Pedestrian Proximity Detection using RGB-D Data

Adam Tupper, Richard Green

adam.tupper@pg.canterbury.ac.nz, richard.green@canterbury.ac.nz

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

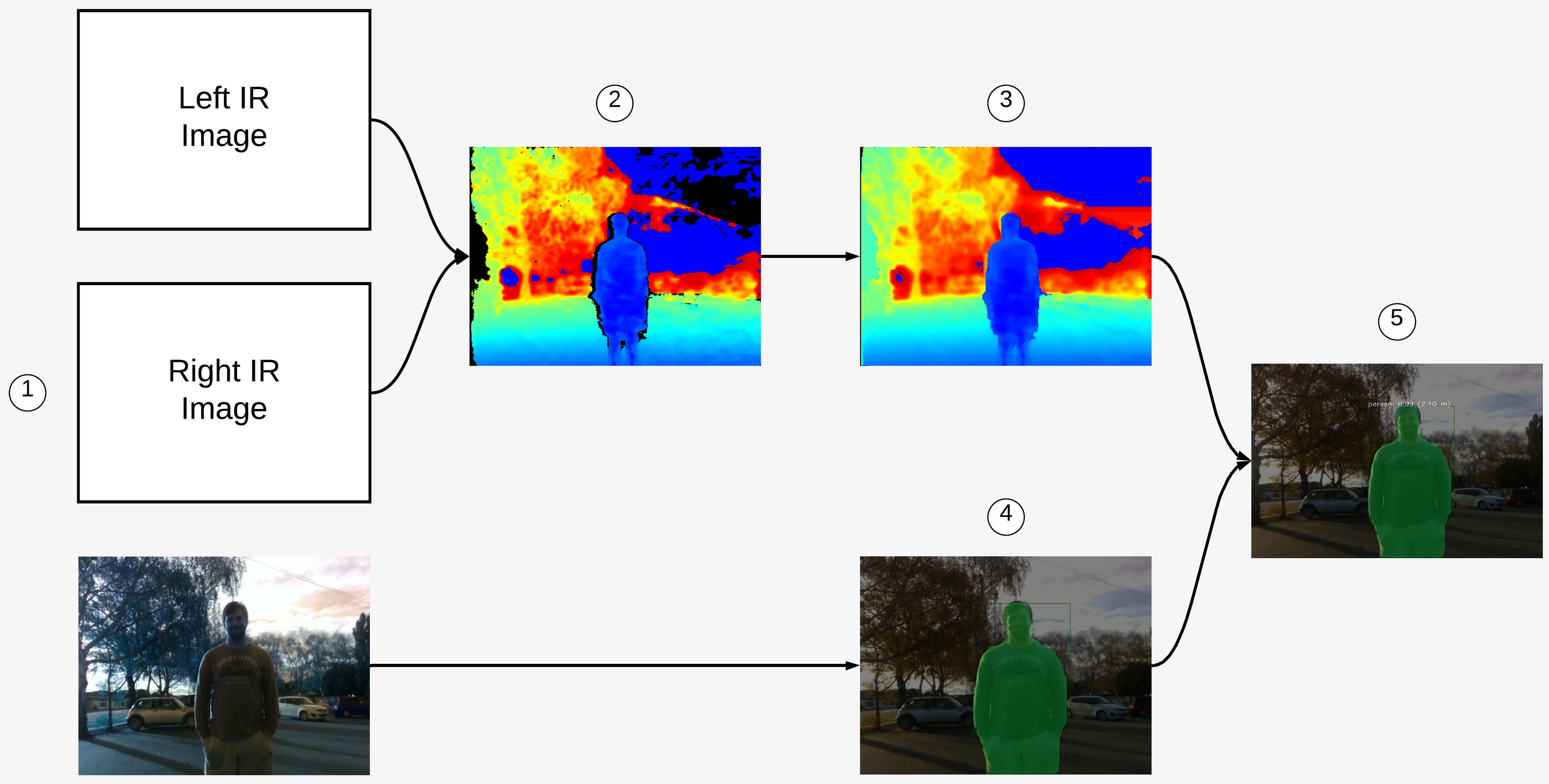
Overview

In 2017, there were 39 pedestrian fatalities and 281 serious injuries as a result of vehicle-related accidents in New Zealand alone [1]. Furthermore, there were 243 workplace fatalities in New Zealand between 2010 and 2018 that were related to vehicles and machinery [2]. In total over 50% of all workplace fatalities over the same period were vehicle or machinery related [2]. **These statistics highlight the need for increased safety measures for vehicles and machines operating in proximity to humans.**

- We propose a method for detecting and monitoring the distance of humans from a machine within a narrow safety envelope using an RGB-D camera.
- Our approach uses human instance segmentation and infrared stereo vision to achieve this.

