In Defense of Color-based Model-free Tracking

Horst Possegger, Thomas Mauthner, Horst Bischof
Institute for Computer Graphics and Vision, Graz University of Technology

We address the problem of model-free online object tracking based on color representations. According to recent benchmark evaluations, such trackers often tend to drift towards regions which exhibit a similar appearance compared to the object of interest. To overcome this limitation, we propose an efficient discriminative object model which allows us to identify potentially distracting regions in advance. We exploit this knowledge to adapt the object representation beforehand so that distractors are suppressed and the risk of drifting is significantly reduced. In addition to its favorable simplicity, this representation also enables accurate scale estimation, as shown in Figure 1.

Combining both models, we obtain the final distractor-aware object model as

\[ P(x \in O|b_x) = \lambda_P P(x \in O|D,b_x) + (1 - \lambda_P) P(x \in O|S,b_x), \]

with the pre-defined weighting parameter \( \lambda_P \).

Similar to recent state-of-the-art trackers, we first localize the object in a new frame and subsequently perform scale estimation. Following the tracking-by-detection principle, we propose an iterative non-maximum suppression strategy to localize the object. Using similarity scores based on visual appearance \( s_v(\cdot) \) and Euclidean distance \( s_d(\cdot) \), we find the current object location as \( O^*_t = \text{argmax} \{ s_v(O_t), s_d(O_t) \} \). Additionally, our strategy yields potentially distracting regions \( D \) which are used to update the object model and thus reduce the risk of drifting. After localization, we segment the object via adaptive thresholding on \( P(x \in O|b_x) \). This allows us to accurately estimate the object scale.

We evaluate our distractor-aware tracking approach on the recent Visual Object Tracking (VOT) benchmark datasets [4], VOT13 and VOT14. These datasets cover many challenging real-world scenarios, allowing us to both perform a fair comparison to the state-of-the-art and draw valid conclusions. Our evaluations show that the discriminative color model without distractor-awareness (noDAT) already achieves state-of-the-art accuracy, see Figure 2. By incorporating the distractor-aware representation (DAT), we can significantly increase the tracking robustness. Moreover, our scale-adaptive version (DAT+scale) outperforms the VOT challenge winner DSST [1]. Overall, our approach performs favorably compared with state-of-the-art approaches and enables robust online object tracking in real-time.

---