Theory of Bouguet's MatLab Camera Calibration Toolbox: Stereo

Yuji Oyamada

[1]HVRL, University

[2]Chair for Computer Aided Medical Procedure (CAMP)
Technische Universität München

June 5, 2012

Variables

Assumption:

- N cameras observing L key-points resulting M images for each sequence.
- The cameras are fixed meaning relative positions among them are same through the sequence.

Variables:

- $\mathbf{a}_j$: Intrinsic parameters of j-th camera.
- $\mathbf{b}_{ij}$: Extrinsic parameters of i-th image of j-th camera.
- $\mathbf{x}_{ijk}$: k-th key-point of i-th image of j-th camera.

where

- $j = 1, \ldots, N$
- $i = 1, \ldots, M$
- $k = 1, \ldots, L_{ij}$

Variables

The set of extrinsic parameters $\{\mathbf{b}_{ij}\}$ has redundancy because the cameras are fixed.
Introduce another variable $\mathbf{r}$:

- $\mathbf{b}_i$: Extrinsic parameters of i-th image of 1st camera.
- $\mathbf{r}_j$: Extrinsic parameters of j-th camera w.r.t. 1st camera.

Note that $\mathbf{r}_1$ is equivalent to identity matrix.

Variables

We have observation vector **x** and parameters vector **p** where

$$\mathbf{x} = (\mathbf{x}_{111}^\top, \ldots, \mathbf{x}_{11L_{11}}^\top, \mathbf{x}_{1N1}^\top, \ldots, \mathbf{x}_{1NL_{1N}}^\top, \ldots, \mathbf{x}_{MN1}^\top, \ldots, \mathbf{x}_{MNL_{MN}}^\top)^\top$$
$$= (\mathbf{x}_{11}^\top, \mathbf{x}_{1N}^\top, \ldots, \mathbf{x}_{MN}^\top, \ldots, \mathbf{x}_{MN}^\top)^\top$$
$$\mathbf{p} = (\mathbf{a}_1^\top, \ldots, \mathbf{a}_N^\top, \mathbf{r}_2^\top, \ldots, \mathbf{r}_N^\top, \mathbf{b}_1^\top, \ldots, \mathbf{b}_M^\top)^\top$$

where $\mathbf{x}_{ij} = (\mathbf{x}_{ij1}^\top, \ldots, \mathbf{x}_{ijL_{ij}}^\top)^\top$

Non-linear optimization

Finds optimal parameters $\mathbf{p}$ as

$$\{\{\hat{\mathbf{a}}_j\}\{\hat{\mathbf{r}}_j\}\{\hat{\mathbf{b}}_i\}\} = \arg \min_{\{\mathbf{a}_j\}\{\mathbf{r}_j\}\{\mathbf{b}_i\}} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=1}^{L_{ij}} v_{ijk} \mathrm{dist}(\hat{\mathbf{x}}_{ijk}, \mathbf{x}_{ijk})^2$$

where

- $\hat{\mathbf{x}}_{ijk} = Q(\mathbf{a}_j, \mathbf{b}_i)$ denotes a reprojected point of $\mathbf{x}_{ijk}$ with parameters $\mathbf{a}_j$ and $\mathbf{b}_i$,
- a visibility term $v_{ijk} = 1$ iff k-th point is visible in i-th image observed by j-th camera.

Normal equations

$$\mathbf{J}^\top \mathbf{J}\delta = -\mathbf{J}^\top \boldsymbol{\epsilon}$$

where

$$\mathbf{J} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{p}} = \left[\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{a}} \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{r}} \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{b}} \right] = [\mathbf{ARB}]$$

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{a}}, \mathbf{R} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{r}}, \mathbf{B} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{b}}$$

$$\mathbf{A}_{ij} = \frac{\partial \hat{\mathbf{x}}_{ij}}{\partial \mathbf{a}_j}, \mathbf{R}_{ij} = \frac{\partial \hat{\mathbf{x}}_{ij}}{\partial \mathbf{r}_j}, \mathbf{B}_{ij} = \frac{\partial \hat{\mathbf{x}}_{ij}}{\partial \mathbf{b}_i}$$

