

Supplementary material for: Depth Attention for Robust RGB Tracking

Yu Liu¹[0009-0009-5898-0113], Arif Mahmood²[0000-0001-5986-9876], and
Muhammad Haris Khan³[0000-0001-9746-276X]

¹ Xinjiang University

7501847851y@gmail.com

² Information Technology University

arif.mahmood@itu.edu.pk

³ Mohamed bin Zayed University of Artificial Intelligence

muhammad.haris@mbzuai.ac.ae

More Attributes Analysis: We present the analysis of additional attributes in Figure 1, 2, and 3. It is evident that the proposed method enhances performance by addressing challenges such as camera motion, tiny targets, and fast motion. Moreover, it maintains performance in scenarios involving low discriminative features, deformation, and out-of-view situations.

Precise Depth is not necessary: The proposed Depth Attention model doesn't heavily depend on precise depth information. To illustrate this, we conducted a comparison between estimated depth and precise depth using a sample from the CDTB dataset [1], which is an RGBD depth dataset. The results of this comparison are depicted in Figure 4.

References

1. Lukezic, A., Kart, U., Kapyla, J., Durmush, A., Kamarainen, J.K., Matas, J., Kristan, M.: Cdtb: A color and depth visual object tracking dataset and benchmark. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 10013-10022 (2019)

