A Unified Approach Combining Photometric and Geometric Information for Pose Estimation

Pierre Georgel
Agenda

• Reminder
  – Triangulation
  – Pose estimation
  – Bundle adjustment

• Hybrid registration
  – New cost function
  – Pitfall to avoid
What you know

Image 1
- Feature Points Extraction

Image 2
- Feature Points Extraction

Matching
- Coord. of Corresp. points

Linear Estimation

Non-linear Estimation

Pose Estimate

Photometric Information

Geometric Information
What you know

Image 1
Feat. Points Extraction

Image 2
Feat. Points Extraction

Matching
Coord. of Corresp. points

Linear Estimation
Non-linear Estimation

Pose Estimate
Triangulation

Photometric Information

Geometric Information

Feat. Points Extraction
Extraction
Matching
Linear Estimation
Non-linear Estimation
Pose Estimate
Triangulation
Epipolar Geometry

Fast Forward

• Compute fundamental matrix using keypoints \( \langle p_i, q_i \rangle \)
• This is done by minimize a cost function, for example the epipolar constraint

\[
q_i^\top F p_i = 0
\]

• Derive a canonical pose \( P \)
Triangulation

\[ P^0 = [I_0] \]
Triangulation

\[ \mathbf{P}^0 = \begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} \]
Triangulation

\[ \begin{array}{c}
M \sim p \\
PM \sim q \\
\end{array} \]
Triangulation

\[
\begin{align*}
\mathcal{M} & \sim p \\
PM & \sim q \\
\mathcal{M} & = \tilde{z}p \\
PM & = \tilde{z}'q
\end{align*}
\]
with \( P^0 \top \mathcal{M} = \tilde{z} \) and \( P_3 \mathcal{M} = \tilde{z}' \)
Triangulation

\[
\begin{align*}
\mathcal{M} &\sim p \\
\mathcal{P}\mathcal{M} &\sim q \\
\mathcal{M} &= \mathcal{z}p \\
\mathcal{P}\mathcal{M} &= \mathcal{z}'q \quad \text{with } \mathcal{P}_0^\top \mathcal{M} = \mathcal{z} \text{ and } \mathcal{P}_3\mathcal{M} = \mathcal{z}' \\
\mathcal{M} &= \mathcal{P}_3^\top \mathcal{M}p \\
\mathcal{P}\mathcal{M} &= \mathcal{P}_3^\top \mathcal{M}q
\end{align*}
\]
Triangulation

\[
\begin{align*}
\mathcal{M} & \sim p \\
\mathbf{P}\mathcal{M} & \sim q \\
\mathcal{M} & = zp \\
\mathbf{P}\mathcal{M} & = z'q
\end{align*}
\]

with \( \mathbf{P}_0^\top \mathcal{M} = z \) and \( \mathbf{P}_3\mathcal{M} = z' \)

\[
\begin{align*}
\mathbf{P}_1^\top \mathcal{M} - x\mathbf{P}_3^0\mathcal{M} & = 0 \\
\mathbf{P}_2^\top \mathcal{M} - y\mathbf{P}_3^0\mathcal{M} & = 0 \\
\mathbf{P}_1^\top \mathcal{M} - x'\mathbf{P}_3^\top \mathcal{M} & = 0 \\
\mathbf{P}_2^\top \mathcal{M} - y'\mathbf{P}_3^\top \mathcal{M} & = 0
\end{align*}
\]
Triangulation

\[
\begin{align*}
(P_1^{0\top} - xP_3^{0\top})M &= 0 \\
(P_2^{0\top} - yP_3^{0\top})M &= 0 \\
(P_1^{1\top} - x'P_3^{3\top})M &= 0 \\
(P_2^{2\top} - y'P_3^{3\top})M &= 0
\end{align*}
\]
Triangulation

\[ \begin{align*}
(\mathbf{P}_1^0 - x\mathbf{P}_3^0)\mathbf{M} &= 0 \\
(\mathbf{P}_2^0 - y\mathbf{P}_3^0)\mathbf{M} &= 0 \\
(\mathbf{P}_1^1 - x'\mathbf{P}_3^1)\mathbf{M} &= 0 \\
(\mathbf{P}_2^1 - y'\mathbf{P}_3^1)\mathbf{M} &= 0
\end{align*} \]

