Application of Computer Vision on Self-driving Car

Shimeng Chen Xinzhe Liu Yi Qin

Abstract

This paper focus on how to use some knowledge of computer vision to solve some detection and estimation problems .such as lane detection and pedestrian detection and distance estimation.

1. Introduction

1.1. Advanced Lane Detection

Identifying lanes on the road is a common task performed by all human drivers to ensure their vehicles are within lane constraints when driving, so as to make sure traffic is smooth and minimize chances of collisions with other cars in nearby lanes. Similarly, it is a critical task for an autonomous vehicle to perform. It turns out that recognizing lane markings on roads is possible using well known computer vision techniques.

1.2. Pedestrians Detection and Binocular

Human detection in video plays an important role in a wide range of applications that intersect with many aspects of our lives. Sometimes we need to detect pedestrian firstly, and then do the next work such as segmentation, tracking, Re-identification, action prediction and so on. In this project, we will detect human firstly and then try to confirm the distance from our car to them by binocular vision.

2. Previous work

2.1. Advanced Lane Detection

Where we learn how to detect edge by using canny edge detector:

John Canny “A Computational Approach to Edge Detection”

2.2. Pedestrians Detection

Where we learned how to use HOG feature extraction algorithm:

N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in CVPR, Vol. 1, 2005, pp. 886-893.

Where we learned how to use CENTRIST visual descriptor:

J. Wu and J. M. Rehg, "CENTRIST: A visual descriptor for scene categorization," IEEE TPAMI, Vol. to appear.

Where we learned how to make some few optimizations and which give us some inspirations for integrating all three result:

Kelvin Lee, Che Yon Choo, Hui Qing See, Zhuan Jiang Tan and Yunli Lee, "Human detection using Histogram of oriented gradients and Human body ratio estimation,"

Where we learned how to combine these two visual descriptor:

X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling,"

Where we learn YOLO detection method:

Joseph Redmon , Santosh Divvala , Ross Girshick , Ali Farhadi,“You Only Look Once: Unified, Real-Time Object Detection”

3. Technical Approach

3.1. Advanced Lane Detection

Step1: convert RGB picture to HSV picture for convenient of detecting yellow and white color.
Step 2: Gaussian blur in order to eliminate noise
Step 3: use canny detector to detect edge
Step 4: create a mask to abandon the line outside the road
Step 5: use Hough transform to detect lane

3.2. Pedestrians Detection

For every frame of the video:

Step1: Use YOLO to get the label, position, and confidence of pedestrians, vehicles or other obstacles in this frame.

Step2: Using the HOG + CENTRIST feature descriptor and SVM classifier, the pedestrian in the frame is detected by sliding-window, and the position information and the confidence are also obtained

Step3: Find the point where the confidence is greater than the threshold and mark the location with a rectangle.

Step4: Try looking for the rectangle that corresponds to a potential target box in the previous frame.

Step4.a: Found the corresponding rectangular; record the location; improve the confidence; determine whether the confidence threshold has been reached.

Step4.a.i: Confidence reaches the target threshold; save as the target.

Step4.a.ii: Confidence reaches a potential target threshold, saved as a potential target.

Step4.a.iii: Confidence does not reach the potential target threshold, discard the rectangle

Step4.b: Did not find the corresponding frame; skipped

Step5: Traverse the rectangular box in this frame without correspond target box in previous frame; determine whether is it in the edge of the image (rectangular box from the edge of the distance is less than the threshold or the frame exist a target box with huge displacement)

Step5.a: At the edge of image; determine whether the confidence threshold has been reached.

Step5.b: Not at the edge of the image; reduce the confidence; determine whether the confidence threshold has been reached.

Step5.a.i: Confidence reaches the target threshold; save as the target.

Step5.a.ii: Confidence reaches a potential target threshold, saved as a potential target.

Step5.a.iii: Confidence does not reach the potential target threshold, discard the rectangle

Step6: Traverse the target box in previous frame without correspond target box in this frame; based on the position information of the first 6 frames, roughly predicted whether the rectangular box is in this frame or not (with polynomial fitting)

Step6.a: Not in this frame; skipped.

Step6.b: In this frame; try to find the corresponding object; reduce the confidence according to the similarity. determine whether the confidence threshold has been reached.

Step6.a.i: Confidence reaches the target threshold; save as the target.

Step6.a.ii: Confidence reaches a potential target threshold, saved as a potential target.

