

MOTCOM: The Multi-Object Tracking Dataset

Complexity Metric

Supplementary Material

Malte Pedersen¹, Joakim Bruslund Haurum^{1,2}, Patrick Dendorfer³, and
Thomas B. Moeslund^{1,2}

¹ Aalborg University, Denmark

² Pioneer Center for AI, Denmark

³ Technical University of Munich, Germany

Computing MOTCOM. We have evaluated four averaging methods for combining the sub-metrics into MOTCOM. The four methods are the arithmetic, quadratic, geometric, and harmonic means and they are presented in Equation (1), Equation (2), Equation (3), and Equation (4), respectively.

$$\text{arithmetic} = \frac{1}{n} \sum_{i=1}^n m_i \quad (1) \qquad \text{quadratic} = \sqrt{\frac{1}{n} \sum_{i=1}^n m_i^2} \quad (2)$$

$$\text{geometric} = \sqrt[n]{\prod_{i=1}^n m_i} \quad (3) \qquad \text{harmonic} = \frac{n}{\sum_{i=1}^n \frac{1}{m_i}} \quad (4)$$

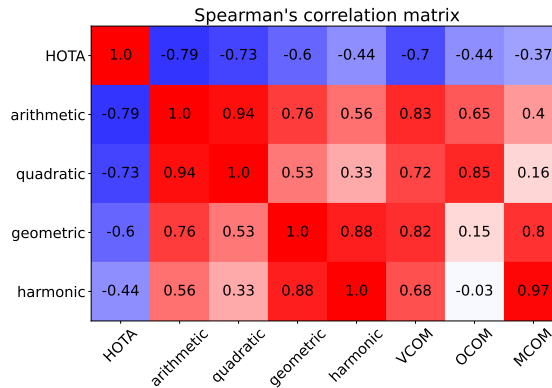
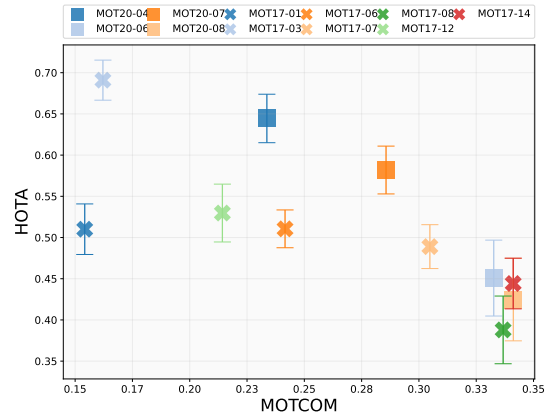


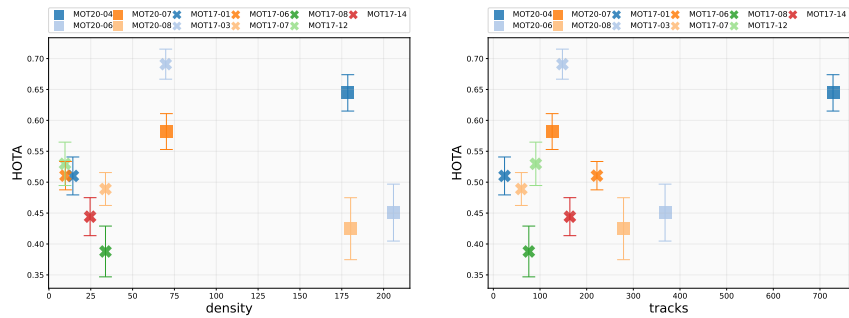
Fig. 1: Spearman's correlation matrix. The entries represent the MOTCOM values when the sub-metrics are combined using the four different averaging methods. The HOTA performance is the average of the top-30 ranked trackers. The scores are based on the combined MOT17 and MOT20 test splits.

We present the four variations of MOTCOM in Figure 1 computed on the combined MOT17 and MOT20 test splits. We see that they all correlate negatively with the HOTA score. However, the arithmetic mean has the strongest negative correlation and it correlates positively with all the sub-metrics. Therefore, we suggest to compute MOTCOM as the arithmetic mean of the sub-metrics.

Complexity Score Plots for MOT17 and MOT20 In the main paper we evaluate MOTCOM, *density*, and *tracks* on the MOT17 and MOT20 test splits. We focus mainly on the ranking capabilities of the metrics as we expect tracker performance to have a monotonic, but not necessarily linear, relationship with complexity. The ranks of MOTCOM, *density*, and *tracks* presented in the main paper are based on the scores displayed in Figure 2.



(a) MOTCOM vs. HOTA.



(b) Density vs. HOTA.

(c) Tracks vs. HOTA.

Fig. 2: Average HOTA performance of the top-30 trackers on MOT17 and MOT20 test split against a) MOTCOM, b) *density*, and c) *tracks*. Square markers represent MOT20 sequences and crosses are MOT17 sequences.

