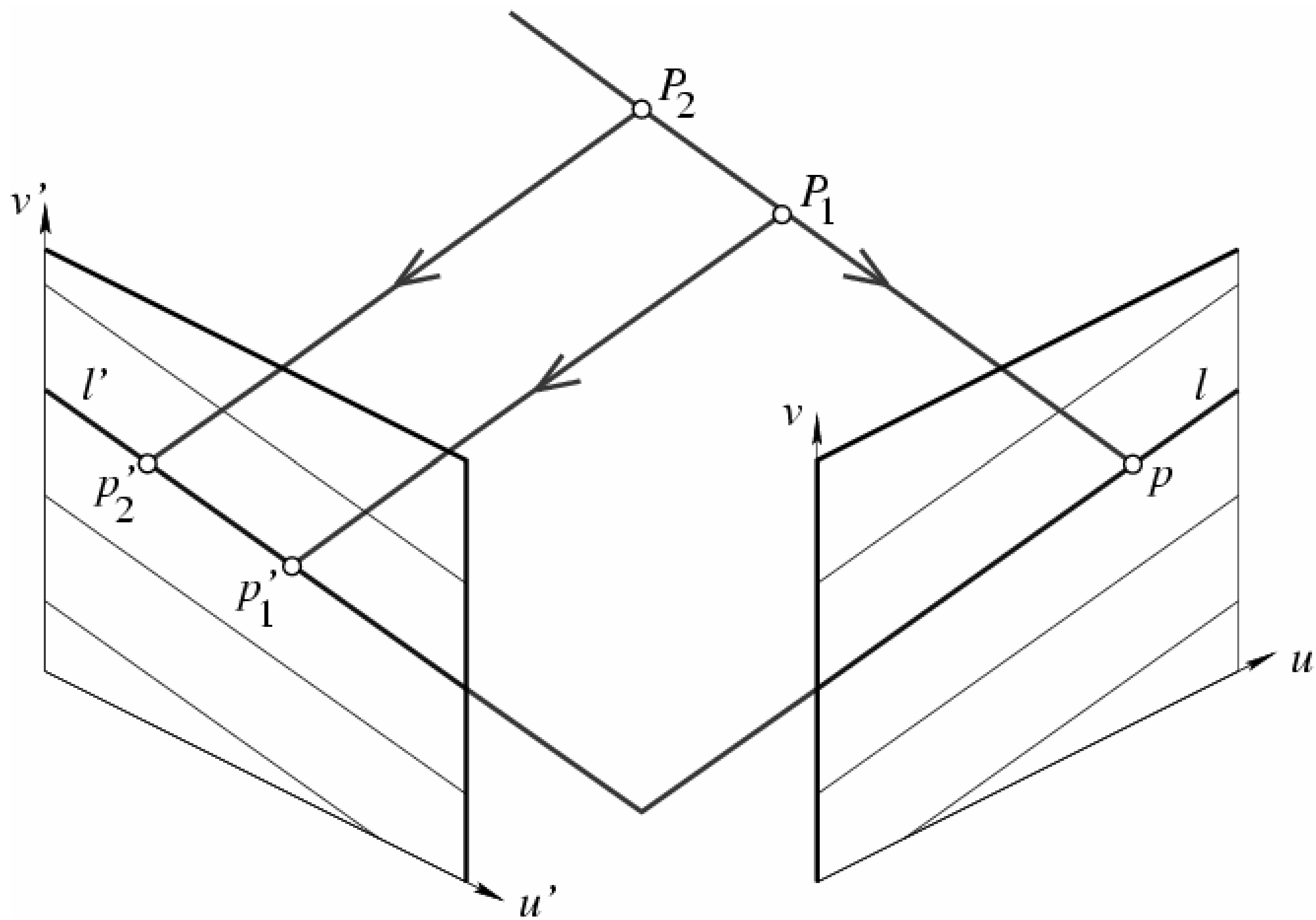
$$\Downarrow$$

$$\operatorname{Det} \begin{pmatrix} \mathcal{A} & \boldsymbol{p} - \boldsymbol{b} \\ \mathcal{A}' & \boldsymbol{p}' - \boldsymbol{b}' \end{pmatrix} = 0$$

$$\Downarrow$$

$$\boxed{\alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0}$$

Affine Epipolar Geometry



Note: the epipolar lines are parallel.