Step6.a.iii: Confidence does not reach the potential target threshold, discard the rectangle

3.3. Distance estimation by binocular vision

Step1: fix two camera onto shelf

Step2: use chessboard to calibrate the camera

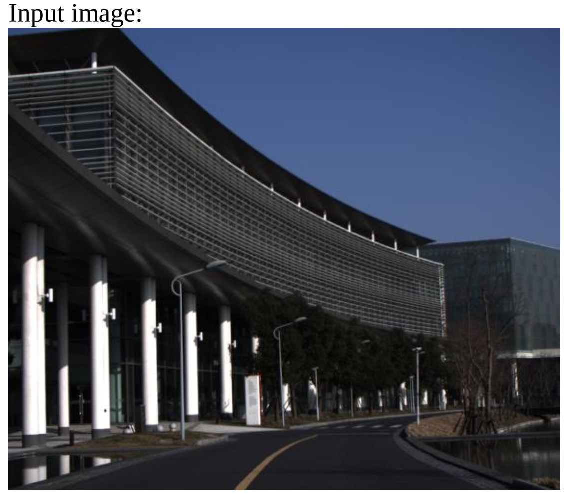
Step3: calculate the disparity map based on intrinsic parameters and extrinsic parameters

Step4: estimate the pedestrian distance based on disparity map

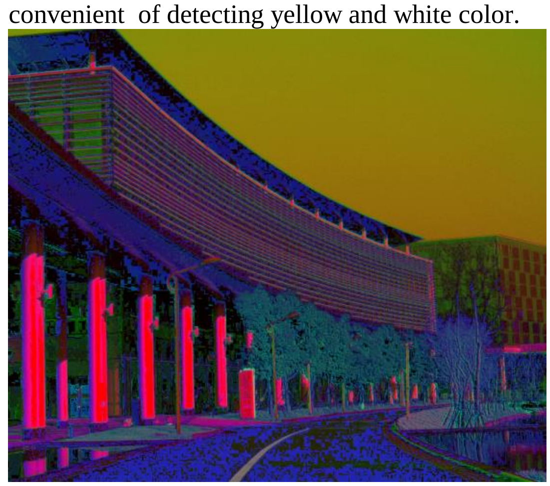
4. Experiment

4.1. Lane detection

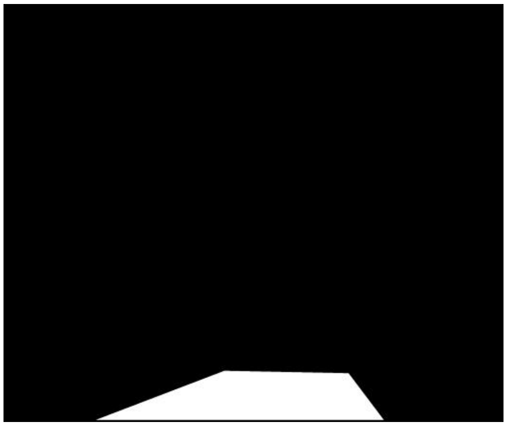
Input image:



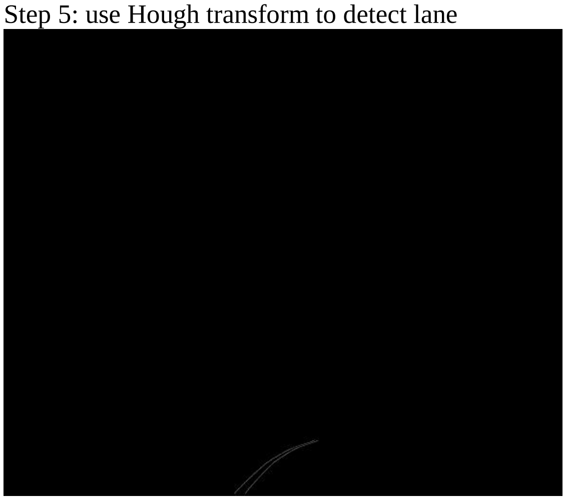
Step1:convert RGB picture to HSV picture for convenient of detecting yellow and white color.



