Autocalibration from Planar Scenes

Projective Direction Frames

• Simpler, more direct formulation of theory of autocalibration

Planar Autocalibration

- Camera calibration + Euclidean structure + motion from a *moving projective* camera viewing a *planar scene*
 - unknown scene & motion, unknown but constant intrinsic parameters
- ≥ 5 images needed for a full projective camera model (f, a, s, u_0, v_0)
- ≥ 3 images suffice if only focal length f is estimated
- Numerical optimization based algorithm, choice of initialization methods

Why Use Planes?

- 1 A common, easy to recognize primitive "every wall is a calibration grid"
- 2 Feature extraction & matching are relatively simple
- 3 A singular case for projective cameras, but not for calibrated ones
 - no camera motion information can be extracted
 - no 3D projective structure (only 2D planar structure)
- 4 *Effective* scene planarity is surprisingly common
 - distant scenes, small motions, dominant ground plane, ...

Direction Frames

Direction Frames are just **orthonormal sets of 3D directions**

- *i.e.* "orthonormal bases for the plane at infinity"
- In homogeneous Euclidean coordinates they become 4×3 orthogonal matrices

 $\mathbf{D} = \begin{pmatrix} \mathbf{R} \\ \mathbf{0} \end{pmatrix}$ where \mathbf{R} is a 3×3 rotation

- In projective 3D coordinates they become *arbitrary* 4×3 *rank* 3 matrices
- They are defined only up to orthogonal mixing of their columns

 $\mathbf{D} \longrightarrow \mathbf{D} \mathbf{R}$ where \mathbf{R} is a 3×3 rotation

Direction Frames and the Plane at Infinity

• The *null space* of a direction frame matrix is the *plane at infinity*

$$\mathbf{p}_{\infty}\,\mathbf{D}~=~oldsymbol{ heta}$$

• In projective coordinates where the first camera matrix is $\mathbf{P}_1 \simeq (\mathbf{I}_{3\times 3} | \mathbf{0})$ and the plane at infinity is $\mathbf{p}_{\infty} = (\mathbf{p}^{\top} 1)$, we can choose

$$\mathbf{D} = \begin{pmatrix} I \\ -p^{\top} \end{pmatrix} K$$

where K is the *internal camera calibration matrix*

$$\boldsymbol{K} = \begin{pmatrix} f & f s & u_0 \\ & f a & v_0 \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & s & u_0 \\ & a & v_0 \\ & & 1 \end{pmatrix} \begin{pmatrix} f & \\ & f & \\ & & 1 \end{pmatrix}$$

Basic Autocalibration Constraint

Orthogonal 3D directions project to orthogonal directions in the camera frame

The calibrated projection of a direction frame is orthogonal

 $K^{-1} P D \simeq 3 \times 3$ rotation

where

- K = internal camera calibration
- $\mathbf{P} = 3 \times 4$ camera projection
- $\mathbf{D} = 4 \times 3$ direction frame matrix

Derived Autocalibration Constraints

• To eliminate the unknown rotation, multiply by the transpose on left or right

 $\mathbf{D}^{ op} \mathbf{P}^{ op} \boldsymbol{\omega}^{-1} \mathbf{P} \mathbf{D} \simeq I_{3 imes 3}$ $K^{-1} \mathbf{P} \mathbf{\Omega} \mathbf{P}^{ op} K^{- op} \simeq I_{3 imes 3}$ $\mathbf{P} \mathbf{\Omega} \mathbf{P}^{ op} \simeq \boldsymbol{\omega}$ where $\mathbf{\Omega} = \mathbf{D} \mathbf{D}^{ op} = D$ ual Absolute 3D Quadric $\boldsymbol{\omega} = K K^{ op} = D$ ual Absolute Image Conic

- 5 constraints/image on 8 projective structure d.o.f. + \leq 5 unknown camera pars. \Rightarrow need $m \geq$ 3 images for full projective camera, $m \geq$ 2 if only f is estimated
- Resolve by *numerical optimization much* stabler than algebraic elimination