The Affine Fundamental Matrix

$$\alpha u + \beta v + \alpha' u' + \beta' v' + \delta = 0$$

$$\Downarrow$$

$$(u, v, 1)\mathcal{F}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \qquad \text{where} \qquad \mathcal{F} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & \alpha \\ 0 & 0 & \beta \\ \alpha' & \beta' & \delta \end{pmatrix}$$

The Affine Ambiguity of Affine SFM

When the intrinsic and extrinsic parameters are unknown

If M_i and P_j are solutions,

$$\boldsymbol{p}_{ij} = \mathcal{M}_i \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix} = (\mathcal{M}_i \mathcal{Q}) \left(\mathcal{Q}^{-1} \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix}\right) = \mathcal{M}'_i \begin{pmatrix} \boldsymbol{P}'_j \\ 1 \end{pmatrix}$$

So are M'_i and P'_j where

$$\mathcal{M}'_i = \mathcal{M}_i \mathcal{Q} \quad \text{and} \quad \begin{pmatrix} \boldsymbol{P}'_j \\ 1 \end{pmatrix} = \mathcal{Q}^{-1} \begin{pmatrix} \boldsymbol{P}_j \\ 1 \end{pmatrix}$$

and

$$\mathcal{Q} = \begin{pmatrix} \mathcal{C} & \boldsymbol{d} \\ \boldsymbol{0}^T & 1 \end{pmatrix} \quad \text{with} \quad \mathcal{Q}^{-1} = \begin{pmatrix} \mathcal{C}^{-1} & -\mathcal{C}^{-1}\boldsymbol{d} \\ \boldsymbol{0}^T & 1 \end{pmatrix}$$

Q is an affine transformation.

An Affine Trick.. Algebraic Scene Reconstruction Method

$$\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b}) \qquad \mathcal{M}' = (\mathcal{A}' \quad \boldsymbol{b}') \qquad \boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \qquad \tilde{\boldsymbol{P}}$$

$$\boxed{\mathrm{Det}\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v'-d \end{pmatrix} = au - bv + cu' + v' - d = 0}$$

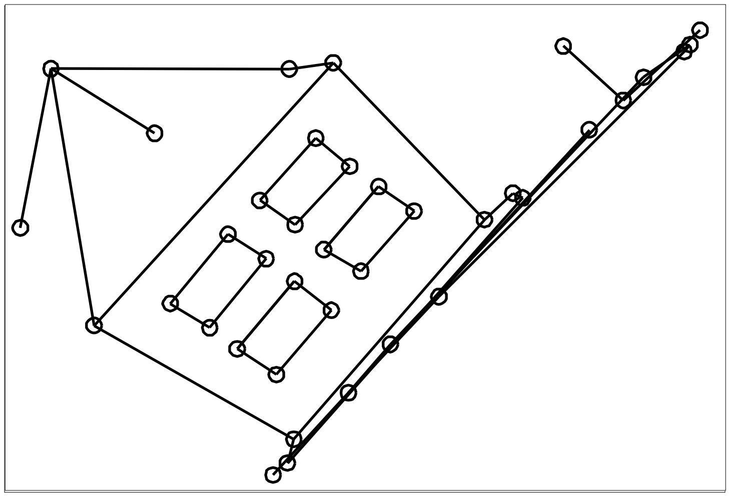
An Affine Trick.. $\Rightarrow$ Algebraic Scene Reconstruction Method

$$\mathcal{M} = \begin{pmatrix} \mathcal{A} & \boldsymbol{b} \end{pmatrix} \qquad \mathcal{M}' = \begin{pmatrix} \mathcal{A}' & \boldsymbol{b}' \end{pmatrix} \qquad \boldsymbol{P}$$

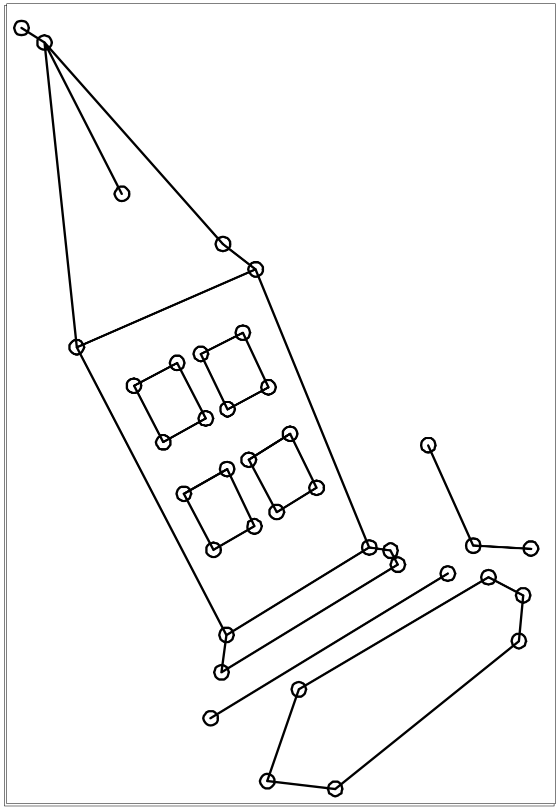
$$\tilde{\mathcal{M}} = \mathcal{M}\mathcal{Q} \qquad \tilde{\mathcal{M}}' = \mathcal{M}'\mathcal{Q} \qquad \tilde{\boldsymbol{P}} = \mathcal{Q}^{-1}\boldsymbol{P}$$

