Supplementary Material: Real-time 6-DoF Pose Estimation by an Event-based Camera using Active LED Markers

Anonymous WACV Applications Track submission

Paper ID 282

Abstract

This appendix provides extra materials for the paper "Real-time 6-DoF Pose Estimation by an Event-based Camera using Active LED Markers". In particular, we first discuss the kinematic relations and the optimization process for computing the two transformation matrices used for the quantitative analyses in the experimental section. Then, we discuss in more detail the bias adjustment process in Section 4.2 of the main paper. We provide an additional comparison experiment with the ORB-SLAM2 algorithm. Finally, we give more information on the recording data, including videos of aggressive maneuvering of the quadcopter with our proposed event-based camera using active led markers (ALM).

1. Kinematic Relations of the Experimental Setup

For the description of the kinematic relations of the experimental setup, the notation of a homogeneous transformation from the coordinate frame $\mathcal{A}$ to the coordinate frame $\mathcal{B}$

$$\mathbf{H}_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} \mathbf{R}_{\mathcal{A}}^{\mathcal{B}} & \mathbf{d}_{\mathcal{A}}^{\mathcal{B}} \\ 0 & 1 \end{bmatrix}, \tag{1}$$

comprising the displacement $\mathbf{d}_{\mathcal{A}}^{\mathcal{B}} \in \mathbb{R}^3$ and the rotation matrix $\mathbf{R}_{\mathcal{A}}^{\mathcal{B}} \in \mathrm{SO}(3)$, is utilized.

Figure 1 gives an overview of the kinematic relations of the setup. Therein, the coordinate system $\mathcal{W}$ is the base frame of the OptiTrack system, $\mathcal{C}_b$ denotes the frame of the OptiTrack markers forming the rigid body of the event-based camera. Similarly, to detect the pose of the ALMs on the board, OptiTrack markers are attached to each i-th ALM, forming the corresponding rigid body $\mathcal{M}_{b,i}$.

The measurements of the OptiTrack system yield the transformations $\mathbf{H}_{\mathcal{W}}^{\mathcal{C}_b}$ and $\mathbf{H}_{\mathcal{W}}^{\mathcal{M}_{b,i}}$ for the rigid body attached to the camera and the i-th ALM. On the other hand, the ALM pose estimation using the event-based camera yields the transformation $\mathbf{H}_{\mathcal{C}}^{\mathcal{M}_i}$ of the i-th ALM with respect to the event-based camera's optical center $\mathcal{C}$.

To fully determine the kinematic relations in Fig. 1 additional transformations, namely $\mathbf{H}_{\mathcal{C}_b}^{\mathcal{C}}$ and $\mathbf{H}_{\mathcal{M}_{b,i}}^{\mathcal{M}_i}$, are computed using optimization by fitting the closed kinematic chain to the ground truth. To this end, the pose of the i-th marker is expressed as

$$\bar{\mathbf{H}}_{\mathcal{W}}^{\mathcal{M}_i} = \mathbf{H}_{\mathcal{W}}^{\mathcal{C}_b} \hat{\mathbf{H}}_{\mathcal{C}_b}^{\mathcal{C}} \mathbf{H}_{\mathcal{C}}^{\mathcal{M}_i}, \tag{2}$$

also equivalent to

$$\bar{\bar{\mathbf{H}}}_{\mathcal{W}}^{\mathcal{M}_i} = \mathbf{H}_{\mathcal{W}}^{\mathcal{M}b,i} \hat{\mathbf{H}}_{\mathcal{M}_{b,i}}^{\mathcal{M}_i}. \tag{3}$$

Using (2) and (3), the residual transformation reads as

$$\tilde{\mathbf{H}}_i = \left(\bar{\bar{\mathbf{H}}}_{\mathcal{W}}^{\mathcal{M}_i}\right)^{-1} \bar{\mathbf{H}}_{\mathcal{W}}^{\mathcal{M}_i} = \begin{bmatrix} \tilde{\mathbf{R}}_i & \tilde{\mathbf{d}}_i \\ 0 & 1 \end{bmatrix}. \tag{4}$$

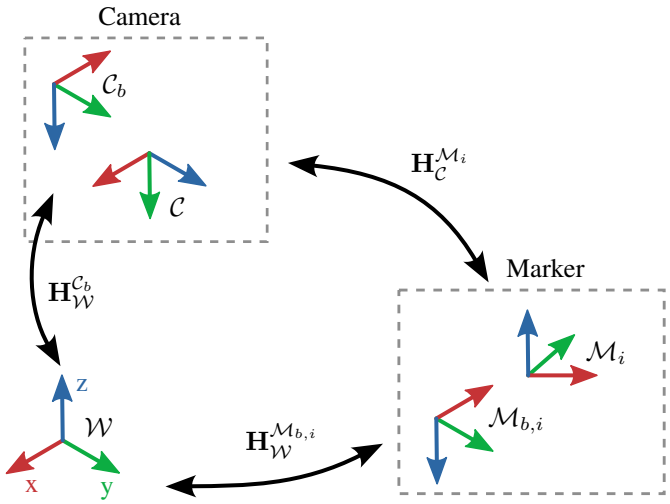

Figure 1. Kinematic relations of the experimental setup. The coordinate system $\mathcal{W}$ indicates the base coordinate system of the OptiTrack system with $\mathcal{C}_b$ and $\mathcal{M}_{b,i}$ designating the OptiTrack markers of the camera base and the marker base i, respectively. The proposed method estimates the pose of the marker coordinate system $\mathcal{M}_i$ with respect to the optical center of the camera $\mathcal{C}$.