Structure of Jacobian matrix $\mathbf{J}$

$$
\begin{array}{c|ccccccccccc|}
 & \boldsymbol{a}_1^\top & \boldsymbol{a}_2^\top & \cdots & \boldsymbol{a}_N^\top & \boldsymbol{r}_2^\top & \cdots & \boldsymbol{r}_N^\top & \boldsymbol{b}_1^\top & \boldsymbol{b}_2^\top & \cdots & \boldsymbol{b}_M^\top \\
\hline
\boldsymbol{x}_{11}^\top & \boldsymbol{A}_{11}^\top & & & & & & & \boldsymbol{B}_{11}^\top & & & \\
\boldsymbol{x}_{12}^\top & & \boldsymbol{A}_{12}^\top & & & \boldsymbol{R}_{12}^\top & & & \boldsymbol{B}_{12}^\top & & & \\
\vdots & & & \ddots & & & \ddots & & \vdots & & & \\
\boldsymbol{x}_{1N}^\top & & & & \boldsymbol{A}_{1N}^\top & & & \boldsymbol{R}_{1N}^\top & \boldsymbol{B}_{1N}^\top & & & \\
\boldsymbol{x}_{21}^\top & \boldsymbol{A}_{21}^\top & & & & & & & & \boldsymbol{B}_{21}^\top & & \\
\boldsymbol{x}_{22}^\top & & \boldsymbol{A}_{22}^\top & & & \boldsymbol{R}_{22}^\top & & & & \boldsymbol{B}_{22}^\top & & \\
\vdots & & & \ddots & & & \ddots & & & \vdots & & \\
\boldsymbol{x}_{2N}^\top & & & & \boldsymbol{A}_{2N}^\top & & & \boldsymbol{R}_{2N}^\top & & \boldsymbol{B}_{2N}^\top & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\boldsymbol{x}_{M1}^\top & \boldsymbol{A}_{M1}^\top & & & & & & & & & & \boldsymbol{B}_{M1}^\top \\
\boldsymbol{x}_{M2}^\top & & \boldsymbol{A}_{M2}^\top & & & \boldsymbol{R}_{M2}^\top & & & & & & \boldsymbol{B}_{M2}^\top \\
\vdots & & & \ddots & & & \ddots & & & & & \vdots \\
\boldsymbol{x}_{MN}^\top & & & & \boldsymbol{A}_{MN}^\top & & & \boldsymbol{R}_{MN}^\top & & & & \boldsymbol{B}_{MN}^\top
\end{array}
$$

Sparse LM

- LM is suitable for minimization w.r.t. a small number of parameters.
- The central step of LM, solving the normal equations,
  - has complexity N^3 in the number of parameters and
  - is repeated many times.
- The normal equation matrix has a certain sparse block structure.

Sparse LM

- Let $\mathbf{p} \in \mathbb{R}^M$ be the parameter vector that is able to be partitioned into parameter vectors as $\mathbf{p} = (\mathbf{a}^\top, \mathbf{b}^\top)^\top$.
- Given a measurement vector $\mathbf{x} \in \mathbb{R}^N$
- Let $\sum_{\mathbf{x}}$ be the covariance matrix for the measurement vector.
- A general function $f : \mathbb{R}^M \to \mathbb{R}^N$ takes $\mathbf{p}$ to the estimated measurement vector $\hat{\mathbf{x}} = f(\mathbf{p})$.
- ϵ denotes the difference $\mathbf{x} - \hat{\mathbf{x}}$ between the measured and the estimated vectors.

Sparse LM

The set of equations $\mathbf{J}\boldsymbol{\delta} = \boldsymbol{\epsilon}$ solved as the central step in the LM has the form

$$\mathbf{J}\boldsymbol{\delta} = [\mathbf{A}|\mathbf{B}]\left(\frac{\boldsymbol{\delta}_a}{\boldsymbol{\delta}_b}\right) = \boldsymbol{\epsilon}.$$

Then, the normal equations $\mathbf{J}^\top \sum_x^{-1} \mathbf{J}\boldsymbol{\delta} = \mathbf{J}^\top \sum_x^{-1} \boldsymbol{\epsilon}$ to be solved at each step of LM are of the form

$$\left[\begin{array}{c|c} \mathbf{A}^\top \sum_{\mathbf{x}}^{-1} \mathbf{A} & \mathbf{A}^\top \sum_{\mathbf{x}}^{-1} \mathbf{B} \\ \hline \mathbf{B}^\top \sum_{\mathbf{x}}^{-1} \mathbf{A} & \mathbf{B}^\top \sum_{\mathbf{x}}^{-1} \mathbf{B} \end{array}\right] \left(\frac{\boldsymbol{\delta}_a}{\boldsymbol{\delta}_b}\right) = \left(\frac{\mathbf{A}^\top \sum_{\mathbf{x}}^{-1} \boldsymbol{\epsilon}}{\mathbf{B}^\top \sum_{\mathbf{x}}^{-1} \boldsymbol{\epsilon}}\right)$$