$$\tilde{\mathcal{M}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \tilde{\mathcal{M}}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix} \quad \tilde{\boldsymbol{P}}$$

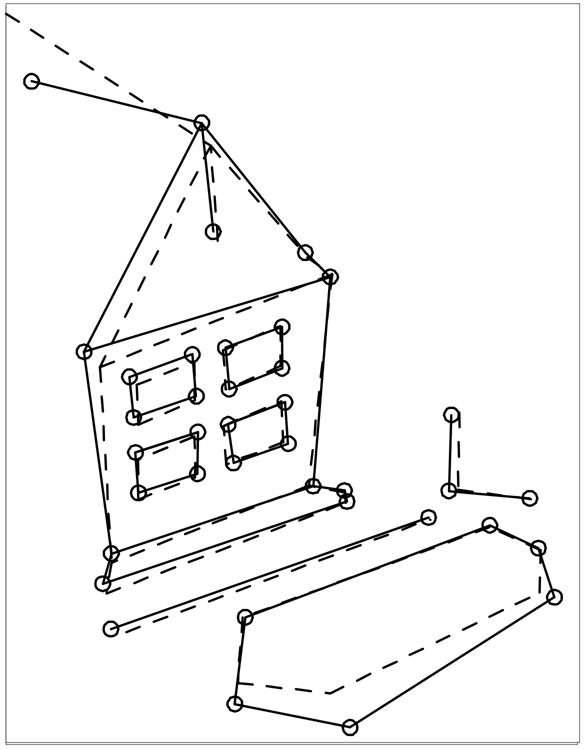
$$\begin{pmatrix} 1 & 0 & 0 & u \\ 0 & 1 & 0 & v \\ 0 & 0 & 1 & u' \\ a & b & c & v' - d \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{P}} \\ -1 \end{pmatrix} = 0 \quad \Rightarrow \quad \boxed{\tilde{\boldsymbol{P}} = \begin{pmatrix} u \\ v \\ u' \end{pmatrix}}$$



First reconsruction. Mean reprojection error: 1.6pixel



Second reconsruction. Mean reprojection error: 7.8pixel



Suppose we observe a scene with *m* fixed cameras..

$$\boldsymbol{p}_i = \mathcal{M}_i \begin{pmatrix} \boldsymbol{P} \\ 1 \end{pmatrix} = \mathcal{A}_i \boldsymbol{P} + \boldsymbol{b}_i \quad \text{for} \quad i = 1, \ldots, m$$

$$\begin{aligned} \boldsymbol{P} &\longrightarrow \boldsymbol{P} - \boldsymbol{P}_0 \\ \boldsymbol{p} &\longrightarrow \boldsymbol{p} - \boldsymbol{p}_0 \end{aligned} \quad \Longrightarrow \quad \boldsymbol{p}_i = \mathcal{A}_i \boldsymbol{P} + \boldsymbol{b}_i \longrightarrow \boldsymbol{p}_i = \mathcal{A}_i \boldsymbol{P}$$

$$\boxed{\begin{matrix} u_{11}\ u_{12} \ldots u_{1n} \\ v_{11}\ v_{12} \ldots v_{1n} \\ \ldots\ldots\ldots \\ u_{m1}\ u_{m2} \ldots u_{mn} \\ v_{m1}\ v_{m2} \ldots v_{mn} \end{matrix}} = \boxed{\begin{matrix} A_1 \\ A_2 \\ \ldots \\ A_m \end{matrix}} \boxed{P_1\ P_2 \ldots P_n} = \mathcal{D}$$

What if we could factorize D? (Tomasi and Kanade, 1992)

$$\mathcal{A}, \mathcal{P} \rightarrow \mathcal{D}$$

Affine SFM is solved!

$$\mathcal{D} \rightarrow \mathcal{A}, \mathcal{P}$$

Singular Value Decomposition

Let $\mathcal{A}$ be an $m \times n$ matrix, with $m \geq n$, then $\mathcal{A}$ can always be written as

$$\mathcal{A} = \mathcal{U}\mathcal{W}\mathcal{V}^T,$$

where:

