PAUL: Uncertainty-Guided Partition and Augmentation for Robust Cross-View Geo-Localization under Noisy Correspondence

Supplementary Material

6. Additional Experimental Results and Discussions

6.1. Impact of Partial Correspondence

Solving the "partial correspondence" problem is our core motivation. Beyond the baseline failure shown in Figure 2 in the main text, our new analysis in Figure 11 demonstrates that partial correspondence is valuable: discarding it leads to performance degradation. Unlike baseline methods, PAUL effectively utilizes these clues instead of simply filtering out noise, thereby transforming partial correspondence into a performance improvement.

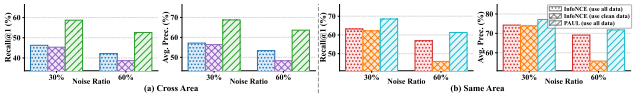

Figure 11. Effectiveness of Partial Correspondence.

6.2. Model Generalization and Robustness to Extreme Noise

To validate the model's generalization ability, we added cross-dataset evaluation, i.e., training the model on the GTA-UAV dataset and testing it directly on the UAV-VisLoc dataset. The results in Table 4 show that PAUL outperforms the baseline model, confirming its strong robustness on unseen data.

Furthermore, regarding the model's performance at a 60% noise ratio, we observed a drawback in the "consistency-driven" approach due to its strong regularization. While this regularization suppresses extreme noise, it also limits the model's generalization ability. As our cross-dataset results (Table 4) show, methods like CREAM exhibit significant performance degradation in unseen domains, indicating overfitting to specific noise distributions. In contrast, PAUL avoids this overfitting, trading improved performance at a 60% noise level for superior stability and generalization across a wider range of scenarios.

Table 4. Cross-dataset Evaluation (GTA-UAV→UAV-VisLoc)

Method	R@1	R@5	R@10	R@top1	AP	Thpt (img/s)
GSC	21.87$_{\pm 1.02}$	41.33$_{\pm 0.85}$	50.13$_{\pm 0.92}$	82.67$_{\pm 0.76}$	31.35$_{\pm 1.04}$	217.53$_{\pm 3.73}$
RCL	18.13$_{\pm 0.86}$	40.27$_{\pm 1.22}$	46.40$_{\pm 0.87}$	81.33$_{\pm 0.97}$	27.97$_{\pm 1.13}$	216.93$_{\pm 1.64}$
CREAM	13.33$_{\pm 0.96}$	27.47$_{\pm 0.64}$	34.67$_{\pm 1.05}$	67.73$_{\pm 1.21}$	20.27$_{\pm 0.53}$	224.67$_{\pm 17.04}$
PAUL	**22.40**$_{\pm 0.93}$	**43.07**$_{\pm 0.85}$	**53.20**$_{\pm 0.69}$	**84.93**$_{\pm 0.84}$	**39.61**$_{\pm 0.82}$	**214.44**$_{\pm 10.21}$

6.3. Insights into Ground-Satellite Cross-View Geo-Localization

While our primary focus is on UAV-Satellite Cross-View Geo-Location (UCVGL), as the lightweight GPS and unconstrained motion of UAVs naturally lead to GPS drift and unique "partial mismatch" characteristics, we also explore its impact on Ground-Satellite Cross-View Geo-Localization (GCVGL). Ground vehicles typically use high precision sensors, thus the likelihood of GPS drift is much lower. Furthermore, due to the large view difference (panoramic vs. top-down view), misalignment in GCVGL can lead to "complete mismatch," in which general purpose noise correspondence methods are generally more suitable than the PAUL method, specifically designed for NC-CVGL.

Supplementary experiments on the VIGOR dataset confirm this (Table 5): due to the large difference hindering accurate alignment, PAUL's advantage in GCVGL is less significant than in UCVGL. However, we observe an interesting "asymmetric processing" phenomenon (Fig 12): augmented data by removing ground queries significantly improves performance. This indicates that specific strategies can mitigate interference caused by view gaps, providing a new approach to the noise correspondence problem in GCVGL.

Table 5. Performance Comparisons on VIGOR(GCVGL dataset).

Noise	Method	R@1	R@5	R@10	R@top1	AP
30%	InfoNCE	37.44$_{\pm 0.60}$	75.59$_{\pm 1.33}$	85.18$_{\pm 1.88}$	**99.80**$_{\pm 0.14}$	52.96$_{\pm 0.52}$
	GSC	39.10$_{\pm 1.86}$	75.50$_{\pm 1.65}$	86.39$_{\pm 1.84}$	99.02$_{\pm 0.08}$	55.42$_{\pm 1.38}$
	CREAM	38.70$_{\pm 0.95}$	76.35$_{\pm 1.38}$	86.39$_{\pm 1.35}$	99.48$_{\pm 0.09}$	55.10$_{\pm 1.29}$
	PAUL	39.36$_{\pm 0.93}$	**81.69**$_{\pm 1.06}$	**92.66**$_{\pm 1.54}$	99.01$_{\pm 0.18}$	56.37$_{\pm 1.33}$
	PAUL (Asys Aug)	**41.27**$_{\pm 1.02}$	80.97$_{\pm 1.15}$	90.16$_{\pm 1.25}$	99.60$_{\pm 0.15}$	**57.43**$_{\pm 1.09}$
60%	InfoNCE	26.35$_{\pm 0.87}$	65.40$_{\pm 0.87}$	78.11$_{\pm 1.29}$	99.73$_{\pm 0.22}$	41.50$_{\pm 0.99}$
	GSC	27.00$_{\pm 0.66}$	65.15$_{\pm 0.95}$	78.34$_{\pm 1.38}$	99.33$_{\pm 0.34}$	42.55$_{\pm 0.91}$
	CREAM	26.60$_{\pm 1.23}$	65.55$_{\pm 0.69}$	78.33$_{\pm 0.89}$	99.41$_{\pm 0.13}$	42.25$_{\pm 1.25}$
	PAUL	27.44$_{\pm 1.02}$	69.81$_{\pm 0.85}$	83.92$_{\pm 0.68}$	99.83$_{\pm 0.12}$	44.30$_{\pm 0.87}$
	PAUL (Asys Aug)	**29.94**$_{\pm 0.85}$	**72.50**$_{\pm 0.76}$	**85.96**$_{\pm 0.81}$	**99.86**$_{\pm 0.09}$	**46.54**$_{\pm 0.76}$

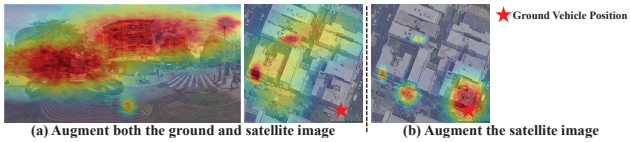

Figure 12. Visualizing Asymmetric Processing.

6.4. Orientation Error and Feature Alignment

Our model learns rotation invariance, a property that is inherited by the saliency maps it derives. This is confirmed by the stable high-response areas shown in Fig. 13. Likewise,

the Grad-CAM visualizations provided in the main text further verify the proper alignment of consistent key regions across the UAV and Satellite views.

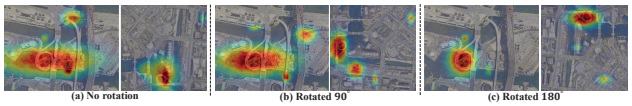

Figure 13. Grad-CAM under Rotations.