Sparse LM

Let

- $\mathbf{U} = \mathbf{A}^\top \sum_{\mathbf{x}}^{-1} \mathbf{A}$
- $\mathbf{W} = \mathbf{A}^\top \sum_{\mathbf{x}}^{-1} \mathbf{B}$
- $\mathbf{V} = \mathbf{B}^\top \sum_{\mathbf{x}}^{-1} \mathbf{B}$

and $\cdot^*$ denotes augmented matrix by λ.
The normal equations are rewritten as

$$\begin{bmatrix} \mathbf{U}^* & \mathbf{W} \\ \mathbf{W}^\top & \mathbf{V}^* \end{bmatrix} \begin{pmatrix} \boldsymbol{\delta}_a \\ \boldsymbol{\delta}_b \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_A \\ \boldsymbol{\epsilon}_B \end{pmatrix}$$

$$\rightarrow \begin{bmatrix} \mathbf{U}^* - \mathbf{W}\mathbf{V}^{*-1}\mathbf{W}^\top & \mathbf{0} \\ \mathbf{W}^\top & \mathbf{V}^* \end{bmatrix} \begin{pmatrix} \boldsymbol{\delta}_a \\ \boldsymbol{\delta}_b \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_A - \mathbf{W}\mathbf{V}^{*-1}\boldsymbol{\epsilon}_B \\ \boldsymbol{\epsilon}_B \end{pmatrix}$$

This results in the elimination of the top right hand block.

Sparse LM

The top half of this set of equations is

$$(\mathbf{U}^* - \mathbf{W}\mathbf{V}^{*-1}\mathbf{W}^\top)\boldsymbol{\delta}_a = \boldsymbol{\epsilon}_A - \mathbf{W}\mathbf{V}^{*-1}\boldsymbol{\epsilon}_B$$

Subsequently, the value of $\boldsymbol{\delta}_a$ may be found by back-substitution, giving

$$\mathbf{V}^*\boldsymbol{\delta}_b = \boldsymbol{\epsilon}_B - \mathbf{W}^\top\boldsymbol{\delta}_a$$

Sparse LM

$\mathbf{p} = (\mathbf{a}^\top, \mathbf{b}^\top)^\top$, where $\mathbf{a} = (fc^\top, cc^\top, alpha\_c^\top, kc^\top)^\top$ and $\mathbf{b} = (\{omc\_i^\top \; Tc\_i^\top\})^\top$

The Jacobian matrix is

$$\mathbf{J} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{p}} = \left[\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{a}}, \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{b}}\right] = [\mathbf{A}, \mathbf{B}]$$

where

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{a}} = \left[\frac{\partial \hat{\mathbf{x}}}{\partial fc}, \frac{\partial \hat{\mathbf{x}}}{\partial cc}, \frac{\partial \hat{\mathbf{x}}}{\partial alpha\_c}, \frac{\partial \hat{\mathbf{x}}}{\partial kc}\right]$$

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{b}} = \left[\frac{\partial \hat{\mathbf{x}}}{\partial omc\_1}, \frac{\partial \hat{\mathbf{x}}}{\partial Tc\_1}, \cdots, \frac{\partial \hat{\mathbf{x}}}{\partial omc\_i} \frac{\partial \hat{\mathbf{x}}}{\partial Tc\_i} \cdots, \frac{\partial \hat{\mathbf{x}}}{\partial omc\_N} \frac{\partial \hat{\mathbf{x}}}{\partial Tc\_N}\right]$$

Sparse LM

The normal equation is rewritten as

$$\left(\begin{bmatrix}\mathbf{A}^\top\\ \mathbf{B}^\top\end{bmatrix}\begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}+\lambda\mathbf{I}\right)\Delta_{\mathbf{p}}=-\begin{bmatrix}\mathbf{A}^\top\\ \mathbf{B}^\top\end{bmatrix}\epsilon_{\mathbf{x}}$$

$$\rightarrow\left(\begin{bmatrix}\sum_{i=1}^{N}\mathbf{A}_i^\top\mathbf{A}_i & \sum_{i=1}^{N}\mathbf{A}_i^\top\mathbf{B}_i\\ \sum_{i=1}^{N}\mathbf{B}_i^\top\mathbf{A}_i & \sum_{i=1}^{N}\mathbf{B}_i^\top\mathbf{B}_i\end{bmatrix}+\lambda\mathbf{I}\right)\Delta_{\mathbf{p}}=-\begin{bmatrix}\mathbf{A}^\top\epsilon_{\mathbf{x}}\\ \mathbf{B}^\top\epsilon_{\mathbf{x}}\end{bmatrix}$$