- $\mathcal{U}$ is an $m \times n$ column-orthogonal matrix, i.e., $\mathcal{U}^T\mathcal{U} = \mathrm{Id}_m$,
- $\mathcal{W}$ is a diagonal matrix whose diagonal entries w_i $(i = 1, \ldots, n)$ are the singular values of $\mathcal{A}$ with $w_1 \geq w_2 \geq \ldots \geq w_n \geq 0$,
- and $\mathcal{V}$ is an $n \times n$ orthogonal matrix, i.e., $\mathcal{V}^T\mathcal{V} = \mathcal{V}\mathcal{V}^T = \mathrm{Id}_n$.

What if we could factorize D? (Tomasi and Kanade, 1992)

$$\mathcal{A}, \mathcal{P} \rightarrow \mathcal{D}$$

Affine SFM is solved!

$$\mathcal{D} \rightarrow \mathcal{A}, \mathcal{P}$$

Singular Value Decomposition

Theorem: The singular values of the matrix $\mathcal{A}$ are the eigenvalues of the matrix $\mathcal{A}^T\mathcal{A}$ and the columns of the matrix $\mathcal{V}$ are the corresponding eigenvectors.

What if we could factorize D? (Tomasi and Kanade, 1992)

$\mathcal{A}, \mathcal{P} \to \mathcal{D}$

Affine SFM is solved!

$\mathcal{D} \to \mathcal{A}, \mathcal{P}$

Singular Value Decomposition

When $\mathcal{A}$ has rank $p < n$, then the matrices $\mathcal{U}$, $\mathcal{W}$, and $\mathcal{V}$ can be written as

$$\mathcal{U} = \begin{array}{|c|c|} \hline \mathcal{U}_p & \mathcal{U}_{n-p} \\ \hline \end{array} \quad \mathcal{W} = \begin{array}{|c|c|} \hline \mathcal{W}_p & 0 \\ \hline 0 & 0 \\ \hline \end{array} \quad \text{and} \quad \mathcal{V}^T = \begin{array}{|c|} \hline \mathcal{V}_p^T \\ \hline \mathcal{V}_{n-p}^T \\ \hline \end{array},$$

and

- the columns of $\mathcal{U}_p$ form an orthonormal basis of the space spanned by the columns of $\mathcal{A}$, i.e., its *range*,
- and the columns of $\mathcal{V}_{n-p}$ for a basis of the space spanned by the solutions of $A\boldsymbol{x} = 0$, i.e., the *null space* of this matrix.

In addition, $\mathcal{A} = \mathcal{U}_p \mathcal{W}_p \mathcal{V}_p^T$.

What if we could factorize D? (Tomasi and Kanade, 1992)

$$\mathcal{A}, \mathcal{P} \rightarrow \mathcal{D}$$

Affine SFM is solved!

$$\mathcal{D} \rightarrow \mathcal{A}, \mathcal{P}$$

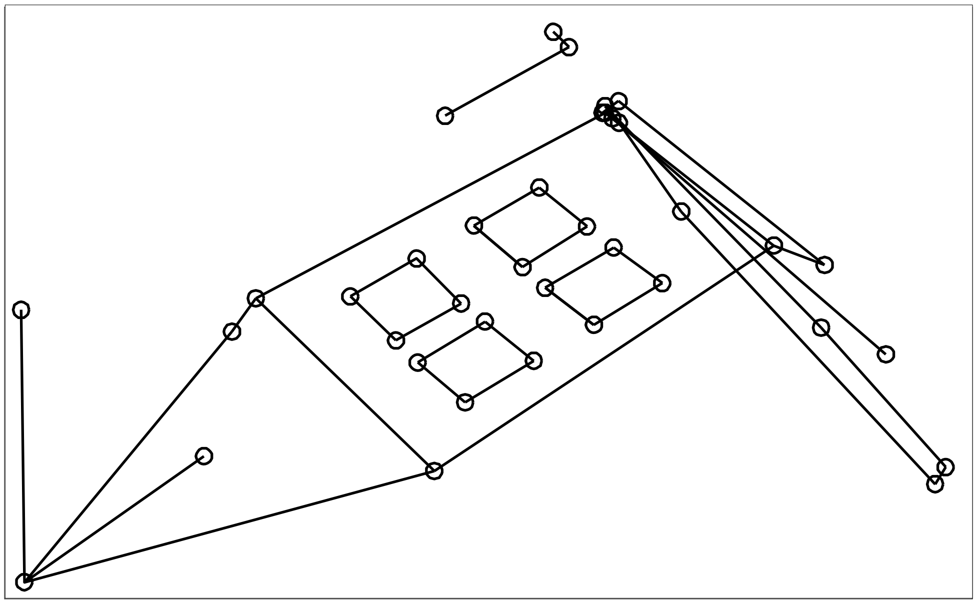
$$E \stackrel{\text{def}}{=} \sum_{i,j} |\boldsymbol{p}_{ij} - \mathcal{A}_i \boldsymbol{P}_j|^2 = \sum_j |\boldsymbol{q}_j - \mathcal{A}\boldsymbol{P}_j|^2 = |\mathcal{D} - \mathcal{A}\mathcal{P}|^2$$

