On Camera and LiDAR Positions in End-to-End Autonomous Driving (Supplementary Material)

Malte Stelzer[1], Jan Pirklbauer[1], Jan Bickerdt[2], Volker P. Schomerus[2], Jan Piewek[2], Thorsten Bagdonat[2], and Tim Fingscheidt[1]

[1] Technische Universität Braunschweig, Germany
`{malte.stelzer, jan.pirklbauer, t.fingscheidt}@tu-bs.de`
[2] Group Innovation Volkswagen AG, Germany
`{jan.bickerdt, volker.patricio.schomerus, jan.piewek, thorsten.bagdonat}@volkswagen.de`

A Statistical Calculations

This section provides a detailed insight into the calculations of the metrics and their corresponding statistical properties. Each run $b \in \mathcal{B} = \{1, ..., B\},\ B = 3$, of the `Longest6 Benchmark` consists of $n \in \mathcal{N} = \{1, 2, ..., N\},\ N = 36$ routes[3]. For each run and route, the individual route completion $\mathrm{RC}_{b,n}$, and infraction score $\mathrm{IS}_{b,n}$ are tracked, while the driving score $\mathrm{DS}_{b,n}$ is calculated as

$$\mathrm{DS}_{b,n} = \mathrm{RC}_{b,n} \cdot \mathrm{IS}_{b,n}. \tag{1}$$

For each run b, the metrics are averaged as

$$\mathrm{RC}_b = \frac{1}{N} \sum_{n \in \mathcal{N}} \mathrm{RC}_{b,n}, \tag{2}$$

$$\mathrm{IS}_b = \frac{1}{N} \sum_{n \in \mathcal{N}} \mathrm{IS}_{b,n}, \tag{3}$$

$$\mathrm{DS}_b = \frac{1}{N} \sum_{n \in \mathcal{N}} \mathrm{DS}_{b,n}. \tag{4}$$

While $\mathrm{DS}_{b,n}$ is entangled with $\mathrm{RC}_{b,n}$ and $\mathrm{IS}_{b,n}$ on a route level (see (1)), the resulting DS_b is disentangled with RC_b and IS_b on a run level, since

$$\mathrm{DS}_b = \frac{1}{N} \sum_{n \in \mathcal{N}} \mathrm{DS}_{b,n} = \frac{1}{N} \sum_{n \in \mathcal{N}} \mathrm{RC}_{b,n} \cdot \mathrm{IS}_{b,n} \neq \mathrm{RC}_b \cdot \mathrm{IS}_b.$$

In the results reported in Table 1 of the main paper, the metrics are shown as average of $B = 3$ benchmark runs with the respective sample standard deviation. To provide an example for DS, these are calculated as

$$\mathrm{DS} = \frac{1}{B} \sum_{b \in \mathcal{B}} \mathrm{DS}_b \tag{5}$$

[3] For simplicity, the term *routes* is used instead of *route-based scenarios*.

and

$$s(\mathrm{DS}_b) = \sqrt{\tfrac{1}{B-1}\sum_{b\in\mathcal{B}}(\mathrm{DS}-\mathrm{DS}_b)^2}, \tag{6}$$

where the computation of the averages for the other two metrics RC and IS, as well as their sample standard deviations $s(\mathrm{RC}_b)$ and $s(\mathrm{IS}_b)$, respectively, is analogous. This means that the sample standard deviations of the three metrics are also disentangled, since

$$\mathrm{DS} \neq \mathrm{RC}\cdot\mathrm{IS}$$

holds. An alternative to (6) to compute the sample standard deviation would be

$$s(\mathrm{DS}_{b,n}) = \sqrt{\tfrac{1}{B\cdot N-1}\sum_{b\in\mathcal{B}}\sum_{n\in\mathcal{N}}(\mathrm{DS}-\mathrm{DS}_{b,n})^2}. \tag{7}$$

However, this comes along with several disadvantages. While (6) compares B identical sets of routes, (7) would give us a standard deviation of the driving score not only over the (non-)deterministic benchmark runs $b \in \mathcal{B}$ (desired), but also over the diverse test routes n (not of interest here!). In consequence, (7) would naturally result in larger standard deviations. Therefore, (6) is used for the calculations, as it allows for a fair comparison of identical sets of routes at a benchmark level, thereby reporting statistical variability of `CARLA` during our $B = 3$ benchmark runs. That is what is of interest for us.