Sparse LM

$$
\mathbf{J}^\top \mathbf{J} = \left[\begin{array}{c|ccccc}
\sum_{i=1}^{N} \mathbf{A}_i^\top \mathbf{A}_i & \mathbf{A}_1^\top \mathbf{B}_1 & \cdots & \mathbf{A}_i^\top \mathbf{B}_i & \cdots & \mathbf{A}_N^\top \mathbf{B}_N \\
\hline
\mathbf{B}_1^\top \mathbf{A}_1 & \mathbf{B}_1^\top \mathbf{B}_1 & & & & \\
\vdots & & \ddots & & & \\
\mathbf{B}_i^\top \mathbf{A}_i & & & \mathbf{B}_i^\top \mathbf{B}_i & & \\
\vdots & & & & \ddots & \\
\mathbf{B}_N^\top \mathbf{A}_N & & & & & \mathbf{B}_N^\top \mathbf{B}_N
\end{array}\right]
$$

Sparse LM

$$\mathbf{J}^\top \epsilon_{\mathbf{x}} = \left[\begin{array}{c} \sum_{i=1}^{N} \mathbf{A}_i^\top \epsilon_{\mathbf{x}} \\ \hline \mathbf{B}_1^\top \epsilon_{\mathbf{x}} \\ \vdots \\ \mathbf{B}_i^\top \epsilon_{\mathbf{x}} \\ \vdots \\ \mathbf{B}_N^\top \epsilon_{\mathbf{x}} \end{array}\right]$$

Sparse LM

When each image has different number of corresponding points ($M_i \neq M_j$, if $i \neq j$), each $\mathbf{A}_i$ and $\mathbf{B}_i$ have different size as

$$\begin{bmatrix} \begin{bmatrix} \frac{dx_1}{dfc(1)} & \frac{dy_1}{dfc(1)} & \cdots\cdots & \frac{dx_{M_1}}{dkc(5)} & \frac{dy_{M_1}}{dkc(5)} \end{bmatrix} & \begin{bmatrix} \frac{dx_1}{domc\_1(1)} & \frac{dy_1}{domc\_1(1)} & \cdots\cdots & \frac{dx_{M_1}}{dTc\_1(3)} & \frac{dy_{M_1}}{dTc\_1(3)} \end{bmatrix} \\ \begin{bmatrix} \frac{dx_1}{dfc(1)} & \frac{dy_1}{dfc(1)} & \cdots\cdots\cdots & \frac{dx_{M_1}}{dkc(5)} & \frac{dy_{M_1}}{dkc(5)} \end{bmatrix} & \begin{bmatrix} \frac{dx_1}{domc\_1(1)} & \frac{dy_1}{domc\_1(1)} & \cdots\cdots\cdots & \frac{dx_{M_1}}{dTc\_1(3)} & \frac{dy_{M_1}}{dTc\_1(3)} \end{bmatrix} \\ \begin{bmatrix} \frac{dx_1}{dfc(1)} & \frac{dy_1}{dfc(1)} & \cdots & \frac{dx_{M_1}}{dkc(5)} & \frac{dy_{M_1}}{dkc(5)} \end{bmatrix} & \begin{bmatrix} \frac{dx_1}{domc\_1(1)} & \frac{dy_1}{domc\_1(1)} & \cdots & \frac{dx_{M_1}}{dTc\_1(3)} & \frac{dy_{M_1}}{dTc\_1(3)} \end{bmatrix} \end{bmatrix}$$

Sparse LM

However, the difference does not matter because

$$\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{d_{\text{int}} \times d_{\text{int}}}$$
$$\mathbf{A}^\top \mathbf{B} \in \mathbb{R}^{d_{\text{int}} \times d_{\text{ex}}}$$
$$\mathbf{B}^\top \mathbf{A} \in \mathbb{R}^{d_{\text{ex}} \times d_{\text{int}}}$$
$$\mathbf{B}^\top \mathbf{B} \in \mathbb{R}^{d_{\text{ex}} \times d_{\text{ex}}}$$

where d_{int} denotes dimension of intrinsic params and d_{ex} denotes dimension of extrinsic params.