Singular Value Decomposition

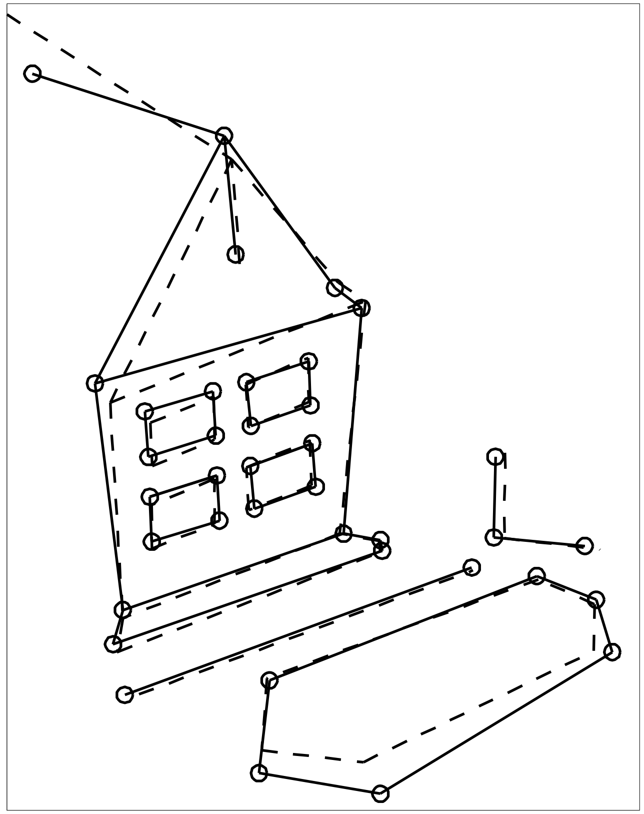
Theorem: When $\mathcal{A}$ has a rank greater than p, $\mathcal{U}_p \mathcal{W}_p \mathcal{V}_p^T$ is the best possible rank-p approximation of $\mathcal{A}$ in the sense of the Frobenius norm.

$$\mathcal{D} = \mathcal{U}_3 \mathcal{W}_3 \mathcal{V}_3^T$$

$$\boxed{\begin{cases} \mathcal{A}_0 = \mathcal{U}_3 \\ \mathcal{P}_0 = \mathcal{W}_3 \mathcal{V}_3^T \end{cases}}$$



Mean reprojection error:
2.4pixel



From uncalibrated to calibrated cameras

Weak-perspective camera:

$$\mathcal{M} = \frac{1}{z_r}\begin{pmatrix} k & s \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathcal{R}_2 & \boldsymbol{t}_2 \end{pmatrix}$$

$$\hat{\mathcal{M}} = \mathcal{M}\mathcal{Q}$$

Calibrated camera:

$$\hat{\mathcal{M}} = \begin{pmatrix} \hat{\mathcal{A}} & \hat{\boldsymbol{b}} \end{pmatrix} = \frac{1}{z_r}\begin{pmatrix} \mathcal{R}_2 & \boldsymbol{t}_2 \end{pmatrix}$$

Problem: what is *Q* ?

$$\hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{a}}_2 = 0 \quad \text{and} \quad |\hat{\boldsymbol{a}}_1|^2 = |\hat{\boldsymbol{a}}_2|^2$$

$$\mathcal{Q} = \begin{pmatrix} \mathcal{C} & \boldsymbol{d} \\ \boldsymbol{0}^T & 1 \end{pmatrix} \Rightarrow \begin{cases} \boldsymbol{a}_{i1}^T \mathcal{C}\mathcal{C}^T \boldsymbol{a}_{i2} = 0, \\ \boldsymbol{a}_{i1}^T \mathcal{C}\mathcal{C}^T \boldsymbol{a}_{i1} = 1, \\ \boldsymbol{a}_{i2}^T \mathcal{C}\mathcal{C}^T \boldsymbol{a}_{i2} = 1, \end{cases} \quad \text{for} \quad i = 1, \ldots, m,$$