- We have a linear system express with $\mathbf{M}$ as unknown
Triangulation

\[
\begin{align*}
(P_1^0 \mathbf{P}_3^0)^\top - xP_3^0 \mathbf{P}_3^0)^\top \mathbf{M} &= 0 \\
(P_2^0 \mathbf{P}_3^0)^\top - yP_3^0 \mathbf{P}_3^0)^\top \mathbf{M} &= 0 \\
(P_1^1 \mathbf{P}_3^1)^\top - x'P_3^1 \mathbf{P}_3^1)^\top \mathbf{M} &= 0 \\
(P_2^2 \mathbf{P}_3^2)^\top - y'P_3^2 \mathbf{P}_3^2)^\top \mathbf{M} &= 0
\end{align*}
\]

- We have a linear system express with \( \mathbf{M} \) as unknown
- You can extend this to \( n \) views
Triangulation

\[
\begin{align*}
(P_1^0 - xP_3^0)M &= 0 \\
(P_2^0 - yP_3^0)M &= 0 \\
(P_1^1 - x'P_3^1)M &= 0 \\
(P_2^2 - y'P_3^2)M &= 0
\end{align*}
\]

- We have a linear system express with \( M \) as unknown
- You can extend this two \( n \) views
- But what it ultimately does is an intersection in 3D which has little to do with observation
Triangulation

Non-linear estimation

You want to optimize the re-projection error

$$\arg \min_M \sum_i \| w (P^i M) - p^i \|^2$$
Triangulation

\[ P^0 = [I_0] \]
Triangulation

Non-linear estimation

• You want to optimize the re-projection error

\[ \arg \min_{\mathcal{M}} \sum_{i} \left\| \mathbf{w} \left( \mathbf{P}^i \mathcal{M} \right) - \mathbf{p}^i \right\|^2 \]

• This is a least square problem
What you know

Image 1
- Feat. Points Extraction

Image 2
- Feat. Points Extraction

Matching

Coord. of Corresp. points

Linear Estimation

Non-linear Estimation

Pose Estimate

Photometric Information

Geometric Information
What you know

Image 1
- Feature Points Extraction
- Photometric Information

Image 2
- Feature Points Extraction

Matching
- Coord. of Corresp. Points

Linear Estimation

Non-linear Estimation
- Pose Estimate

Geometric Information
Pose Estimation

First Image

Target Image

\( p_i \)

\( M_i \)

\( q_i \)
Pose Estimation

Target Image

$M_i$

$q_i$

$P$
Pose Estimation

- Using the same idea as for triangulation you can build a linear system.
Pose Estimation

- Using the same idea as for triangulation you can build a linear system
- Least square related cost function

$$\arg\min_{P} \sum_{j} \|w \cdot (PM_{j}) - p_{j}\|^2$$
Bundle Adjustment

First Image

\[ \mathbf{p_i} \]

Target Image

\[ \mathbf{q_i} \]

\[ \arg \min_{\mathcal{M}_j} \sum_{i,j} \| \mathbf{w}(\mathbf{P}^i \mathcal{M}_j) - \mathbf{p}_j \|^2 \]

\[ \arg \min_{\mathcal{P}} \sum_j \| \mathbf{w}(\mathbf{P} \mathcal{M}_j) - \mathbf{p}_j \|^2 \]
Bundle Adjustment

\[
\arg\min_{p_i, M_j} \sum_{i,j} \| w (P^i M_j) - p^i_{ij} \|^2
\]
What you know

![Diagram of feature point extraction, matching, and estimation]

Photometric Information

\[ \arg \min_{\mathcal{X}_i, T} \sum_i d(p_i, K w(\mathcal{X}_i))^2 + d(p'_i, K' w(T \mathcal{X}_i))^2 \]

Geometric Information
Is Geometric Only a Good Idea?
What can we do more?

