Lane Detection based on Object Detection and Image-to-image Translation

Hiroyuki Komori, Kazunori Onoguchi
Hirosaki University, Japan
Many vision-based lane detection methods have been proposed for advanced driver assistance system or autonomous driving system. Most of these methods detect lane markings such as white lines drawn on the road surface. Traffic lanes are divided by various roadside objects. The proposed method detects various types of lane markings and road boundaries simultaneously from a monocular camera image.
Outline of the proposed method

- Processing flow

```
Input image → IPM image → Object detection → Boundary detection
```

- IPM image
- Curb: Solid line
- Green: Bounding box
- Red: Boundary

Techniques:
- YOLOv3
- pix2pix
Detection of lane boundary object

- **Object detection network**
  - **YOLOv3**

  Darknet: Conv layer + residual block  
  Detection: Conv layer + Upsampling  
  Conv layer: filter + Batch Normalization + Leaky ReLU

8 categories are identified by the object detection network:
- **Lane marking**: Solid line, Dashed line, Zebra line
- **Road boundary**: Curb, Grass, Guardrail, Sidewall, Snow sidewalk
Datasets for object detection

- New datasets in which bounding boxes are annotated in the IPM image are built and used for training

Examples in datasets for object detection
Boundary generation

- **Image-to-image translation network**
  - In each bounding box extracted by the object detection network, boundary lines with a lane marking or a roadside object are generated by pix2pix.
  - Different generators are created for each category.
  - Boundaries are generated using a generator that corresponds to the category of the bounding box.

![Diagram showing the process of boundary generation](Diagram.png)
Datasets for boundary generation

- **A pair images**
  - Original image: a bounding box image obtained from object detection network
  - Boundary image: an image adding line borders to an original image manually
    - In solid line and dashed line, line borders are drawn on both sides of a white line

- **Zero padding**
  - The image size input to pix2pix is 256 x 256 pixels
    - When the height and width of the cut-out image are less than 256 pixels, the periphery of the image is filled with pixel value 0

![Curb](image1)
![Grass](image2)
![Solid line](image3)
![Dashed line](image4)

Boundary image of each class
Experiments (1)

- Detection of lane boundary object
  - Datasets
    - IPM images annotating bounding boxes around the lane markings and road boundaries
    - Image size: 720 x 560 pixels → 608 x 608 pixels (resize)
    - Train: 3,240 images, Test: 1,172 images
  - Quantitative evaluation
    - True positive (IoU > 0.5 and Category reliability > 0.25)
  - mAP for testing: 92%
  - Processing time: 16ms

<table>
<thead>
<tr>
<th></th>
<th>curb</th>
<th>grass</th>
<th>guardrail</th>
<th>sidewalk</th>
<th>snow</th>
<th>snow</th>
<th>white line (dashed)</th>
<th>white line (zebra)</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>train</td>
<td>98.9</td>
<td>99.2</td>
<td>98.2</td>
<td>98.5</td>
<td>99.1</td>
<td>99.5</td>
<td>99.1</td>
<td>99.7</td>
<td>99.0</td>
</tr>
<tr>
<td>test</td>
<td>91.7</td>
<td>97.1</td>
<td>93.3</td>
<td>87.2</td>
<td>92.4</td>
<td>95.4</td>
<td>96.4</td>
<td>82.8</td>
<td>92.0</td>
</tr>
</tbody>
</table>
Experiments (2)

- Detection of lane boundary object
  - Examples of true positive detection
Boundary generation

Three generators for a curb, grass and a white line (solid line and dashed line) were created and evaluated.

Datasets

- Curb: train 1,000 images, test 110 images
- Grass: train 800 images, test 100 images
- White line (Solid line and dashed line): train 3,200 images, test 1,060 images
- Image size: 256 x 256 pixels
White line (solid line, dashed line)

- Quantitative evaluation
  - The degree of overlap between white line area of the ground truth image and white line area of the generated image are evaluated by IoU.
  - The ground truth areas of the white line is created manually.
  - F-measure: 99.3% (IoU > 0.3), 96.4% (IoU > 0.5)

- Qualitative evaluation

- Processing time: 3.0ms
Curb and Grass

Quantitative evaluation
- When the average error in the horizontal direction between the generated boundary and the ground truth is less than about 5 pixels, it's determined as the true positive.
- True positive: 80% of 110 images for curb
  93% of 100 images for grass

Qualitative evaluation

Processing time
- Curb: 2.8ms  Grass: 2.9ms
Conclusion

- We proposed a method to detect various types of lane markings and road boundaries simultaneously from monocular camera image
  - First, the object detection network detects bounding boxes surrounding a lane marking or the boundary with roadside object
  - Next, in each bounding box, lane marking boundaries and road boundaries are drawn by the image-to-image translation network

- Experimental results using our own dataset show the effectiveness of the proposed method for detecting lane markings and road boundaries