From uncalibrated to calibrated cameras

Weak-perspective camera:

$$\mathcal{M} = \frac{1}{z_r}\begin{pmatrix} k & s \\ 0 & 1 \end{pmatrix}(\,\mathcal{R}_2 \quad \boldsymbol{t}_2\,)$$

$$\hat{\mathcal{M}} = \mathcal{M}\mathcal{Q}$$

Calibrated camera:

$$\hat{\mathcal{M}} = (\,\hat{\mathcal{A}} \quad \hat{\boldsymbol{b}}\,) = \frac{1}{z_r}(\,\mathcal{R}_2 \quad \boldsymbol{t}_2\,)$$

Problem: what is *Q* ?

$$\hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{a}}_2 = 0 \quad \text{and} \quad |\hat{\boldsymbol{a}}_1|^2 = |\hat{\boldsymbol{a}}_2|^2$$

$$\mathcal{D} = \mathcal{C}\mathcal{C}^T \Rightarrow \begin{cases} \boldsymbol{a}_{i1}^T \mathcal{D} \boldsymbol{a}_{i2} = 0, \\ \boldsymbol{a}_{i1}^T \mathcal{D} \boldsymbol{a}_{i1} = \boldsymbol{a}_{i2}^T \mathcal{D} \boldsymbol{a}_{i2} \end{cases} \quad \text{for} \quad i = 1, \ldots, m.$$

From uncalibrated to calibrated cameras

Weak-perspective camera: $$\mathcal{M} = \frac{1}{z_r}\begin{pmatrix} k & s \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathcal{R}_2 & \boldsymbol{t}_2 \end{pmatrix}$$

$$\hat{\mathcal{M}} = \mathcal{M}\mathcal{Q}$$

Calibrated camera: $$\hat{\mathcal{M}} = \begin{pmatrix} \hat{\mathcal{A}} & \hat{\boldsymbol{b}} \end{pmatrix} = \frac{1}{z_r}\begin{pmatrix} \mathcal{R}_2 & \boldsymbol{t}_2 \end{pmatrix}$$

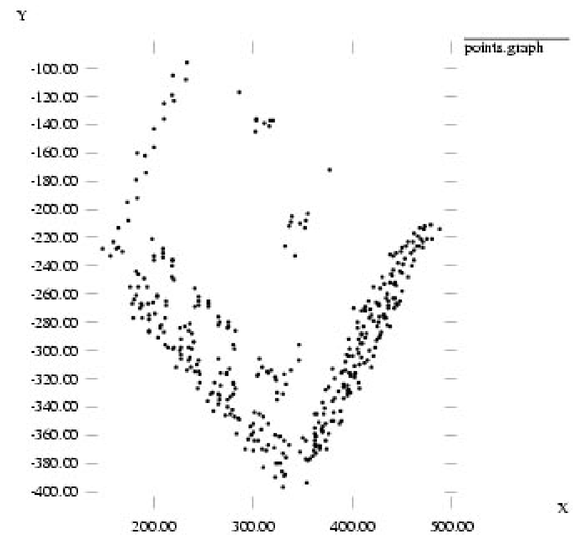
Problem: what is *Q* ? $$\hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{a}}_2 = 0 \quad \text{and} \quad |\hat{\boldsymbol{a}}_1|^2 = |\hat{\boldsymbol{a}}_2|^2$$

Note: Absolute scale cannot be recovered. The Euclidean shape (defined up to an arbitrary similitude) is recovered.

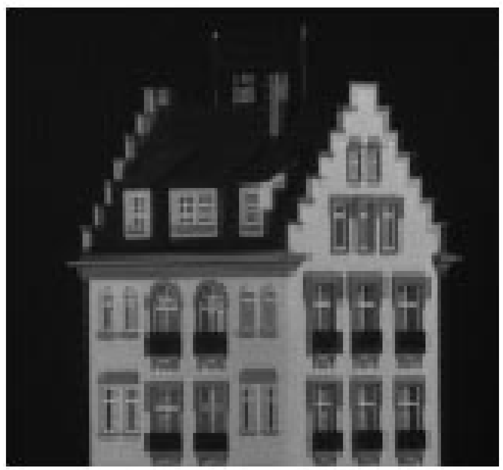
Reconstruction Results (Tomasi and Kanade, 1992)

