ShapeFit and ShapeKick for Robust, Scalable Structure from Motion

Location Recovery Problem:
Given: relative directions \( \{v_{ij}\} \) between cameras \( i, j \) (for known camera orientations)
Find: camera locations \( \{t_i\} \)
Difficulty: many directions are outliers

Mathematical Formulation:
Let: \( t_1 \ldots t_n \in \mathbb{R}^3 \)
\( G = ([n], E = E_g \cup E_b) \)
\( v_{ij} = \frac{t_i - t_j}{\|t_i - t_j\|_2} \) for \( i, j \in E_g \)
\( v_{ij} \in S^2 \) for \( i, j \in E_b \)

Given: \( G, \{v_{ij}\} \)
Find: \( \{t_i\} \) up to translation and scale

ShapeFit:
A convex program for location recovery with outliers

\[
\begin{align*}
\text{minimize} & \quad \sum_{ij \in E} \|P_{ij}(t_i - t_j)\|_2 \\
\text{subject to} & \quad \sum_{ij \in E} (t_i - t_j, v_{ij}) = 1, \quad \sum_i t_i = 0
\end{align*}
\]

ShapeKick:
A fast ADMM approach for ShapeFit using kicking

Augmented Lagrangian:
\[
\mathcal{L}_\alpha(T, Y, \lambda) = \sum_{ij \in E} \|P_{ij}(y_{ij})\|_2 + \frac{\tau}{2} \sum_{ij \in E} \|t_i - t_j - y_{ij} + \lambda_{ij}\|^2
\]
\[
\begin{align*}
T & \leftarrow \arg\min_{T} \mathcal{L}_\alpha(T, Y, \lambda) \\
Y & \leftarrow \arg\min_{Y} \mathcal{L}_\alpha(T, Y, \lambda) \\
\lambda_{ij} & \leftarrow \lambda_{ij} + \tau(t_i - t_j - y_{ij})
\end{align*}
\]
Kicking: increase \( \tau \) when convergence slows

ShapeFit: provably robust to outliers

Let: \( t_1 \ldots t_n \sim \mathcal{N}(0, I_{3 \times 3}) \) be i.i.d.
\( ij \in E \) with prob. \( p = \Omega(n^{-1/5} \log^3/5 n) \)
\( E_b \subset E \) be an arbitrary subset
\( v_{ij} \in S^2 \) be arbitrary for \( ij \in E_b \)
\( \gamma = cp^5/\log^3 n \) for some \( c > 0 \)

Theorem
If \( n \) is large enough, and \( \max \deg(E_b) \leq \gamma n \),
then with probability at least \( 1 - \frac{1}{n^2} \),
the minimizer of Shapefit is unique and
exactly equals \( \{t_i\} \) up to translation and scale.

References:


Global SfM Pipelines:

<table>
<thead>
<tr>
<th>Standard Pipeline:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate graph of</td>
</tr>
<tr>
<td>Epipolar Geometries</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Triangle structure-</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bipartite Pipeline:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate graph of</td>
</tr>
<tr>
<td>Epipolar Geometries</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>problem</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>camera locations</td>
</tr>
</tbody>
</table>

ShapeKick: comparable accuracy & 10x faster than state of the art