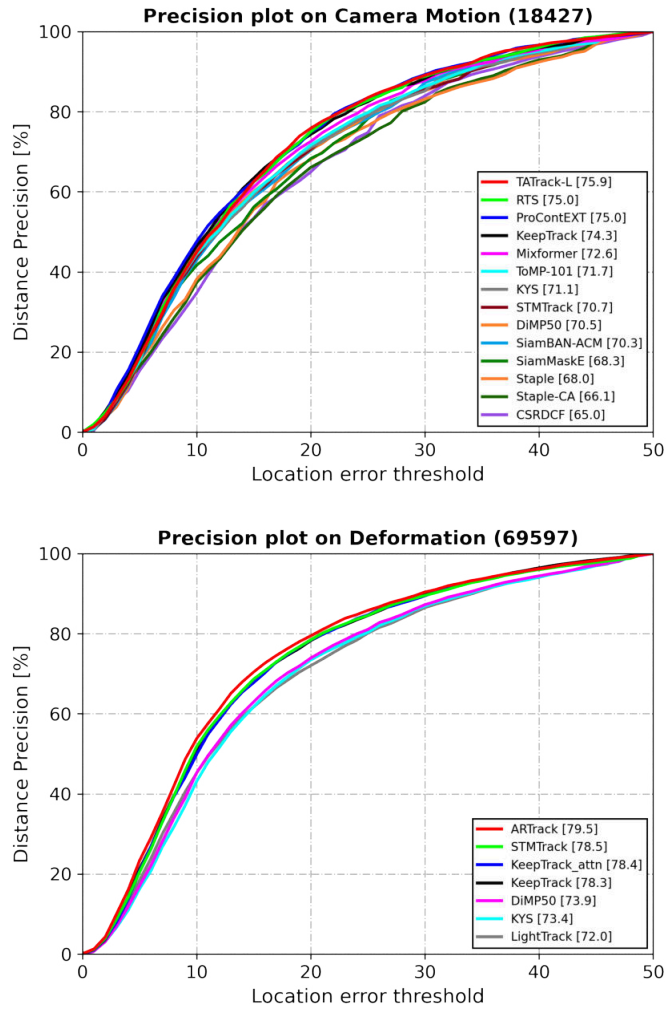


Fig. 1. Additional attributes analysis on camera motion and deformation

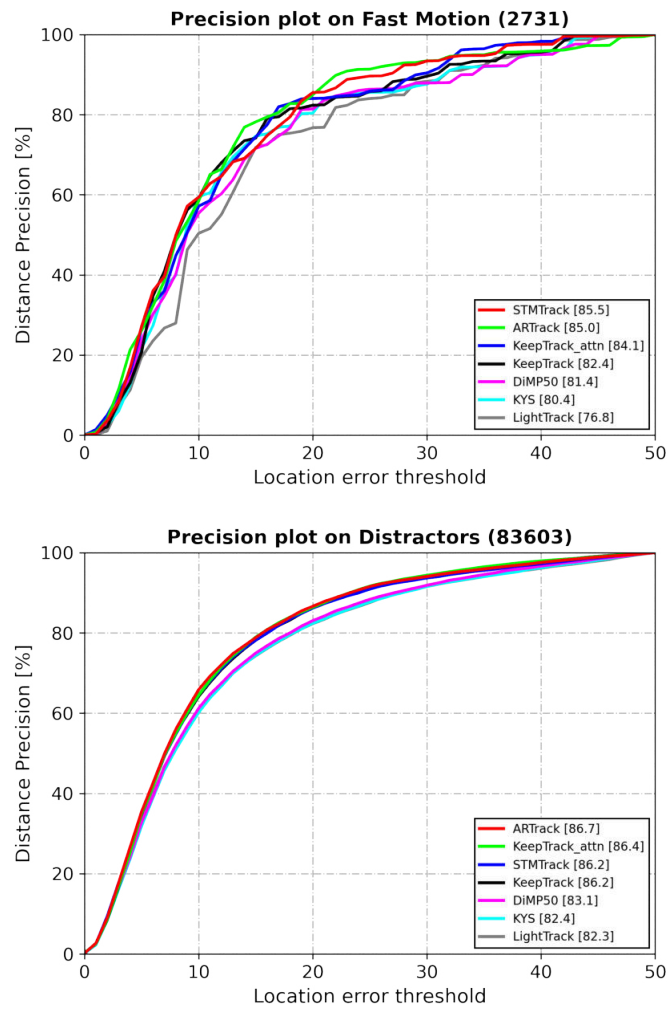


Fig. 2. Additional attributes analysis on fast motion and distractors

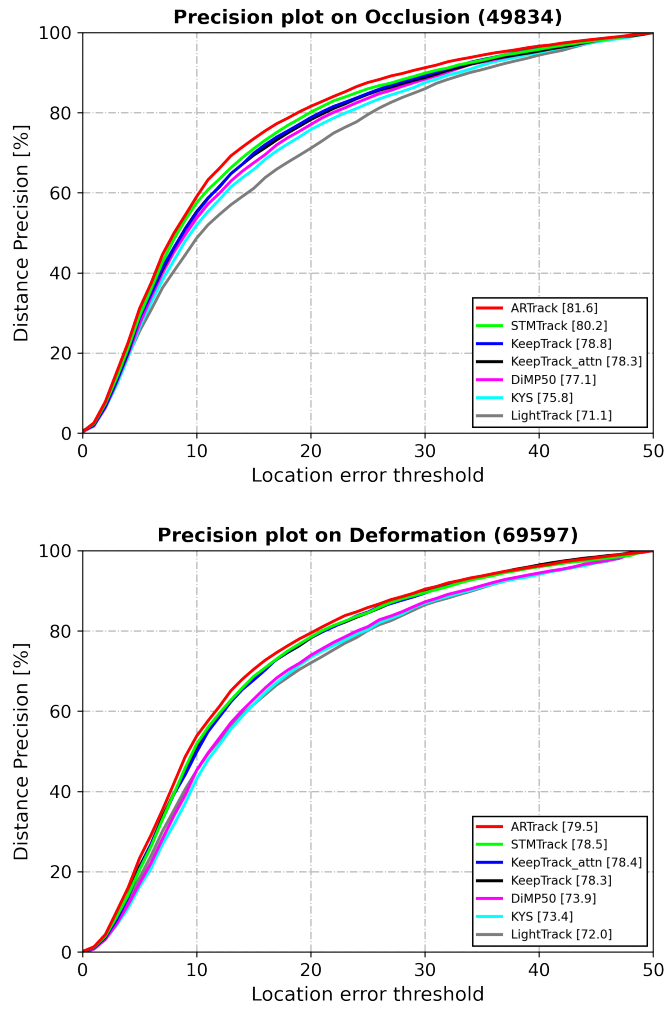


Fig. 3. Additional attributes analysis on occlusion and deformation

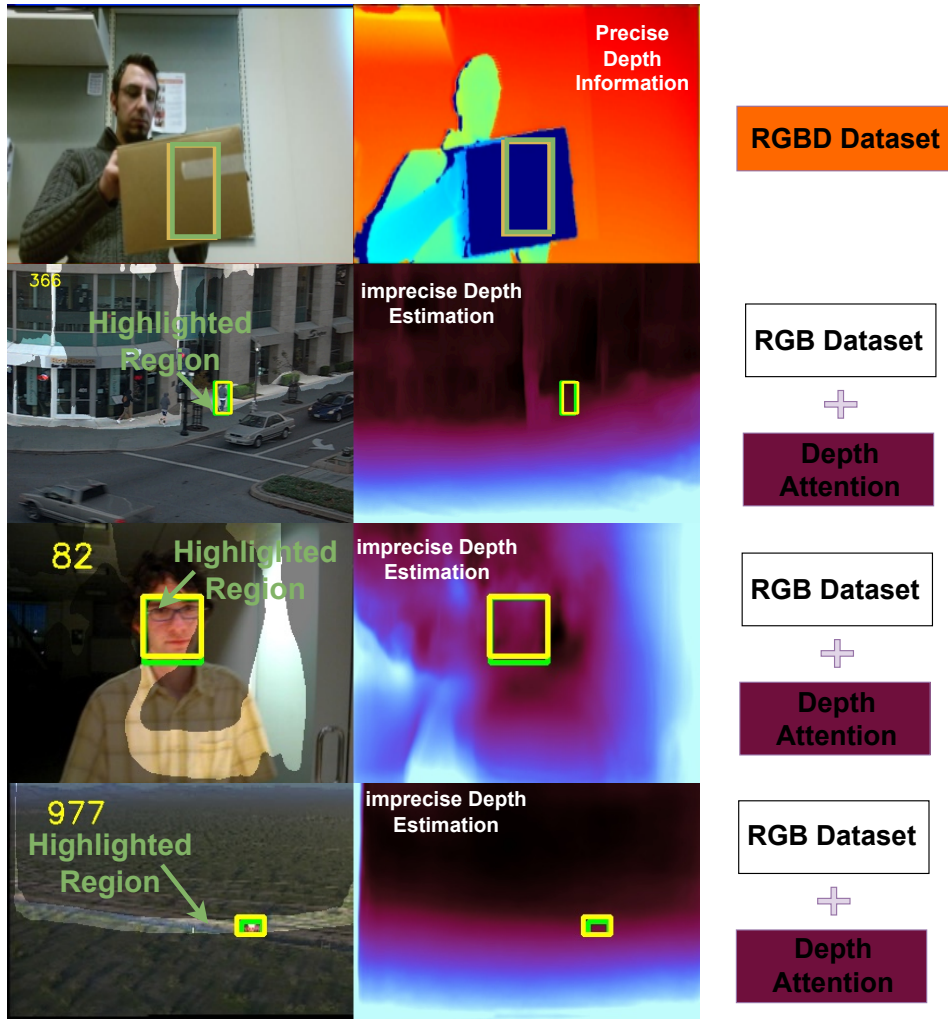


Fig. 4. The proposed method is not reliant on the precise depth information