Planar Autocalibration Constraint

• Choose two vectors \mathbf{x}, \mathbf{y} of the direction frame parallel to the 3D plane

The calibrated projections of \mathbf{x}, \mathbf{y} are orthonormal up to scale $\|\boldsymbol{u}\|^2 = \|\boldsymbol{v}\|^2$ $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ where $(\boldsymbol{u}, \boldsymbol{v}) \equiv \boldsymbol{K}^{-1} \mathbf{P}(\mathbf{x}, \mathbf{y})$

• Equivalently, the plane's two *circular points* project onto the *image of the absolute conic* ω^{-1}

$$(\mathbf{P} \mathbf{x}_{\pm})^{\top} \boldsymbol{\omega}^{-1} (\mathbf{P} \mathbf{x}_{\pm}) = 0$$
 where $\mathbf{x}_{\pm} \equiv \mathbf{x} \pm i \mathbf{y}$

Planar Autocalibration from Homographies

- In practice, the plane is represented projectively by one of its images (say image 1)
- The projections \mathbf{P}_i become *inter-image homographies* H_{i1}

$$\begin{aligned} \|\boldsymbol{u}_i\|^2 &= \|\boldsymbol{v}_i\|^2 \\ \boldsymbol{u}_i \cdot \boldsymbol{v}_i &= 0 \end{aligned} \quad \text{where} \quad (\boldsymbol{u}_i, \boldsymbol{v}_i) \equiv \boldsymbol{K}_i^{-1} \boldsymbol{H}_{i1} \left(\boldsymbol{x}, \boldsymbol{y} \right) \end{aligned}$$

Algorithm

- Minimize the *autocalibration constraint residuals* over the 4 d.o.f. of (x, y) and the $n \le 5$ free calibration parameters
- Statistically motivated error model

error =
$$\sum_{i=1}^{m} \left(\frac{(\|\boldsymbol{u}_i\|^2 - \|\boldsymbol{v}_i\|^2)^2}{\|\boldsymbol{x}\|^2 \|\boldsymbol{a}_i\|^2 + \|\boldsymbol{y}\|^2 \|\boldsymbol{b}_i\|^2} + \frac{(\boldsymbol{u}_i \cdot \boldsymbol{v}_i)^2}{\|\boldsymbol{x}\|^2 \|\boldsymbol{b}_i\|^2 + \|\boldsymbol{y}\|^2 \|\boldsymbol{a}_i\|^2} \right)$$

where $(\boldsymbol{a}_i, \boldsymbol{b}_i) \equiv \boldsymbol{K}_i^{-\top}(\boldsymbol{u}_i, \boldsymbol{v}_i)$

- Can also include a *weak prior distribution* on calibration
 - e.g. gives default values for degenerate cases

How to Choose an Algebraic Error Model

- Poorly weighted algebraic error measure \Rightarrow seriously biased results!!
- "Normalization" (preconditioning) helps, but is not the whole answer

A better method

1 Start with an arbitrary algebraic error — here $\boldsymbol{e} = (\|\boldsymbol{u}\|^2 - \|\boldsymbol{v}\|^2, \, \boldsymbol{u} \cdot \boldsymbol{v})$

2 Use *covariance propagation* to estimate covariance of *e*

$$\operatorname{Cov}(\boldsymbol{e}) \approx \epsilon \cdot \begin{pmatrix} \boldsymbol{x}^2 \boldsymbol{a}^2 + \boldsymbol{y}^2 \boldsymbol{b}^2 & (\boldsymbol{x}^2 - \boldsymbol{y}^2) \boldsymbol{a} \cdot \boldsymbol{b} \\ (\boldsymbol{x}^2 - \boldsymbol{y}^2) \boldsymbol{a} \cdot \boldsymbol{b} & \boldsymbol{x}^2 \boldsymbol{b}^2 + \boldsymbol{y}^2 \boldsymbol{a}^2 \end{pmatrix}$$

3 The statistically correct error metric is

$$\chi^2 \approx \boldsymbol{e}^\top \operatorname{Cov}(\boldsymbol{e})^{-1} \boldsymbol{e}$$

4 If this is too complicated, approximate — you now know what to aim for!

Numerical Method

- Work with respect to a *nominal calibration* $(f, a, s, u_0, v_0) = (1, 1, 0, 0, 0)$
- Stabilized Gauss-Newton iteration converges quickly and reliably
 - ... but sometimes to the *wrong solution* !

