Fusion Framework for Moving-Object Classification

Omar Chavez, Trung-Dung Vu (UJF)
Trung-Dung Vu (UJF)
Olivier Aycard (UJF)
Fabio Tango (CRF)
Introduction

- Advance Driver Assistant Systems (ADAS) help drivers to perform complex driving tasks and avoid dangerous situations.
- ADAS generally have three components: perception, reasoning & decision and control.
- Perception provides, by processing sensor measurements, information of the environment the robot is immerse in.
  - Simultaneous localization and mapping (SLAM) deals with modeling static parts.
  - Detection and tracking moving objects (DATMO) is responsible for modeling dynamic parts.

![Diagram of ADAS components]

Array of Sensors -> Perception -> Environment Description
Reasoning & Decision -> Set of actions
Control -> Actuators' Instructions
Our work at interactIVe project aims at providing a reliable list of objects of interest using two modules:

- Frontal Object Perception (FOP) delivers descriptions about relevant objects of interest (e.g.: location, speed) in the frontal area of the ego vehicle
- Moving Object Classification (MOC) aims to provide estimated information about the class of moving objects detected by the FOP module

An object could be categorized into different classes: pedestrian (or group of pedestrians), bike, car and truck
Sensor processing (1)

Target Detection:
- Segmentation based on lidar point clouds from several frames
- Taking into account spatial and temporal information (similar to optical flows technique)

Model-based tracking
- Better tracking result
- Able to estimate object geometry: updated over time from new observations
- Provide a likelihood of object class for moving objects (classes of moving objects are quite limited)
Train multiple binary-classifiers:
- One classifier for each view of object: pedestrian, car(rear, front), truck(rear, front)
- Sparse-HOG features: compact, fast to compute using integral image
- Learning method: Adaboost

Final classifiers:

Classification-based Detectors:
- Input: ROIs computed from Lidar Targets & Radar Targets
- Apply a classic sliding window strategy inside ROIs (not entire image): ~2ms/100 img patches
Fusion for Moving Object Classification

- Generic framework to fuse classification information from different sources (detector modules)
- Based on DS theory
  - Class information is represented as evidence masses
  - Manage conflict situations when evidence sources don’t agree
- Takes into account the reliability of the sources and their precision to classify specific objects
Fusion for Moving Object Classification

- Set of possible class hypothesis per object $\Omega = \{\text{car}, \text{truck}, \text{pedestrian}, \text{bike}\}$

  \[
  m(\emptyset) = 0; \quad \sum_{A \subseteq \Omega} m(A) = 1
  \]

  \[
  m_r(A) = \sum_{B \cap C = A} m_b(B) m_e(C); \quad A \neq \emptyset
  \]

  \[
  K = \sum_{B \cap C = \emptyset} m_b(B) m_e(C)
  \]

  \[
  m_r(\Omega) = m'_r(\Omega) + K
  \]

- Instantaneous fusion: combine information from different sources at current time $t$.

  Applying reliability factor

  \[
  m_b(A) = r_{ab} \times m'_b(A); \quad A \subseteq 2^\Omega, A \neq \Omega
  \]

  \[
  m_b(\Omega) = m'_b(\Omega) + \sum_{A \subseteq 2^\Omega} (1 - r_{ab} \times m(A)); \quad \text{for } A \neq \emptyset, A \neq \Omega
  \]

  Weighting specific hypothesis

  \[
  m_a(A_i) = m'_a(A_i) \times f_i; \quad A_i \subseteq 2^\Omega, A_i \neq \emptyset
  \]

  \[
  m_a(\Omega) = m'_\Omega + \sum_{A_i \subseteq 2^\Omega} (1 - f_i) \times m'_a(A_i); \quad \text{for } A_i \neq \emptyset, A_i \neq \Omega
  \]

- Dynamic fusion: instantaneous result is combined with fused information at time $t-1$
Experiments

Setup

- Four class hypothesis: car, truck, pedestrian, bike
- Three different classification sources:
  - Vehicle detector, pedestrian detector and lidar-based detector
- CRF vehicle demonstrator includes three main sensors: radar, lidar and camera
- Reliability and precision factors are obtained experimentally using real datasets
Frontal Object Perception

Trung-Dung Vu, Olivier Aycard
Université Joseph Fourier Grenoble, France

Contact: Olivier.Aycard@imag.fr
Results

Qualitative performance

Test track

Real scenarios
## Results

### Quantitative performance

- **Real time performance of FOP-MOC**

<table>
<thead>
<tr>
<th>Vehicle class</th>
<th>Dataset</th>
<th>Lidar processing</th>
<th>Vehicle detector</th>
<th>Fusion approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>highway 1</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>highway 2</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>urban 1</td>
<td>15</td>
<td>19</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>urban 2</td>
<td>18</td>
<td>23</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pedestrian class</th>
<th>Dataset</th>
<th>Lidar processing</th>
<th>Pedestrian detector</th>
<th>Fusion approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>urban 1</td>
<td>10</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>urban 2</td>
<td>13</td>
<td>7</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

### General performance

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Total objects</th>
<th>Correct Detection</th>
<th>False Detection</th>
<th>Correct Classification</th>
<th>False Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ped</td>
<td>car</td>
<td>truck</td>
<td>ped</td>
<td>car</td>
</tr>
<tr>
<td>Motorway</td>
<td>0</td>
<td>682</td>
<td>216</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Urban</td>
<td>33</td>
<td>525</td>
<td>87</td>
<td>27</td>
<td>495</td>
</tr>
<tr>
<td>Test track</td>
<td>248</td>
<td>301</td>
<td>0</td>
<td>247</td>
<td>300</td>
</tr>
</tbody>
</table>
Conclusions

- Fusion approach includes information about reliability of the sources and specific precision factors
- Architecture can be extended by including more detector modules
- Several class of objects involved
- Improves the performance of individual single sensor-based modules
- Provides a confidence value along the final object classification
Thank you.