(a) Adjusted biases

(b) Default biases

Figure 2. Visualization of bias adjustment in an event camera using the IMX636ES sensor. The vertical axis represents the flattened indices of the pixels in the region of interest (ROI) around an LED light. The left plot demonstrates an optimal bias adjustment, where event fronts - closely packed clusters of events triggered by a rapid change in the scene (like a sudden LED blink) - are clearly distinct. No unwanted spurious events occur between these event fronts. The right plot, in contrast, displays the event distribution with the camera's default bias settings, showing a less distinct separation of event fronts.

With the residual pose (4), the hidden transformations $\{\hat{\mathbf{H}}_{\mathcal{C}_b}^{\mathcal{C}}, \hat{\mathbf{H}}_{\mathcal{M}_{b,1}}^{\mathcal{M}_1}, \hat{\mathbf{H}}_{\mathcal{M}_{b,2}}^{\mathcal{M}_2}\} \in \mathrm{SE}(3)$ for the experimental setup comprising two markers $i \in \{1, 2\}$ can be estimated over $k = 1, \ldots, N$ measurements with the following optimization problem

$$\min \sum_{k=1}^{N} \sum_{i=1}^{2} \left[\left\| \tilde{\mathbf{d}}_{i,k} \right\|_2 + w \left\| \mathbf{E} - \tilde{\mathbf{R}}_{i,k} \right\|_{\mathrm{F}} \right]^2, \tag{5}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. The choice for the metric of the orientation error is motivated in [2] with an empirically chosen weighting value $w = \frac{1}{10}$.

2. Bias Adjustment Procedure

The IMX636ES sensor has five biases [1]. The first two are the contrast sensitivity thresholds *bias_diff_on* and *bias_diff_off* that set the contrast thresholds for on and off events, respectively. The biases *bias_fo* and *bias_hpf* control the bandwidth by setting the illumination signal's low- and highpass filters. Lastly, *bias_refr* sets the refractory period, determining the duration for which the pixel is blind after each event. In the case of the IMX636ES sensor, the biases are expressed as relative offsets from the factory-trimmed default value, resulting in better portability of bias settings between different sensors.

The goal during bias adjustment is to distinguish event fronts and reduce events caused by noise and dynamic changes in the scene. Event fronts are closely packed clusters of events, usually triggered by rapid, high-contrast changes in the scene, like the sudden blink of an LED light. For illustration, refer to Fig. 2 where the pixels of the ROI around a LED are plotted over time. The vertical axis shows the flattened indices $\mathrm{Ind}(\mathbf{u}_j)$ of the pixels $\mathbf{u}_j$ in the ROI. Figure 2a shows a good bias adjustment with clearly distinct event fronts. In between event fronts, no unwanted spurious events occur.

The selection of biases depends on the application. For active markers, the camera's default biases are an adverse choice, as depicted in Fig. 2b. The event plots for the camera's default bias settings is depicted in Fig. 2b.

To adjust the biases, start with the default settings and prepare an online visualization as in Fig. 2. First, increase *bias_fo* until the event rate peaks. Gradually decrease the *bias_hpf* value until a significant drop in the event rate is observed. Increase the *bias_diff_on* and *bias_diff_off* values until the event rate drops significantly. To detect only single polarity events, increase either *bias_diff_on* or *bias_diff_off* to their maximum values. Adjust the *bias_refr* until the event fronts are as narrow as possible. As the bias settings affect the analog front end of the sensor, biases influence each other. For the fine adjustment, look at the visualization and the event rate and alter the settings until the best separation between event fronts is reached while keeping the event rate high. The biases used in the experiments are listed in Tab. 1.

Table 1. List of biases for pose estimation with active LED markers used in the experiments.

Parameter	Value
bias_diff_off	177
bias_diff_on	0
bias_fo	10
bias_hpf	120
bias_refr	0

3. Additional experiment on 6-DoF position estimation for a quadcopter

In the second experiment, the quadcopter is slowly operated in a space of $1.2\,\mathrm{m}$ in the x-direction, $0.3\,\mathrm{m}$ in y-direction, and $0.1\,\mathrm{m}$ in z-direction for about $40\,\mathrm{s}$. The position estimates and the resulting errors for this case are depicted in Fig. 3 and Fig. 4. The translational errors obtained from the proposed method, *i.e.*, $p_{e,x} \in [-0.06, 0.03]\,\mathrm{m}$, $p_{e,y} \in [-0.02, -0.04]\,\mathrm{m}$, and $p_{e,z} \in [-0.02, 0.02]\,\mathrm{m}$, are smaller than the ones obtained from the ORB-SLAM2, *i.e.*, $p_{e,x} \in [-0.1, 0.1]\,\mathrm{m}$, $p_{e,y} \in [-0.06, -0.05]\,\mathrm{m}$, $p_{e,z} \in [-0.04, -0.02]\mathrm{m}$. On the other hand, the rotational errors obtained from the two methods are nearly similar. However, larger measurement spikes are observed with the ORB-SLAM2 method.

