Wen-jing KANG, Chang LIU, Gong-liang LIU, 2020. A quantitative attribute-based benchmark methodology for single-target visual tracking. *Frontiers of Information Technology & Electronic Engineering*, 21(3):405-421.
https://doi.org/10.1631/FITEE.1900245

A quantitative attribute-based benchmark methodology for single-target visual tracking

Key words: Visual tracking; Performance evaluation; Visual attributes; Computer vision

Corresponding author: Gong-liang LIU
E-mail: liugl@hit.edu.cn
ORCID: Wen-jing KANG, https://orcid.org/0000-0002-7779-0106; Gong-liang LIU, https://orcid.org/0000-0001-7534-4201

Motivation

- Visual tracking is one of the most popular research topics in the field of computer vision. The performance evaluation benchmark of tracking algorithms is particularly important.
- Current tracker evaluation systems focus mainly on the overall performance, or the average value of the whole test dataset. This cannot clearly demonstrate the advantages and disadvantages of the tracker in specific scenarios.
- Moreover, the selection of benchmark measurement lacks convincing explanation.
- As modern algorithms become increasingly powerful, a more detailed evaluation system that meets the need of choosing algorithms based on specific applications is needed.

Main idea

- Visual attributes exist in every frame, and only the intensities are different.
- Different measures seem to reflect different aspects of the tracker's performance; however, some of them are related to each other.
- Find out the influence of each visual attribute on the performance of different trackers (measures showed).

Method

- Inter-frame changes in the video are split into several aspects, which are quantified, normalized, and calculated as 12 quantitative attributes, including illumination change (IL), size change (SC), absolute size change (AS), target movement length (ML), background clutter (BC), camera movement (CM), blur (BL), aspect-ratio change (AR), color change (CC), deformation (DE), scene complexity (SE), and absolute movement (AM).
- The correlation- and weight-based analyses are performed to find out how each attribute impacts the performance of the trackers.
- Seven most representative measures are divided into three categories according to their correlations to describe the accuracy or robustness of the tracker.

Major results

- The measures are clustered clearly into three groups.

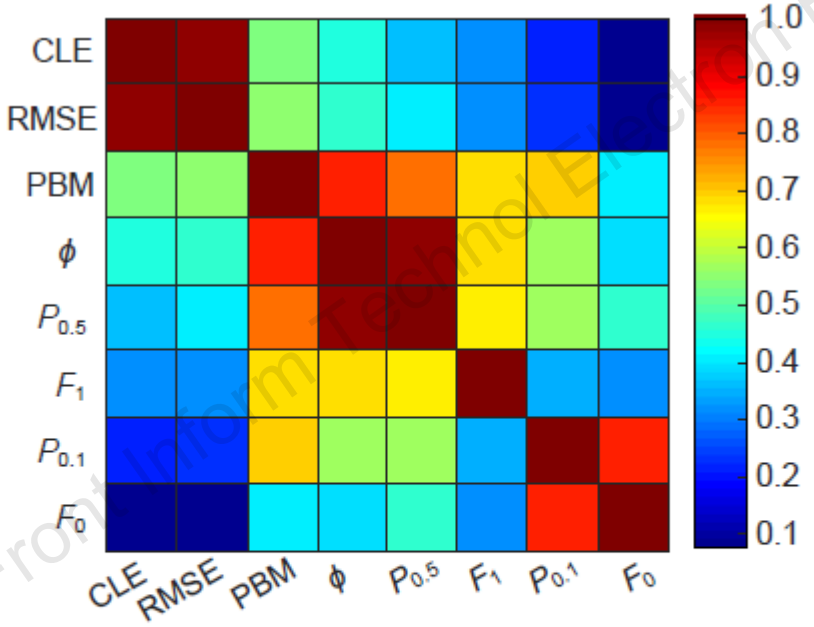

Fig. 7 Correlation matrix for seven measures

CLE: center location error; RMSE: root-mean-square error; PBM: position-based measure; ϕ: average overlap; $P_{0.5}$: success percentage when τ=0.5; F_1: F1-score; $P_{0.1}$: success percentage when τ=0.1; F_0: failure rate. References to color refer to the online version of this figure

Major results (Cont'd)

- SC, ML, BC, BL, AR, SE, and AM are correlated highly with the trackers' performance by influencing at least one measure.
- The ML and AM attributes indicate the tracker's performance in tracking fast-moving objects.
- CLE is much more sensitive than overlap to the SC attribute.
- For the scale-related attributes, the performance gap between deep and traditional trackers is not large.

Conclusions

- The proposed two attribute-based benchmarking methods can provide more detailed information on trackers' performance.
- The current bottleneck in visual tracking lies in the scale-adaptive part of the process.
- The impact of a target's moving speed cannot be ignored either.
- CLE can better indicate the influence of a rapid change of the target size. This shows that we cannot rely on only a simple theoretical conjecture to conclude the characteristics of a test measure or a methodology.