Texture Mapping for 3D Reconstruction with RGB-D Sensor

Yanping Fu, Qingan Yan, Long Yang, Jie Liao, Chunxia Xiao
Motivation

Reconstructing high quality texture models has important significance in areas such as 3D reconstruction, cultural heritage, virtual reality and digital entertainment.
**Problems**

- Due to the noise of depth data, reconstructed 3D models always accompany geometric errors and distortions.
- In camera trajectory estimation, the pose residual would be gradually accumulated and lead to camera drift.
- The timestamp between captured depth frame and color frame is not completely synchronized.
- RGB-D sensors are usually in low resolution, and the color image is also vulnerable to light and motion conditions.
- RGB images from consumer depth cameras typically suffer from optical distortions.
Problems

Ideally, these projected images are photometrically consistent, and thus, combining them produces a high-quality texture map.
Problems

Related Works
- Blending-based methods
- Projection-based methods
- Warping-based methods
Method

We propose a global-to-local correction strategy to compensate for the texture and the geometric misalignment cause by camera pose drift and geometric errors.
**Method**

**Preprocess:** Modify the sparse-sequence fusion method [4] to reconstruct 3D and select texture candidate images by weighting the elements of image clarity ($E_{cla}$), jitter ($E_{jit}$), motion velocity ($E_{vel}$) and viewport overlay ($E_{dif}$).

$$E = E_{jit} + E_{dif} + E_{vel} + E_{cla} + E_{sel}$$

**Texture Image Selection:** To construct high fidelity texture, we select an optimal texture image for each face of the model to avoid the blurring caused by multi-image blending.
Method

- **Global Optimization:** Because both camera pose $T$ and reconstructed $M$ are not absolutely accurate, adjacent faces with different labels usually cannot be completely stitched. We first adjust the camera pose of each texture chart based on the color consistency and geometric consistency between relevant charts.

\[
E(T) = \sum_{l} \sum_{j \in G_l} \sum_{k \in \text{ch}_l} \left( I_l(\Pi(T_lv_k)) - I_j(\Pi(T_jv_k)) \right)^2 + \lambda \sum_{l} \sum_{k \in \text{ch}_l} \left( \varphi(T_lv_k) - D(\Pi(T_lv_k)) \right)^2
\]
Method

- **Local Optimization:** The global optimization can only correct the camera drift of each chart. But the ubiquity of geometry errors makes the only global optimization is insufficient for high fidelity texture mapping. We import an a local adjustment to refine texture coordinates of each vertex on the boundary of chart and make seamlessly stitched textures.

\[
E(T_{ij}) = \sum_{i} \sum_{j} \sum_{k \in A_i} (I_i \left( \prod(T_{ij} T_{ij}) \right) - I_k \left( \prod(T_{kj} T_{kj}) \right))^2 + \lambda \sum_{i} \sum_{j} (T_{ij}^T T_{ij} - I)
\]
Results

The comparisons between the state-of-the-art approaches Waechter et al. [1] (left) Zhou et al. [2] (middle) and ours (right) on several datasets acquired by Kinect.
Results
Results

The performance statistics of Waechter et al. [23], Zhou et al. [30] and our algorithm.

<table>
<thead>
<tr>
<th>model</th>
<th>points</th>
<th>faces</th>
<th>key frames</th>
<th>$T_s$</th>
<th>$T_g$</th>
<th>$T_l$</th>
<th>$T_t$</th>
<th>[23]</th>
<th>[30]</th>
</tr>
</thead>
<tbody>
<tr>
<td>toy</td>
<td>40705</td>
<td>79682</td>
<td>14</td>
<td>1.087</td>
<td>16.591</td>
<td>4.446</td>
<td>27.608</td>
<td>147.068</td>
<td>341.755</td>
</tr>
<tr>
<td>book</td>
<td>178584</td>
<td>352510</td>
<td>16</td>
<td>2.262</td>
<td>84.792</td>
<td>57.053</td>
<td>157.014</td>
<td>486.649</td>
<td>902.207</td>
</tr>
<tr>
<td>hat</td>
<td>70623</td>
<td>137767</td>
<td>10</td>
<td>2.475</td>
<td>29.193</td>
<td>13.263</td>
<td>59.949</td>
<td>214.774</td>
<td>1002.190</td>
</tr>
<tr>
<td>keyboard</td>
<td>68238</td>
<td>134475</td>
<td>13</td>
<td>4.223</td>
<td>26.478</td>
<td>7.594</td>
<td>47.637</td>
<td>321.973</td>
<td>1513.080</td>
</tr>
</tbody>
</table>
Limitations

◆ The texture may be stretched and shrunk on the boundary of charts.

◆ When geometric error is large, the correction would still generate some local texture distortions to final mapping results.


Q&A

个人邮箱：ypfu@whu.edu.cn

课题组主页：http://graphvision.whu.edu.cn/