Chapter 7: Computation of the Camera Matrix P

Arco Nederveen

Eagle Vision

March 18, 2008

1. Chapter 7: Computation of the camera Matrix P
    - Basic Equations
    - Geometric error
    - Restricted camera estimation
    - Radial distortion

Basic equations

- Chapter 7: Numerical methods for estimation of the camera matrix.
- Given point correspondences $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ between 3D points $\mathbf{X}_i$ and 2D image points $\mathbf{x}_i$ find the camera matrix $\mathtt{P}$.
- Where $\mathtt{P}$ is a 3×4 matrix, such that $\mathbf{x}_i = \mathtt{P}\mathbf{X}_i$ for all i
- For each correspondence:

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i^\top & y_i\mathbf{X}_i^\top \\ w_i\mathbf{X}_i^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i^\top \\ -y_i\mathbf{X}_i^\top & x_i\mathbf{X}_i^\top & \mathbf{0}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} \tag{1}$$

Where each $\mathbf{P}^{i\top}$ is a 4-vector, the i-th row of $\mathtt{P}$.

Basic equations

- Alternatively:

$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i^\top & y_i\mathbf{X}_i^\top \\ w_i\mathbf{X}_i^\top & \mathbf{0}^\top & -x_i\mathbf{X}_i^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix}$$

- A set of n correspondences results in $2n \times 12$ matrix $\mathtt{A}$.
- The projection matrix $\mathtt{P}$ is computed by solving. $\mathtt{A}\mathbf{p} = \mathbf{0}$. Where $\mathbf{p}$ contains the entries of $\mathtt{P}$.
- Minimal solution: $\mathtt{P}$ has 12 entries and 11 degrees of freedom, therefore $5\frac{1}{2}$ correspondences are needed.
- Overdetermined solution: Minimize algebraic or geometric error.

Basic equations

- Algebraic error: minimize $||\mathtt{A}\mathbf{p}||$ with a normalization constraint. like $||\mathbf{p}|| = 1$ or $||\hat{\mathbf{p}}^3|| = 1$ where $\hat{\mathbf{p}}^3$ is the vector. $(p_{31}, p_{32}, p_{33})^\top$ from $\mathtt{P}$
- Algebraic error: the residual $\mathtt{A}\mathbf{p}$.
- The DLT algorithm can be used to compute $\mathtt{P}$ in the same manner as computing $\mathtt{H}$.

Basic equations

- Degenerate configurations:
  - The camera and points all lie on a twisted cubic.
  - The points lie on a union of a plane and a single straight line containing the camera center.
- Data normalization:
  - 2D points: Centroid of data at origin and their RMS distance from the origin is $\sqrt{2}$.
  - 3D points: If variation of depth of points is relatively small than position centroid of data on the origin and scale the RMS of the points to $\sqrt{3}$

Basic equations

- Using Line correspondences in the DLT algorithm
  - Line in 3D can be represented by $\mathbf{X}_0$ and $\mathbf{X}_1$
  - The plane formed by back-projecting the image line $\mathbf{l}$ is equal to $\mathtt{P}^\top \mathtt{l}$
  - Than $\mathbf{X}_j$ lies on this plane if:

$$\mathtt{P}^\top \mathtt{l}_j = 0 \texttt{ for } j = 0, 1$$

  - Each choice for j gives a single linear equation which may be added to equation 1.

Geometric error

- If world points $\mathbf{X_i}$ are exactly known and the 2D image points contain noise then the geometric error is given by:

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2$$

  Where $\mathbf{x_i}$ is the measured point and $\hat{\mathbf{x}}_i$ is the point $\mathtt{P}\mathbf{X}_i$.

- If the measurement noise is Gaussian then:

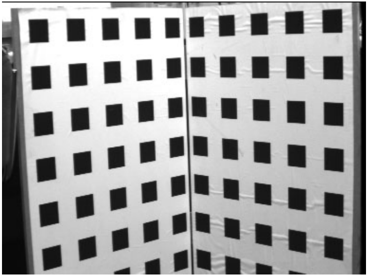
$$\min_{\mathtt{P}} \sum_i d(\mathbf{x_i}, \mathtt{P}\mathbf{X}_i)^2$$

  is the Maximum Likelihood estimate of $\mathtt{P}$.

- To minimize the geometric a iterative method must be used. Such as used in the Gold Standard algorithm in table 7.1.

Geometric Error

- Example 7.1: Camera estimation from a calibration object.
  - DLT algorithm compared to Gold Standard.
  - Rule of thumb: point measurements should exceed the number of unknowns by a factor five.
  - Table 7.1 shows the results: slight improvement with Gold Standard algorithm.



Errors in the world points

- Its possible world points cannot be determined with absolute accuracy.
- If this is the case the 3D geometric error is defined as:

$$\sum_i d(\mathbf{X}_i, \hat{\mathbf{X}}_\mathbf{i})^2$$

Where $\hat{\mathbf{X}}_i$ is the closest point in space to $\mathbf{X}_i$ that maps exactly onto $\mathbf{x}_i$ via $\mathbf{x}_i = \mathtt{P}\hat{\mathbf{X}}_i$.