The position of the marker indicates the average score and the error bar is the standard deviation. The marker of the MOT17 sequences is a cross and the MOT20 sequences are represented by a square. We see that the MOT20 sequences have significantly higher densities and more tracks compared to the MOT17 sequences, but the HOTA performance is not correspondingly low. This illustrates that *density* and *tracks* do not suffice to describe the complexity of MOT sequences.

Complete Spearman's Correlation Matrix for MOT17 and MOT20. In Figure 9 in the main paper we presented a partial Spearman's correlation matrix based on the MOT17 and MOT20 sequences. We used the matrix to evaluate the monotonic relationship between the three complexity metrics (MOTCOM, *density*, and *tracks*) and HOTA, MOTA, and IDF1. In Figure 3 we present the complete Spearman's correlation matrix, which shows additional details on the relationship between the entries.

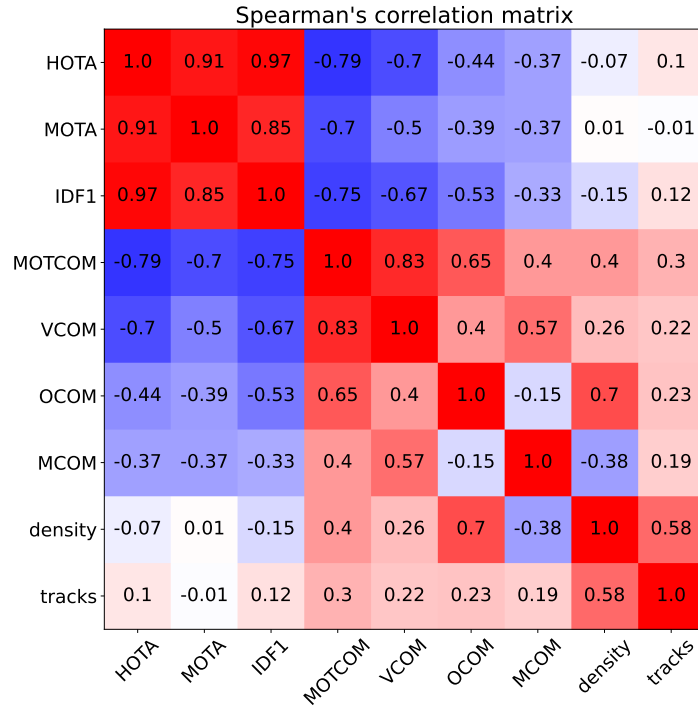


Fig. 3: Spearman's correlation matrix. Based on the average performance of the top-30 trackers on MOT17 and MOT20 test split.

Complete Spearman’s Correlation Matrix for MOTSynth. In the main paper we presented the Spearman’s Footrule Distance and the complexity scores for the MOTSynth sequences. To expand upon this, we include the complete Spearman’s correlation matrix for the MOTSynth train split in Figure 4. The matrix gives a detailed overview of the monotonic relationship between the entries.

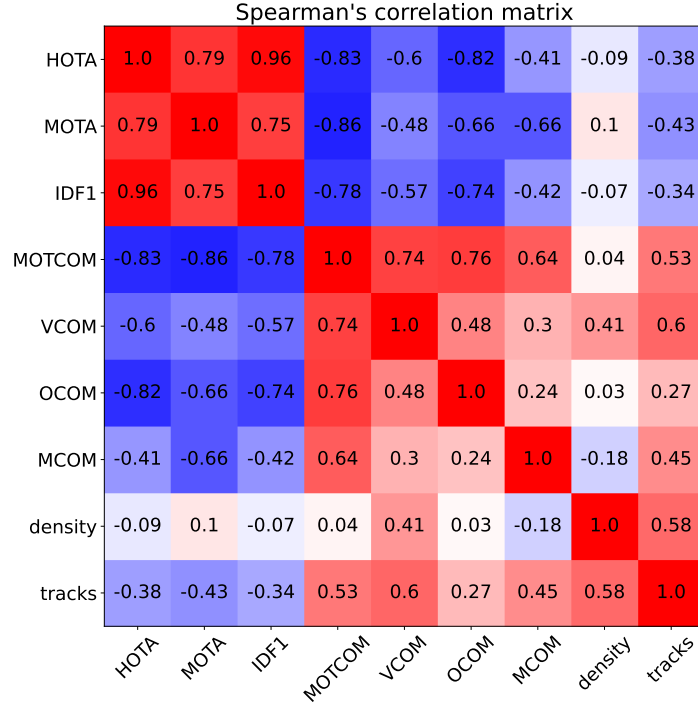


Fig. 4: Spearman’s correlation matrix. Based on the CenterTrack performance on the MOTSynth train split.