Flight Dynamics-based Recovery of a UAV Trajectory using Ground Cameras: Supplementary material

Artem Rozantsev[a] Sudipta N. Sinha[b] Debadeepta Dey[b] Pascal Fua[a]

[a]Computer Vision Laboratory, EPFL [b]Microsoft Research

{artem.rozantsev, pascal.fua}@epfl.ch {sudipta.sinha, dedey}@microsoft.com

In the supplementary material we present more details on one of the baseline methods (**BA-pKF**), which was omitted from the main paper due to the limit on page length. We also report two additional experiments in the supplementary material. In the first experiment, we report a sensitivity analysis regarding parameter settings in our proposed **BA-pDM** method and discuss how this affects the accuracy of the estimated control inputs. In the second experiment, we compare our BA procedure with a baseline where the data association problem was solved before running bundle adjustment by forcing the single-view tracker to output at most one 2D observation in every video frame. Finally, qualitative results on the LAB and FARM datasets are shown in the supplementary video.

1. Kalman filter prior

In Section 5 of the main paper, we described several baselines that we have compared our approach to. Here we provide more details on the method based on the Kalman filter, as the other ones are relatively straightforward to implement.

Classical Kalman filter allows predicting the state of the quadrotor at time $t+1$ from its state at time t. In our case this state contains drone's position and velocity. We then track quadrotor in 3D using the constant acceleration Kalman filter model. Therefore, $\hat{\mathbf{x}}_{t+1}$ from Eq. 4 of the main paper is computed according to the prediction of the Kalman filter. However, as the experiments show, the prior is dependent on the quality of the initialization and **BA-pKF** is not robust to noise in the initial 3D trajectory.

2. Control inputs prediction analysis

In Section 5.1 of the main paper we have shown that our system is capable of inferring the internal state of the quadrotor, which can be further used to estimate the control inputs commanded to the drone by the operator. Fig. 3 of the main paper shows that our approach tends to over smooth the internal parameters $(\mathbf{\Phi}, \mathbf{\Theta}, \mathbf{U})$. Therefore, in this section we investigated the effect of fine-tuning the dynamics-based prior on the accuracy of the estimated $(\mathbf{\Phi}, \mathbf{\Theta}, \mathbf{U})$.

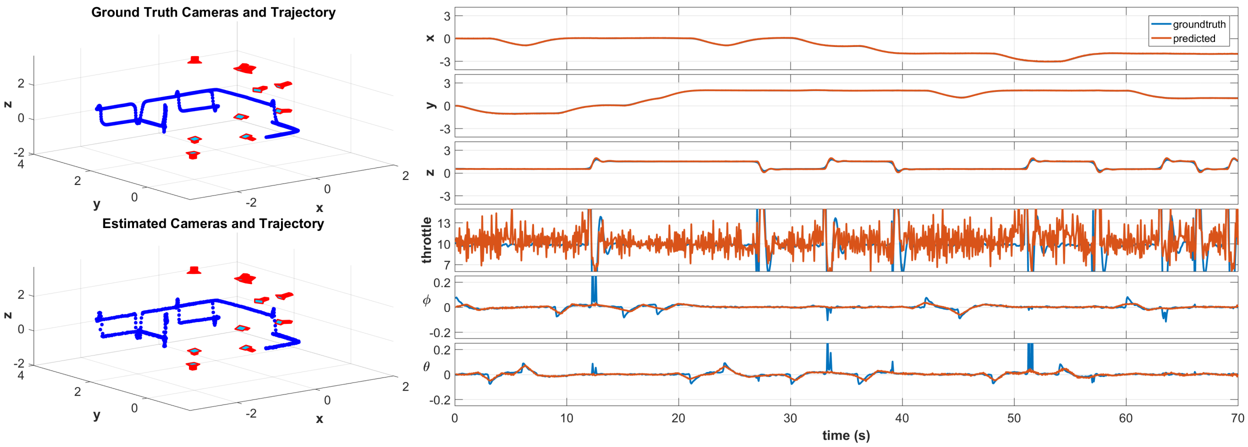
method	Position error (m)
BA-pDM-single	1.998
BA-pDM	1.636

Table 1: Comparison between our method (**BA-pDM**) and a baseline (**BA-pDM-single**) on the FARM dataset. The baseline uses at most one detection in every frame in all the input videos.

