A Novel Self-Calibration Method for a Stereo-ToF System Using a Kinect V2 and Two 4K GoPro Cameras

Yuan Gao*, Sandro Esquivel, Reinhard Koch, Joachim Keinert

Motivation

- The Kinect V2 is impossible to be fixed due to the changeable tilt, which affects the extrinsic parameters;
- Little literature focuses on the stereo- and ToF-depth fusion problem in 4K resolution.

Depth Correction

Analysis of [1] has shown that typically the Kinect V2 has a depth bias of -18 mm. Correcting for the bias improves the result.


Experiments

- Table I. The RMSE results of the self-calibration.
- Table II. The performances of different depth sources for the rendering of virtual views using different evaluation metrics.

Self-Calibration

- Fig. 1. The multi-camera rig.
- Fig. 2. A virtual vertical view of the rig.

Depth Fusion

In the camera image space of $C_A$:
- $D_1$ is projected from the Kinect V2 sensor;
- $D_2$ is the depth result of using the stereo matching method.

$$\begin{align*}
D'_2(i, j) &= \begin{cases} 
D_2(i, j), & \text{if } D_2(i, j) > 0; \\
D_1(i, j), & \text{else}. 
\end{cases}
\end{align*}$$

Notation

- A 2D point of $C_A$: $[x_A^v y_A^v 1]^T$
- A 2D point of $C_C$: $[x_C^v y_C^v 1]^T$
- Camera intrinsic parameters: $K_a, K_i$
- A 2D point of $C_A$ on the normalized image plane: $p = \left[ -\frac{w}{f}x^v, \frac{w}{f}y^v, 1 \right]^T$
- $R_c = R_c^w [R_c^e]^{T}$
- $t_f = t_f^w - R_c^e t_c$

Experiments

- Fig. 5. The manually-labeled points in $C_A$ and $C_C$.
- Fig. 6. The point clouds of a part of the scene in the camera 3D space of $C_A$.
- Fig. 7. Projected virtual color views on the camera image plane of $C_A$.

E-mail: yga@informatik.uni-kiel.de

International Conference on 3D Vision, October 10th-12th 2017, Qingdao, China