Step 4:create a mask to abandon the line outside the road

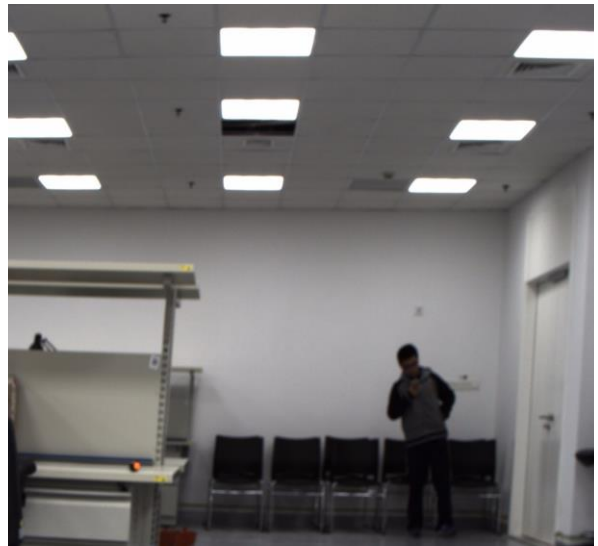


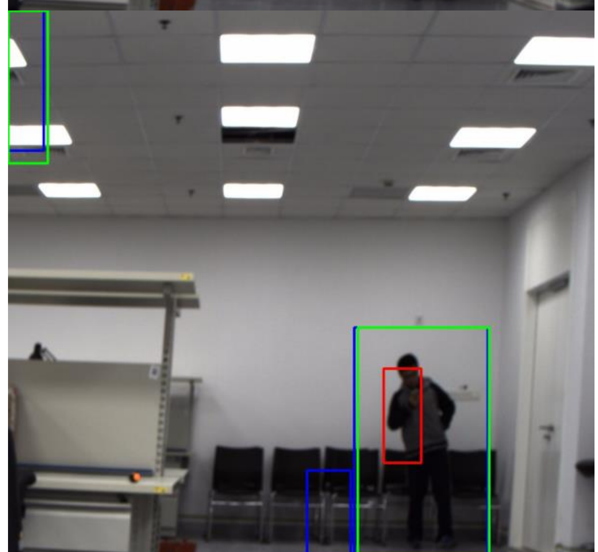
Step 5: use Hough transform to detect lane

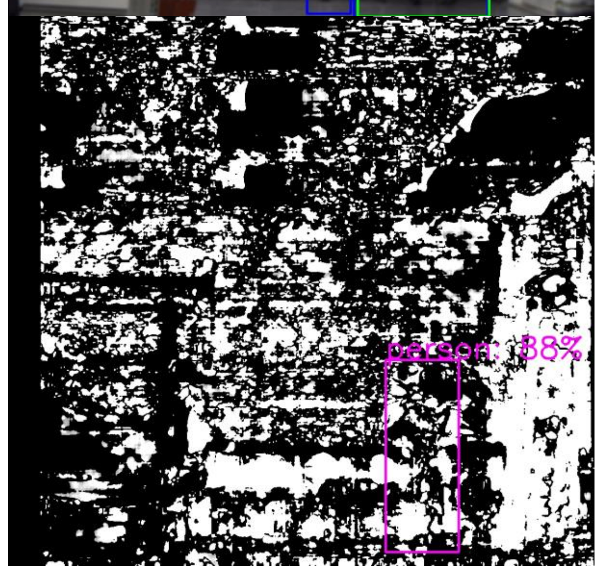


Final result:

4.2. Pedestrians Detection and Distance estimation







5. Conclusion

5.1. Lane detection

Canny edge detector and hough transform are very effective on lane detection. But there is some probability of false alarm on some line not belongings to lanes. If we want to decrease the probability of false alarm, we may use supervised learning to train a classifier to identify which line belongs to the lane

Following is an example of false alarm

5.2. Pedestrians Detection and Distance estimation

In the experiment, our program can return the location of pedestrians during most of time. When the image is blurred, there is a probability of missing pedestrian. In brief, our program can satisfy our need of detecting pedestrians.

Due to the noise of disparity map, we could not estimate the pedestrians’ distance precisely. So there are still a lot work to do in order to improve precision.

6. Reference

[1] John Canny “A Computational Approach to Edge Detection”, 1986 Jun;8(6):679-98..

[2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in CVPR, Vol. 1, 2005, pp. 886-893.

[3] J. Wu and J. M. Rehg, "CENTRIST: A visual descriptor for scene categorization," IEEE TPAMI, Vol. to appear.

[4] Kelvin Lee, Che Yon Choo, Hui Qing See, Zhuan Jiang Tan and Yunli Lee, "Human detection using Histogram of oriented gradients and Human body ratio estimation,"

[5] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling,"

[6] Joseph Redmon , Santosh Divvala , Ross Girshick , Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection”

[7] Tomas Vojir, Jana Noskova and Jiri Matas, “Robust scale-adaptive mean-shift for tracking“. Pattern Recognition Letters 2014.