

Driver Behavior Dataset

The dataset is a collection of smartphone sensor measurements for driving events. An Android application is used to record smartphone sensor data, like accelerometer, linear acceleration, magnetometer and gyroscope, while a driver executed particular driving events.

We performed the experiment in 4 car trips of approximately 13 minutes each, in average. The experiment's conditions, its setup and data labelling were the following:

1. The vehicle was a 2011 Honda Civic;
2. The smartphone was a Motorola XT1058 with Android version 5.1;
3. The smartphone was fixed on the car's windshield by means of a car mount, and was neither moved nor operated while collecting sensor data;
4. The motion sensors sampling rate varied between 50 and 200 Hz, depending on the sensor;
5. Two drivers with more than 15 years of driving experience executed the driving events;
6. The weather was sunny and the roads were dry and paved with asphalt.
7. The Android app which collected sensor data also displayed a real time timer with milliseconds precision. The timer would display the elapsed time in the current trip which could be later converted to match the "uptimeNanos" field in the datasets.
8. During the experiments, there were two people in the car. The driver and an assistant in the back seat who filmed the experiment.
9. The recorded video showed the front window view, the driver's movements, the car steering wheel, velocimeter and tachometer, and the smartphone (displaying the running timer) mounted in the car panel. With this arrangement, someone watching the video later on could tell exactly when the driving event started and ended. That is how the data was labelled.
10. Both drivers were not experts in this field, but they had plenty of driving experience.
11. According to a predefined schedule, the assistant asked the driver to execute a specific driving event type (e.g. aggressive braking, aggressive cornering) and waited for its execution while recording the video. There were no detailed instructions on how the maneuver should be performed. Hence, we relied on the driver's perception and experience on how to perform an aggressive or non-aggressive maneuver.

Our purpose was to establish a set of driving events that represents usual real-world events such as breaking, acceleration, turning, and lane changes. The table below shows the 7 driving events types we used in this work and their number of collected samples. We strived to collect a balanced number of samples for each driving event type. However, we discarded 11 lane change events due to errors during the experiment. Therefore, the number of these events is slightly unbalanced. We also made our best effort to evenly distribute event samples between both drivers.

Driving Event Type	# of samples
Aggressive breaking	12
Aggressive acceleration	12
Aggressive left turn	11
Aggressive right turn	11
Aggressive left lane change	4
Aggressive right lane change	5
Non-aggressive event	14
Total	**69**

Dataset structure

#####

Each folder represents a car trip containing the following 6 files:

-> Acceleration -> accelerometer -> linear file and other non-linear file -> 2 file.

* aceleracaoLinear_terra.csv - File with the Linear Acceleration in 3 axis.

* acelerometro_terra.csv - File with the Acceleration in 3 axis.

-> Magnetometer -> مقياس شدة المجال المغناطيسي

* campoMagnetico_terra.csv - File with the Magnetometer readings in 3 axis.

-> Gyroscope

* giroscopio_terra.csv - File with the Gyroscope readings in 3 axis.

-> time-sensor

* groundTruth.csv - File with the start and end timestamps for each event.

-> phone info

* viagem.json - Info about the smartphone used.

The following fields are common in the files:

* Timestamp - Timestamp when the record is taken.

* UptimeNanos - Time in nanoseconds since the smartphone start.

The *groundTruth.csv* file has the start and end of the events, and the following fields:

- * Evento - Type of event.
- * Inicio - Start of the event in seconds.
- * Fim - End of the event in seconds

The fields *inicio* and *fim* are counted since the start of the trip, so if the **inicio** field is 2 and the **fim** field is 4.5, the event started in the second 2 of the trip and ended in the second 4.5.

Related work

The following work used this data set:

- * [Driver behavior profiling: An investigation with different smartphone sensors and machine learning]
(<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174959>)
- * [Exploiting the Use of Recurrent Neural Networks for Driver Behavior Profiling] (<http://ieeexplore.ieee.org/document/7966230/>)
- * [Análise de Perfil de Motoristas: Detecção de Eventos por meio de Smartphones e Aprendizado de Máquina]
(<http://sbrc2016.ufba.br/downloads/WoCCES/154831.pdf>)

If you use this dataset in your work, fell free to send a PR to keep this list updated.