Calibration and Registration of Consumer-grade RGB-D cameras

Presenter: DENG Teng
2016-06-21
Consumer-grade RGBD sensor

• Simultaneously acquire the appearance and the geometry
  – Structured light sensor
  – Time-of-flight sensor

• Their depth stream contains distortions
  – Originally designed for gesture detection
Existing work - Intrinsic depth calibration

• Direct depth correction, Smisek[1] et al., 2011
  \[ z_{\text{correct}} = z + Z(u, v) \]

• Disparity value correction, Herrera et al.[2], 2012
  \[ Z = \frac{1}{c_1 d + c_0} \]
  \[ d_{\text{correct}} = d + D(u, v) \cdot \exp(\alpha_0 - \alpha_1 d) \]

– Matlab calibration toolbox available

Existing work – Multiple RGB-D registration

• Intrinsic depth calibration cannot guarantee a visually appealing results when using multiple RGB-D
  – Limited to the available positions of checkerboard
  – Still have errors after depth correction and can be obvious on multi-RGBD setup

• RGB-D Registration
  – Local rigid transform, Deng et al.[3], 2014
  – Volumetric calibration + tracking system, Beck et al.[4], 2015


Motivation

optical

geometric

our method
Proposed method

• Local rigid transform: transforming a point \( x \) from camera A into camera B

\[ x' = R_x x + T_x \]

- \( T_x \) denotes the field of translations, and
- \( R_x \) denotes the field of rotations
Proposed method

• The evaluation of the amount of rotation $r := \mathbf{R}_x$ and translation $t := \mathbf{T}_x$
  – Based on the rotation coefficient $r_{ijk}$ and translation coefficient $t_{ijk}$ on vertex $v_{ijk}$
  – From a rectangular regular grid of $n \times m \times o$ vertices within camera A
Proposed method

• A parameter \((u, v, w)\) is computed from the EIGHT corresponding voxel vertices \(v_{ijk}\) which contains \(x\)
  
  – *lerp* and *slerp* operations are used for interpolating translation and rotation, respectively
Computation of transformation coefficients

• Based on correspondences from **checkerboard** observations

• Adaptive scheme to compute coefficients
  – Estimating $r_{ijk}^l$ and $t_{ijk}^l$ from the correspondences from each region $l$
  – Picking the coefficients that have the smallest error
## Quantitative Results

<table>
<thead>
<tr>
<th>Grid Resolution</th>
<th>5x5x5</th>
<th>10x10x10</th>
<th>20x20x20</th>
<th>40x40x40</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reprojection Error (in pixels)</td>
<td>1.3</td>
<td>1.1</td>
<td>0.8</td>
<td>0.8</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Average reprojection error (in pixels). **Red:** Proposed. **Blue:** Single rigid transform.
Manufacturing Param. + Single 3d Transformation

Manufacturing Param. + Proposed Method
Pros & cons

• Pros:
  – Simple algorithm and easy to implement
  – Visually appealing results
  – Fast on depth correction

• Cons:
  – No geometric correction
Q & A
Thank you!