Photometric Information and Geometric information in a unified framework
Modification of the cost function

\[ \sum_{j \in \Omega_i} \left( I_1(Kw(Y_{ij})) - I_2(K'w(TY_{ij})) \right)^2 \]

- How can we define the neighborhood of the 2D points in the two images?
- At which stage the photometric term should be used?
- How should the geometric and the photometric terms be weighted?
Figure 3: The Neighbors $\mathcal{N}_j$ (in blue) defined in 3D around triangulated point $\mathcal{X}_i$ (in green), they are projected in the image to create Patch$_1$ and Patch$_1'$ as shown in left and right pictures.
When to activate the Intensity based cost function?

<table>
<thead>
<tr>
<th>#Points/Noise</th>
<th>0.01</th>
<th>0.1</th>
<th>1.0</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.3619</td>
<td>3.9054</td>
<td>96.3384</td>
<td>179.4612</td>
</tr>
<tr>
<td>10</td>
<td>0.2066</td>
<td>1.9684</td>
<td>17.0605</td>
<td>24.6839</td>
</tr>
<tr>
<td>25</td>
<td>0.1002</td>
<td>1.0511</td>
<td>10.4050</td>
<td>18.5587</td>
</tr>
<tr>
<td>50</td>
<td>0.0660</td>
<td>0.6916</td>
<td>6.0568</td>
<td>14.0087</td>
</tr>
</tbody>
</table>

Table 1: Evolution of the mean residual error (in pixels) over 200 runs after applying the 8-points algorithm and the optimal triangulation, with respect to the number of points and the Gaussian noise.
When to activate the Intensity based cost function?

<table>
<thead>
<tr>
<th>#Points/Noise</th>
<th>0.01</th>
<th>0.1</th>
<th>1.0</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.3619</td>
<td>3.9054</td>
<td>96.3384</td>
<td>179.4612</td>
</tr>
<tr>
<td>10</td>
<td>0.2066</td>
<td>1.9684</td>
<td>17.0605</td>
<td>24.6839</td>
</tr>
<tr>
<td>25</td>
<td>0.1002</td>
<td>1.0511</td>
<td>10.4050</td>
<td>18.5587</td>
</tr>
<tr>
<td>50</td>
<td>0.0660</td>
<td>0.6916</td>
<td>6.0568</td>
<td>14.0087</td>
</tr>
</tbody>
</table>

Table 1: Evolution of the mean residual error (in pixels) over 200 runs after applying the 8-points algorithm and the optimal triangulation, with respect to the number of points and the Gaussian noise.

\[ \delta_i = \begin{cases} 
  1 & \text{if } d(p_i, K'w(T\hat{R}_i)) < \tau_1 \text{ and } NCC(Patch_i, Patch'_i) < \tau_2 \\
  0 & \text{otherwise} 
\end{cases} \]
Unified Cost function

\[
\arg\min_{\mathbf{x}_i, \mathbf{T}} \quad \alpha_g \mathcal{D}_g (\mathbf{x}_i, \mathbf{T}) + \alpha_p \mathcal{D}_p (\mathbf{x}_i, \mathbf{T})
\]

\[
\mathcal{D}_g = \sum_i d (\mathbf{p}_i, \mathbf{Kw}(\mathbf{x}_i))^2 + d (\mathbf{p}_i', \mathbf{K'}\mathbf{w}(\mathbf{T}\mathbf{x}_i))^2
\]

\[
\mathcal{D}_p = \sum_i \delta_i \sum_j \left( \mathcal{I}_1 (\mathbf{Kw} (\mathbf{y}_{ij})) - \mathcal{I}_2 (\mathbf{K'}\mathbf{w} (\mathbf{T}\mathbf{y}_{ij})) \right)^2
\]
Qualitative Results
Quantitative Results

Figure 4: Convergence rate in percent; the first graph shows results with an increasing number of points and a Gaussian noise of 0.5; the second graph shows the effect of an increasing distance from the scene with 50 points and a Gaussian noise of 0.5; the last graph shows the results with 50 points and varying Gaussian noise; it shows that the additional intensity information enables our approach to be more robust even with a fronto-parallel assumption (FP).
Conclusion

• Lesson learned
  – Look at what the linear cost function try to solve
  – Think about what you really want to minimize

  – Using both information we achieve
    • Better convergence rate
    • Better precision
    • Seems to work for calibration
Merry Christmas!

pierre.georgel@gmail.com