Google Street View Images Support the Development of Vision-Based Driver Assistance Systems

Jan Salmen, Sebastian Houben, and Marc Schlipsing
Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany
firstName.lastName@ini.rub.de

Motivation

- The development of vision-based modules typically requires large amounts of data.
- Collecting appropriate data is often more time-consuming and more expensive than the algorithm development itself.
- Sometimes publicly available data sets can be useful.
- Google already did a lot of recording. The retrieval of images from Google Street View becomes a powerful alternative!

The Google Street View Static API

- Google Street View is a network of adjacent 360° high resolution panorama images.
- Those images are intended to be accessed via the Google Maps API. In order to make them usable for algorithm development, some undocumented features have to be considered.
- It is not completely clear that the presented way of usage is compliant with Google’s terms of service.
- Many of the technical details have been described recently by Jamie Thompson in his weblog: http://jamiethompson.co.uk/web/2010/05/15/google-streetview-static-api
- Access to thumbnail images ...
- ... and image tiles can easily be achieved via an URL interface.
- In the largest zoom level, a panorama image with an impressive resolution of 90 mega pixels can be accessed.

Our framework

- GUI for virtual driving in Google Street View.
- Map scrolling and zooming realized with the freely available Marble toolkit: http://edu.kde.org/marble
- Definition of starting point and rectangular search.
- Using a depth-first-search algorithm.
- Saving panorama images, image processing, collection of statistical data, visualization of virtually driven distances and overall progress.
- Mean virtual driving speed up to 300 km/h.

First evaluation

<table>
<thead>
<tr>
<th>City</th>
<th>Distance [km]</th>
<th>#Panoramas</th>
<th>Time [h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>San Francisco, USA</td>
<td>3023.8</td>
<td>268,127</td>
<td>10:03</td>
</tr>
<tr>
<td>Penghu Islands, Taiwan</td>
<td>514.8</td>
<td>44,736</td>
<td>1:41</td>
</tr>
<tr>
<td>Port Elizabeth, South Africa</td>
<td>2191.1</td>
<td>175,369</td>
<td>5:58</td>
</tr>
<tr>
<td>Belo Horizonte, Brazil</td>
<td>1426.3</td>
<td>129,459</td>
<td>4:33</td>
</tr>
<tr>
<td>Alcalá de Henares, Spain</td>
<td>499.9</td>
<td>32,445</td>
<td>1:01</td>
</tr>
<tr>
<td>Gold Coast, Australia</td>
<td>2509.8</td>
<td>219,558</td>
<td>8:34</td>
</tr>
</tbody>
</table>

Example application

- Collecting traffic sign images from Google Street View using a Viola-Jones detector.
- Semi-supervised training procedure, final detector with a precision of 49.3% and a recall of 81.6%.
- For the very first detector training, we used images from the “German Traffic Sign Recognition Benchmark” (GTSRB, see benchmark.ini.rub.de) as positive examples and randomly chosen images from Street View as negative examples.
- If no data had been available, one could have collected a few training examples manually from Street View panorama images or any other image source.
- We were able to collect more than 10,000 images of traffic signs within a single week...
- ... already more different sign instances than in our GTSRB dataset.
- Collecting data for GTSRB took our group several months!
- Virtual driving speed of approx. 40 km/h on our test system.

Discussion

- The available data is biased – Recordings were predominantly made during daytime with stable weather conditions.
- No context between recordings – There are only few recordings of the same object and sometimes acquisitions were done months apart even with connected panoramas.
- A data set of this size has to be processed automatically – False Possible results are excluded systematically.

⇒ It becomes much easier for scientists and developers to train and test their approaches on large data sets.
⇒ In future work, we plan to collect datasets for object detection and classification in different countries.