Adding Synchronization and Rolling Shutter in Multi-Camera Bundle Adjustment

Thanh-Tin Nguyen

Maxime Lhuillier
http://maxime.lhuillier.free.fr

Institut Pascal
CNRS UMR 6602, Université Blaise
Pascal, IFMA
Aubière, FR

This paper introduces a new bundle adjustment for multi-cameras, that simultaneously estimates not only the usual parameters (poses and points) but also the synchronization and the rolling shutter. We start from an initial calibration with a global shutter camera model (GS) and a frame-accurate synchronization provided by previous self-calibration methods [1]. Our BA provides subframe-accurate synchronization (SFA), *i.e.* it estimates the residual time offsets Δ_j between a reference camera and the others. It also estimate the rolling shutter (RS) coefficient, *i.e.* the time delay τ between two adjacent lines of a frame.

We present a continuous camera trajectory model that provides the multi-camera pose at every time. Let $\mathcal{R}$ be a $\mathcal{C}^1$ continuous function that maps $\Omega \subseteq \mathbb{R}^k$ to rotation set $SO(3)$. The camera trajectory is modeled by a $\mathcal{C}^3$ continuous function $M : [0,1] \to \mathbb{R}^3 \times \Omega$ such that $M(t)^T = \begin{pmatrix} T_M(t)^T & E_M(t)^T \end{pmatrix}$ and $T_M(t) \in \mathbb{R}^3$ is the translation and $\mathcal{R}(E_M(t)) \in SO(3)$ is the rotation. We approximate M at every time t by using M at few times $t_i \in [0,1]$ corresponding to the key-frames provided by standard GS-multi-camera structure-from-motion: $M(t)$ is a linear combination of three $M(t_i)$ thanks to Taylor approximations and by neglecting their remainders. The y-th line of the j-th camera in the i-th keyframe is taken at time $t = t_i + \Delta_j + y\tau$.

Following [2], a minimal parametrization $\mathcal{R}$ is preferred to avoid any constraints on the $\mathcal{R}$ entry and limit the number of estimated parameters: $k = 3$. However, $\mathcal{R}$ is a global parametrization of all rotations of the continuous camera motion and $\mathcal{R}$ has singularities like every 3D parametrization of $SO(3)$. We propose to define $\mathcal{R}$ using a careful use of the Euler parametrization and keep away from the Euler singularities thanks to an assumption on the camera (all yaw angles are possible, but the pitch and roll are small). This assumption is reasonable for an user exploring the environment without a special objective like grasping at objects on the ground.
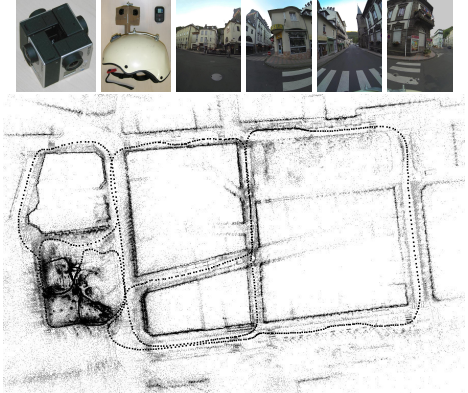
We experiment in cases that we believe use-

Figure 1: Helmet-held multi-camera formed by four GoPro Hero3 cameras, images taken at a viewpoint, reconstruction of a 900m long video sequence (walking in a town) by RS-SFA bundle adjustment without loop closure.

ful: several and identical consumer cameras mounted on a helmet under varying conditions: bike riding, walking, and flying. Ground truth is available for τ (using a strobe) and Δ_j (synthetic videos). At first glance, our approximations seem hazardous if the user does a motion that is not consistent with the neighboring keyframes. Anyway, the majority of keyframes provides accurate enough approximation to obtain the following results in our non trivial datasets. The relative error of the estimated line delay is less than 7.9% except in the most difficult case (flying) with faster head motions; the simultaneous estimation of line delay and time offsets can provide bias but it also provides the best result (5.1%) for the most difficult case. The best time offsets are given by the simultaneous estimation.

[1] M. Lhuillier and T.T. Nguyen. Synchronization and self-calibration for helmet-held consumer cameras, applications to immersive 3d modeling and 360 videos. In *3DV'15*.

[2] L. Oth, P. Furgale, L. Kneip, and R. Siegwart. Rolling shutter camera calibration. In *CVPR'13*.