Initialization

- Algebraic initialization seems difficult too many images & variables
- Instead use one of three *numerical initialization methods:*
- 1 Start search at nominal calibration
- 2 Line search over f, with nominal (a, s, u_0, v_0)
 - for each f, the problem reduces to relative orientation of two calibrated cameras from a planar scene (stable new SVD based method for this)
- 3 Hartley's 'rotating camera' method







Rules for Calibration Accuracy

• The *natural geometric error measures* are

$ \Delta f /f$	relative focal length error
$ \Delta u_0 /f, ~ \Delta v_0 /f$	principal point error relative to focal length
$ \Delta a $	absolute error in dimensionless aspect ratio
$ \Delta s $	absolute error in dimensionless skew

• For most auto/calibration methods with reasonably strong geometry

 $|\Delta f|/f \approx |\Delta u_0|/f \approx |\Delta v_0|/f$ $|\Delta a| \approx |\Delta s|$ are about one order of magnitude smaller

• If these rules fail, the geometry is probably weak (or the algorithm!)

— *e.g.* insufficient camera rotation during autocalibration.

A Nice Idea, but . . .

- All homographies are with respect to a *key image* this could bias results
- Use **homography factorization** to choose a more neutral frame
 - analogous to depth + factorization based projective reconstruction
 - the key is to find *consistent scale factors* for the homographies
- A nice idea, but in the end it *didn't improve the results*
 - ... this suggests that using a key image doesn't cause too much bias

Homography Factorization Method

1 Estimate homographies between all image pairs ($H_{ii} = I$, $H_{ij} = H_{ji}^{-1}$)

2 Find self-consistent relative scalings $H_{ij} \rightarrow \lambda_{ij} H_{ij}$:

— choose a key image 1 and define $\lambda_{i1}=\lambda_{1j}=1$

- enforce $H_{ij} \approx H_{i1} H_{1j}$ by choosing $\lambda_{ij} = \frac{\text{Trace}(H_{i1} H_{1j} \cdot H_{ij}^{\top})}{\text{Trace}(H_{ij} \cdot H_{ij}^{\top})}$

- **balance** the scales for numerical stability

3 Rescale and factorize to rank 3

$$\begin{pmatrix} \lambda_{11}\boldsymbol{H}_{11} & \cdots & \lambda_{1m}\boldsymbol{H}_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{m1}\boldsymbol{H}_{m1} & \cdots & \lambda_{mm}\boldsymbol{H}_{mm} \end{pmatrix} \approx \begin{pmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_m \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{H}}_1^{-1} & \cdots & \tilde{\boldsymbol{H}}_m^{-1} \end{pmatrix}$$



	f only	f	a	S	u_0	v_0
calibration	-	1515±4	0.9968±0.0002	-	271±3	264±4
6 images	1584±63	1595±63	0.9934±0.0055	0.000±0.001	268±10	271±22
8 images	1619±25	1614±42	$0.9890 {\pm} 0.0058$	-0.005±0.005	289±3	320±26
10 images	1612±19	1565±41	1.0159±0.0518	-0.004 ± 0.006	273±5	286±27

Summary

Direction Frame approach to Autocalibration

- Conceptually simpler avoids most of the abstract algebraic geometry
- Much stabler & fewer degeneracies than Kruppa approach

Planar Autocalibration

- Accuracy seems reasonable
- Degeneracies seem to be those of 3D autocalibration
- Many images required 8–10 recommended in practice
- Occasional convergence to false solutions remains a problem

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