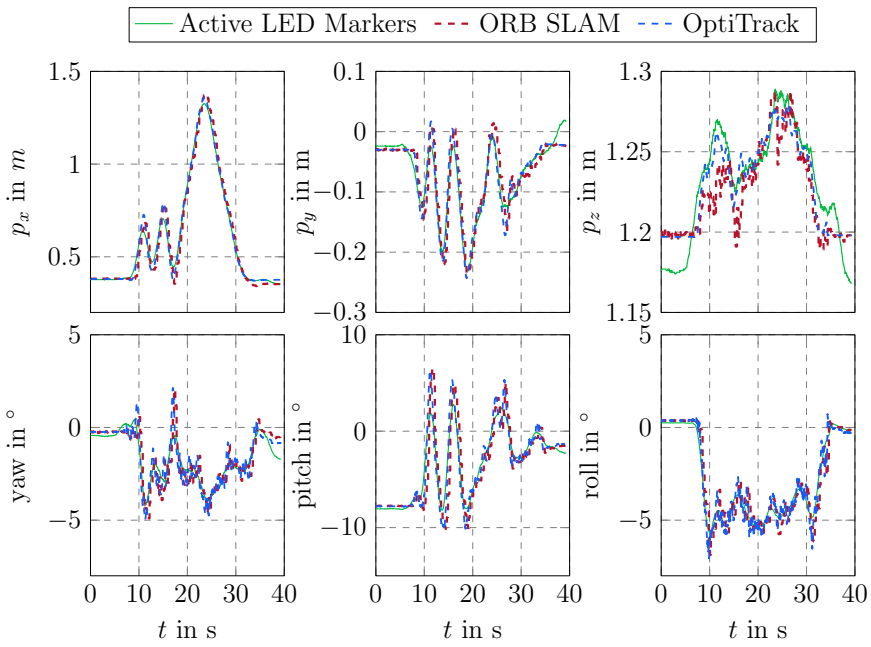

Figure 3. Experiment 2: The estimated poses from the proposed method, the ORB-SLAM2, and the ground truths from OptiTrack are illustrated in green, red, and blue, respectively.

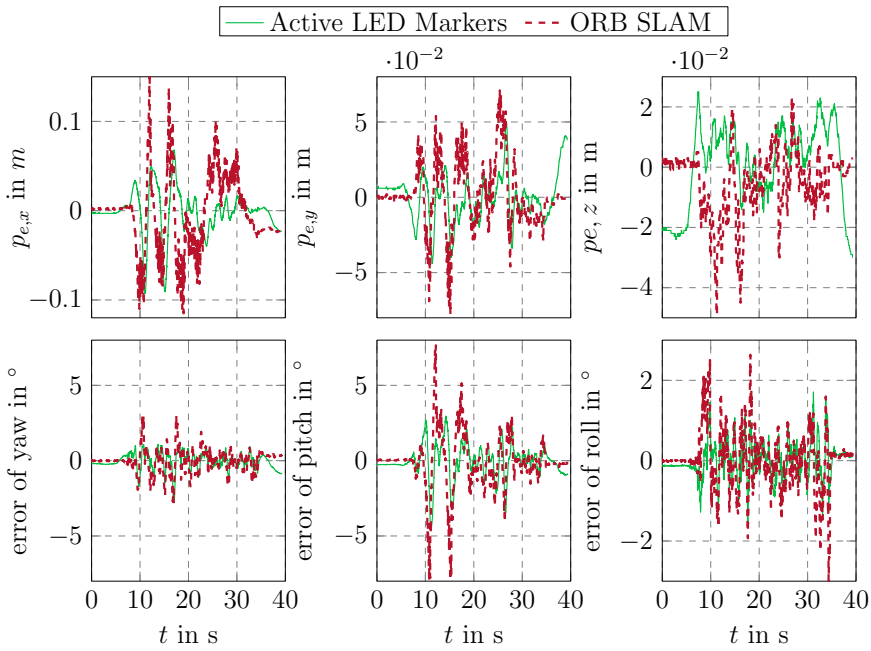

Figure 4. Experiment 2: The error plots of the corresponding estimates, depicted in Fig. 3 with respect to the ground-truth measurements from the OptiTrack.

References

[1] Biases — metavision SDK docs 4.1.0 documentation. 2

[2] Du Q. Huynh. Metrics for 3D Rotations: Comparison and Analysis. *Journal of Mathematical Imaging and Vision*, 35(2):155–164, 2009. 2