- If both the world and image coordinates contain errors the sum of world and image errors is minimized.

$$\sum_i^n d_{\mathrm{Mah}}(\mathbf{x}_i, \mathtt{P}\hat{\mathbf{X}}_i)^2 + d_{\mathrm{Mah}}(\mathbf{X}_i, \hat{\mathbf{X}}_i)^2$$

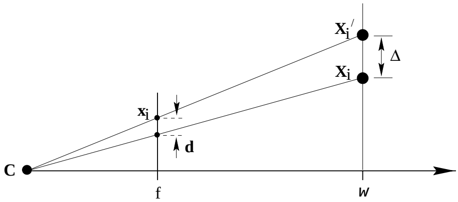
Where d_{Mah} represents the Mahalanobis distance.

Geometric interpretation of algebraic error

- Points $\mathbf{X}_i$ in the DLT algorithm are normalized such that $\mathbf{X}_i = (X_i, Y_i, Z_i, 1)^\top$ and $\mathbf{x_i} = (x_i, y_i, 1)^\top$
- The quantity minimized is then given by: $\sum_i (\hat{w}_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i))^2$ where $\hat{w}_i(\hat{x}_i, \hat{y}_i, 1)^\top = \mathtt{P}\mathbf{X}_i$
- But given, $\hat{w}_i = \pm||\hat{\mathbf{p}}^3||\mathrm{depth}(\mathbf{X}; \mathtt{P})$ we see that $\hat{w}_i$ can be seen as the depth of point $\mathbf{X}_i$ from the camera in the direction of principle ray if $||\hat{\mathbf{p}}^3||^2 = 1$.

Geometric interpretation of algebraic error

- As seen in figure 7.2 $\hat{w}_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)$ is proportional to $f d(\mathbf{X}', \mathbf{X})$
- The algebraic distance being minimized is equal to $f \sum_i d(\mathbf{X}_i, \mathbf{X}'_i)^2$



Transformation invariance

- Scaling and translation in either the image of world coordinates have no influence on minimizing $||\mathtt{A}\mathbf{p}||$.

Estimation of an affine camera

- Affine camera is one for which the projection matrix has $(0,0,0,1)$ as last row.
- Given $\mathbf{X}_i = (X_i, Y_i, Z_i, 1)^\top$ and $\mathbf{x_i} = (x_i, y_i, 1)^\top$ equation 1 reduces to:
$$\begin{bmatrix} \mathbf{0}^\top & -w_i\mathbf{X}_i^\top \\ \mathbf{X}_i^\top & \mathbf{0}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \end{pmatrix} + \begin{pmatrix} y_i \\ -x_i \end{pmatrix}$$
which shows that the squared algebraic error in this case equals the squared geometric error
$||\mathtt{A}\mathbf{p}||^2 = \sum_i (x_i - \mathbf{P}^{1\top}\mathbf{X}_i)^2 + (y_i - \mathbf{P}^{2\top}\mathbf{X}_i) = \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2.$
- A linear estimation algorithm is given in algorithm 7.2.

Restricted camera estimation

- Matrix $\mathtt{P}$ with centre at finite point $\mathtt{P} = \mathtt{K}[\mathtt{R} | -\mathtt{R}\tilde{\mathbf{C}}]$. With

$$\mathtt{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

- With common assumptions:
  - The skew s is zero.
  - The pixels are square: $\alpha_x = \alpha_y$.
  - The principal point (x_0, y_0) is known.
  - The complete camera calibration matrix $\mathtt{K}$ is known.

Restricted camera estimation

- Some assumptions can make it possible to estimate the restricted camera matrix with a linear algorithm.
- Example: fit a pinhole camera model ($s = 0$ and $\alpha_x = \alpha_y$) to a set of point measurements.
- Minimize geometric error.
  - Parameterize the camera model using the remaining 9 parameters. ($x_0, y_0, \alpha, 6$ from $\mathtt{R}$ and $\tilde{\mathbf{C}}$ denoted collectively by $\mathbf{q}$)
  - Iterative minimization of geometric error.
- Minimizing algebraic error.
  - Iterative minimization problem becomes smaller.
  - Equivalent to minimizing $||\mathtt{A}g(\mathbf{q})||$.

Restricted camera estimation

- The reduced measurement matrix.
  - Possible to replace matrix $\mathtt{A}$ with a square 12×12 matrix $\hat{\mathtt{A}}$ such that $||\mathtt{A}\mathbf{p}|| = \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} = ||\hat{\mathtt{A}}\mathbf{p}||$.
  - By SVD: Let $\mathtt{A} = \mathtt{UDV}^\top$ be the SVD of $\mathtt{A}$, and define $\hat{\mathtt{A}} = \mathtt{DV}^\top$. Then

$$\mathtt{A}^\top \mathtt{A} = (\mathtt{VDU}^\top)(\mathtt{UDV}^\top) = (\mathtt{VD})(\mathtt{DV}^\top) = \hat{\mathtt{A}}^\top \hat{\mathtt{A}}$$

  - Given a set of n world to image correspondence, $\mathbf{X_i} \leftrightarrow \mathbf{x_i}$, the problem of finding a constrained camera matrix $\mathtt{P}$ that minimizes the sum of algebraic distances $\sum_i d_{\mathrm{alg}}(\mathbf{x_i}, \mathtt{P}\mathbf{X}_i)^2$ reduces to the minimization of a function $\mathbb{R}^9 \rightarrow \mathbb{R}^{12}$, independent of the number n correspondences.