1

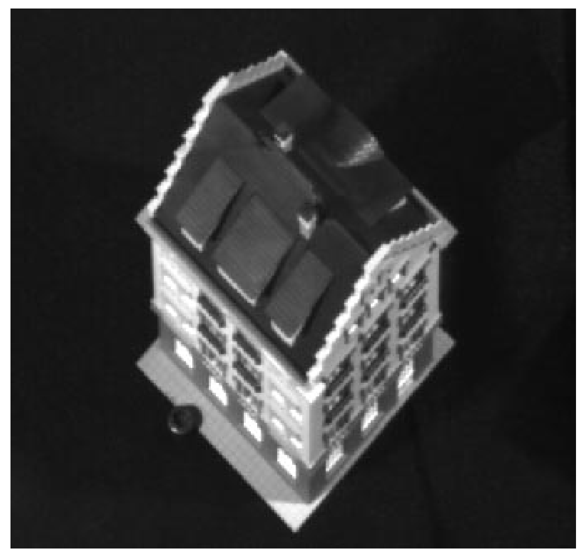
60



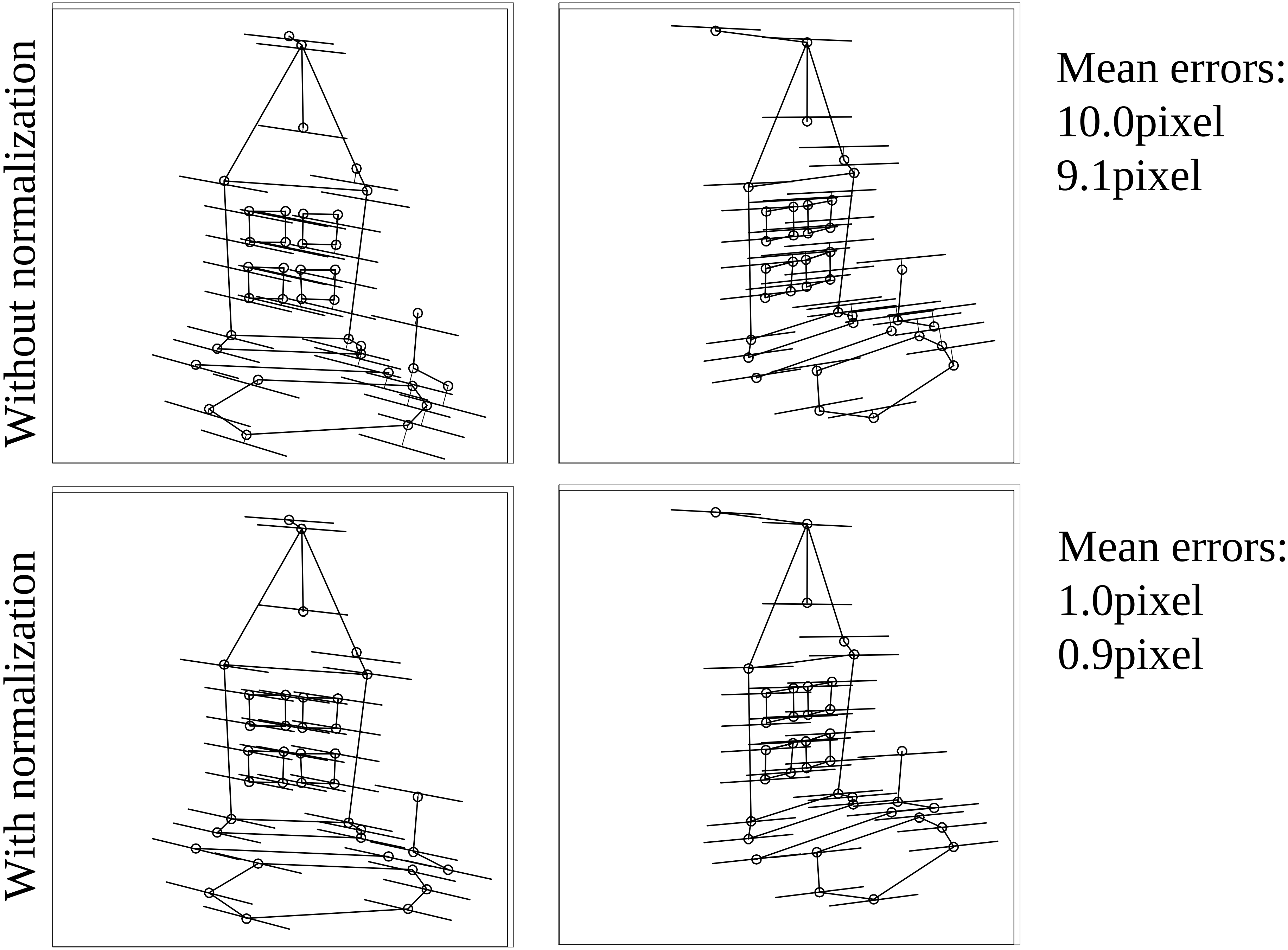
120

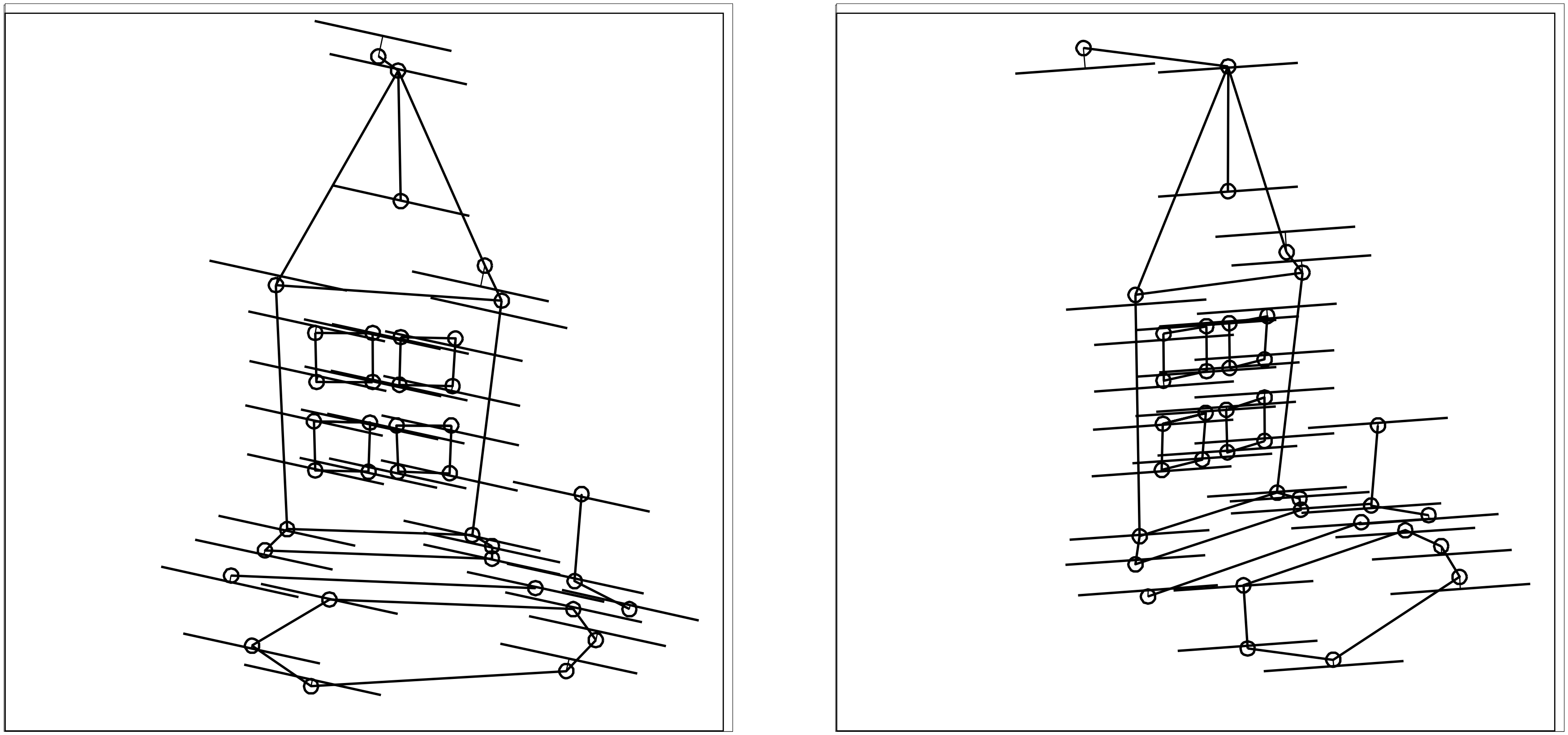


150



Reprinted from "Factoring Image Sequences into Shape and Motion," by C. Tomasi and T. Kanade, Proc. IEEE Workshop on Visual Motion (1991). © 1991 IEEE.


Without normalization
Mean errors:
10.0pixel
9.1pixel
With normalization
Mean errors:
1.0pixel
0.9pixel



Mean errors: 3.24 and 3.15 pixels