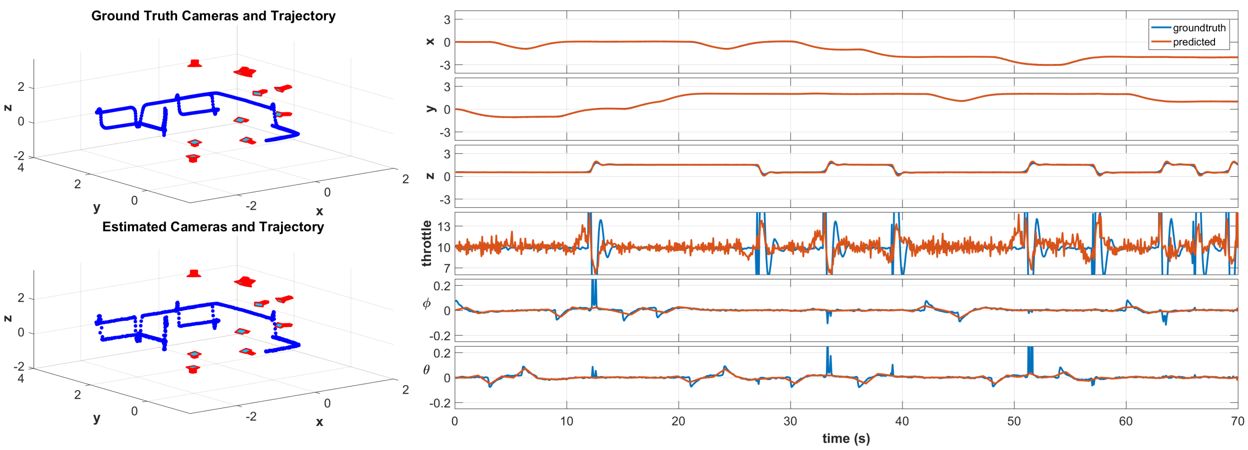
Fig. 1 illustrates the influence of the dynamic-based prior on internal quadrotor state. Here the weight λ from the Eq. 3 of the main paper together with the smoothing coefficient σ of the gaussian kernel $\mathcal{H}(\mathbf{\Gamma}) = (g * \mathbf{\Gamma})$ from Eq. 11 increases from top most plot to the bottom one. We can see in Fig. 1(a) that if the weight of the prior is small (0.01), we can quite reliably recover the sharp peaks of the throttle $(\mathbf{U})$ command, however, there is some residual noise for time periods when the true throttle command remains fixed at a constant value. This happens because the prior does not have enough influence on the optimization to make it robust to the measurement noise in the image. On the other hand increasing the weight of the prior to (0.03) (Fig. 1(c)) allows us to compensate for the initialization noise and recover a smoother estimation of $\mathbf{U}$. However, this tends to oversmooth $\mathbf{U}$ especially when it changes abruptly corresponding to times when the drone suddenly changes direction or altitude.

3. Advantage of the new cost function

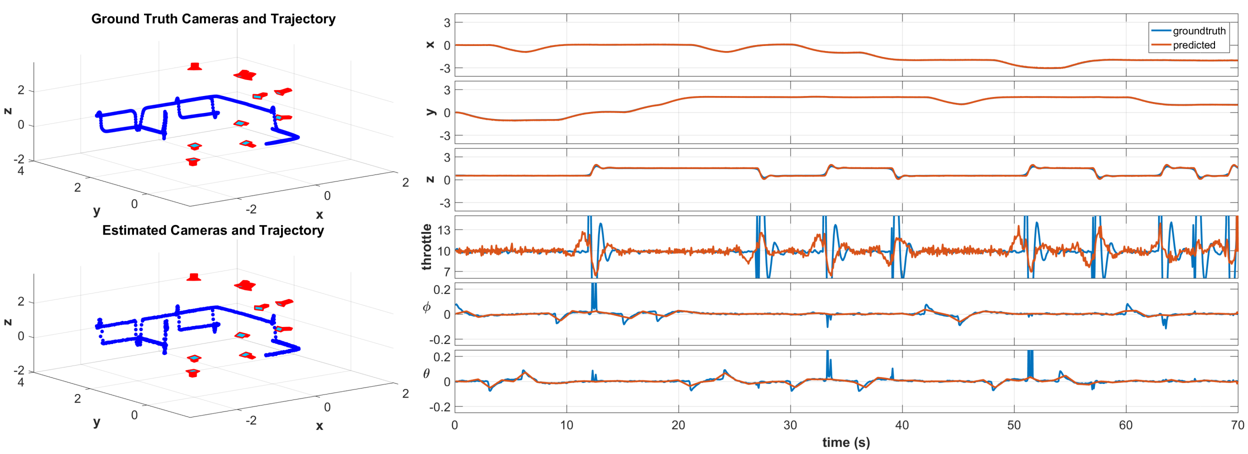
Recall that in Eq. 2 in the main paper, we introduced a cost function based on a more general form of reprojection error that did not rely on perfect correspondence between different views. We did this to handle multiple candidate detections in every frame produced by the single-view



(a) $\lambda = 0.01, \sigma = 0.8$



(b) $\lambda = 0.02, \sigma = 1.1$



(c) $\lambda = 0.03, \sigma = 1.4$

Figure 1: Influence of the dynamics-based prior on the prediction of the internal state of the quadrotor. In different plots we have varied the weight of the prior λ and the smoothing factor σ of the gaussian kernel that smooths $(\mathbf{\Phi}, \mathbf{\Theta}, \mathbf{U})$. (a) corresponds to $(\lambda, \sigma) = (0.01, 0.8)$, (b) illustrates the case when $(\lambda, \sigma) = (0.02, 1.1)$ and (c) depicts the experiment with $(\lambda, \sigma) = (0.03, 1.4)$.

tracker running on the input videos. To quantify the benefit of this new cost function, we compared our method (**BA-pDM**) which uses this new cost function $E(C, X, O)$ with another baseline which we refer to as (**BA-pDM-single**). This baseline uses at most one measurement (detection) in every video frame. In that case, $E(C, X, O)$ becomes identical to the cost function $E_{BA}(C, X, O)$ (see Eq. 1 in the main paper) used in conventional bundle adjustment. These unique detections used in the baseline, were selected during the 3D trajectory initialization step, which uses the RANSAC-based multi-view triangulation method we have proposed.

Table 1 reports the final average position error for 3D points sampled on the trajectories estimated by our method (**BA-pDM**) and by the baseline (**BA-pDM-single**) respectively. These correspond to the FARM dataset. Note that both methods use the same dynamics-based prior but **BA-pDM** produces a more accurate result because the new cost function allows the selection of the 2D measurements (amongst the multiple detection candidates) to be refined during the bundle adjustment procedure.