Beyond Thinking in Common Categories: Predicting Obstacle Vulnerability using Large Random Codebooks

Johannes Rühle, Erik Rodner, and Joachim Denzler
Computer Vision Group, Friedrich Schiller University of Jena, Germany
http://www.inf-cv.uni-jena.de/vulnerability

Motivation and Goal

- increase road safety and driver awareness
- more detailed understanding of obstacles in front of the vehicle in emergency situations
- **How vulnerable are the present obstacles?**
  classify obstacle hypotheses from a vehicle stereo camera and obstacle detection unit

Vulnerability Classes [1]:

- **none**, **medium**, **heavy**, **fatal**
- forming supersets of object categories

Challenges and Method

**Challenges:**
- use of complex generic features to represent unique vulnerability patterns
- no rigid feature constellation due to the variability of obstacle hypotheses in size and coverage

**Bag of Words (BoW) Advantages:**
- histogram representation as distribution of visual features
- robust against changing amount of visual features in a hypothesis
- spatial pooling suitable to handle occlusions

**Image Acquisition from Stereo Camera with Obstacle Detection**

**Multi-Scale Dense RootSIFT [2] Feature Extraction**

**Multi-Cue Feature Fusion**

**Codebook Matching Histogram Generation**

**Vulnerability Classification using SVM**

**Evaluation Results and Dataset Details**

**Optimized Evaluation Parameters:**
- Large Codebooks: 25k random codebook entries
- Multi-Scale: using 9 different scales
- Multi-Cue: using disparity and intensity information

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>ACC</th>
<th>ARR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Guessing</td>
<td>43.9</td>
<td>25.0</td>
</tr>
<tr>
<td>Human Expert</td>
<td>91.7</td>
<td>85.6</td>
</tr>
</tbody>
</table>

**Evaluation Details:**
- Imageset: dataset1
- results averaged over 5 trials per run

**Dataset Details:**
- 17 authentic real-world street sequences
- 10k obstacles in 2402 images and disparity images
- obstacles labeled with vulnerability classes

**Runtime:**
- 8.57s

**Multi-cue: image and disparity**
- Fusion of SIFT descriptors 83.6
- Fusion of BoW histograms 83.5

**Visual Results**

- Good Results
- Failure Cases

**Bicyclist Sequence**

**Conclusion**

- vulnerability classification of arbitrary obstacle hypotheses
- classification using a Bag of Words approach with large random codebooks best on intensity and disparity data
- successful application in 17 real-world street sequences using data from a vehicle stereo camera

---