Restricted camera estimation

- Initialization of the iteration
  - Use a linear algorithm such as DLT to find initial camera matrix.
  - Clamp fixed parameters
  - Set variable parameter to their values obtained by decomposition of initial camera matrix
- The values obtained from the DLT can differ significant from the real values. Therefore use soft constraints and add extra terms to the cost function. E.g. if $s = 0$ and $\alpha_x = \alpha_y$ the geometric error becomes:

$$\sum_i d(\mathbf{x}_i, \mathtt{P}\mathbf{X}_i)^2 + ws^2 + w(\alpha_x - \alpha_y)^2$$

Restricted camera estimation

- Exterior orientation
  - All internal parameters of the camera are known.
  - Only position and orientation of the camera unknown.
  - Six degrees of freedom remain and can be determined with three points.
- Experimental evaluation (table 7.2)
  - Same results
  - Algrabraic method is quicker (12 errors vs. $2n = 396$)

Restricted camera estimation

- Covariance estimation
  - Similar to 2D homography case.
  - Assuming all errors are in the image measurements the residual error is equal to:

$$\epsilon_{\text{res}} = \sigma(1 - d/2n)^{1/2}$$

    With d the number of camera parameters being fitted. (Example 7.1: $\epsilon = 0.365$ results in $\sigma = 0.37$).

Restricted camera estimation

- Covariance ellipsoid for an estimated camera.
- Backpropogate covariance of the point measurements back to camera models. Which results in:

$$\Sigma_{\text{camera}} = (\mathtt{J}^\top \Sigma_{\text{points}}^{-1} \mathtt{J})^{-1}$$

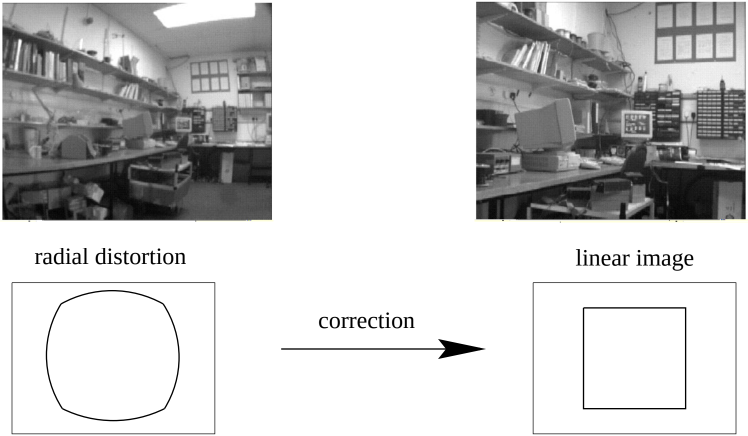
  With $\mathtt{J}$ Jacobian of measured points.
- Error bounds of ellipsoid can now be computed.

$$(\mathbf{C} - \bar{\mathbf{C}})^\top \Sigma_{\mathbf{C}}^{-1} (\mathbf{C} - \bar{\mathbf{C}}) = k^2$$

  Where $k^2 = F_n^{-1}(\alpha)$ from the inverse cumulative χ_n^2 with confidence level α and n the number of variables.

Radial distortion

- Assumed that a linear model is accurate for imaging.
- Not the case for real lenses.
- Radial distortion



Radial distortion

- Radial distortion is modeled by:

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = L(\tilde{r}) \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}$$

  - $(\tilde{x}, \tilde{y})$ is the ideal position
  - (x_d, y_d) actual image position
  - $\tilde{r}$ is the radial distance $\sqrt{\tilde{x}^2 + \tilde{x}^2}$ from the centre of distortion
  - $L(\tilde{r})$ is a distortion factor

Radial distortion

- Correction of the coordinates in pixel coordinates is given by:

$$\begin{aligned}\hat{x} &= x_c + L(r)(x - x_c)\\ \hat{y} &= y_c + L(r)(y - y_c)\end{aligned}$$

  Where (x, y) are the measured coordinates, $(\hat{x}, \hat{y})$ corrected coordinates, (x_c, y_c) the center of radial distortion, and $r^2 = (x - x_c)^2 + (y - y_c)^2$.

- $L(r)$ is approximated by a Taylor expansion
  $L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \kappa_3 r^3 + \ldots$

Radial distortion

- Principal point is taken as the centre of radial distortion.
- $L(r)$ may be computed by minimizing a cost based on a the deviation from a linear mapping.
- Parameters κ_i can be computed together with $\mathtt{P}$ during minimization of the geometric error.
- Alternatively one can use known straight lines to guide minimization process.
- Example 7.3 and 7.4 show an example of radial distortion compensation and the effect on the residual error (Table 7.3).