Figure 1: An example output for multiple pedestrians from our proposed method.

Method



1) A stereo pair of infrared images and a colour image are captured by a RealSense D435 camera.
2) A depth map is computed using the pair of infrared images using the **Semiglobal Matching algorithm** [3].
3) The depth map is post-processed using **edge-preserving spatial filtering**, **spatial hole-filling** and **temporal filtering** [4] to smooth depth noise while preserving object edges and to fill holes in the depth map.
4) The colour image is passed through a **Mask R-CNN human segmentation model**, trained initially on the COCO dataset [5] and then refined on the Supervisely Persons dataset [6].
5) The instance masks for each pedestrian are overlaid onto the depth map and the **median distance estimate for the identified region** is computed.

Evaluation

- For each of the lighting conditions listed in Table 1, a person was placed at 1m intervals within the range of 1m to 5m.

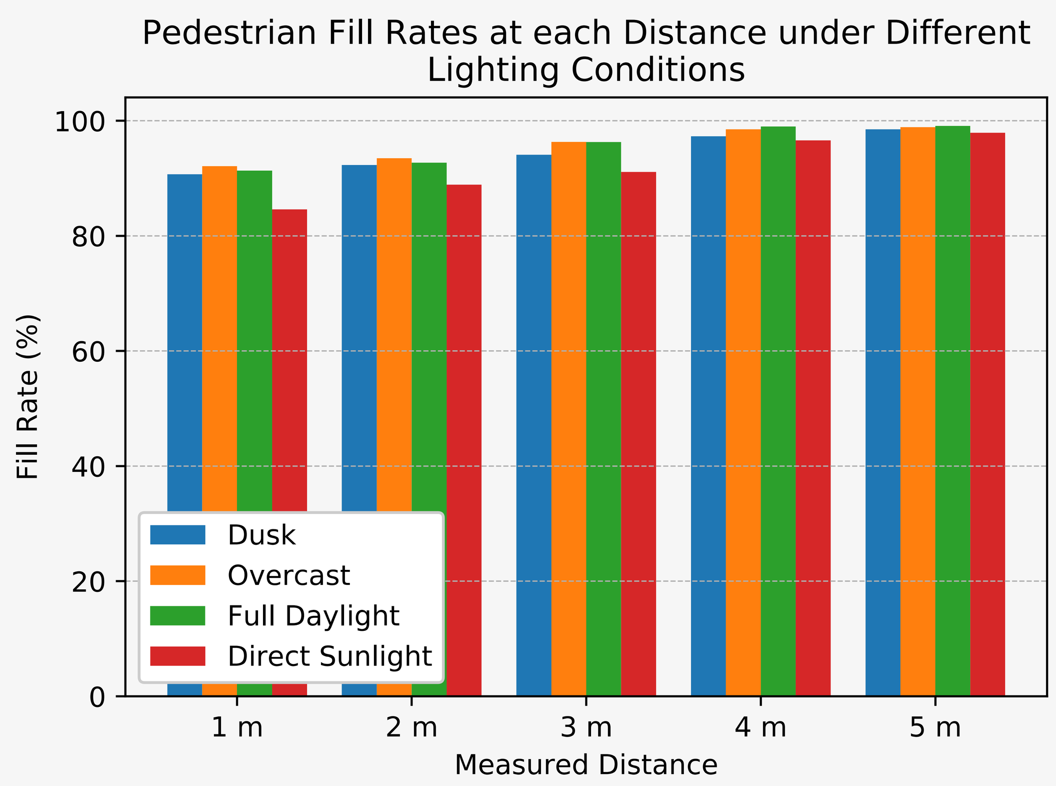
Table 1: Lighting conditions tested in our system evaluation [12].

Lighting Condition	Lux Range
Dawn/Dusk	< 1000 lux
Overcast	1000 - 10,000 lux
Full Daylight	10,000 - 32,000 lux
Direct Sunlight	> 32,000 lux

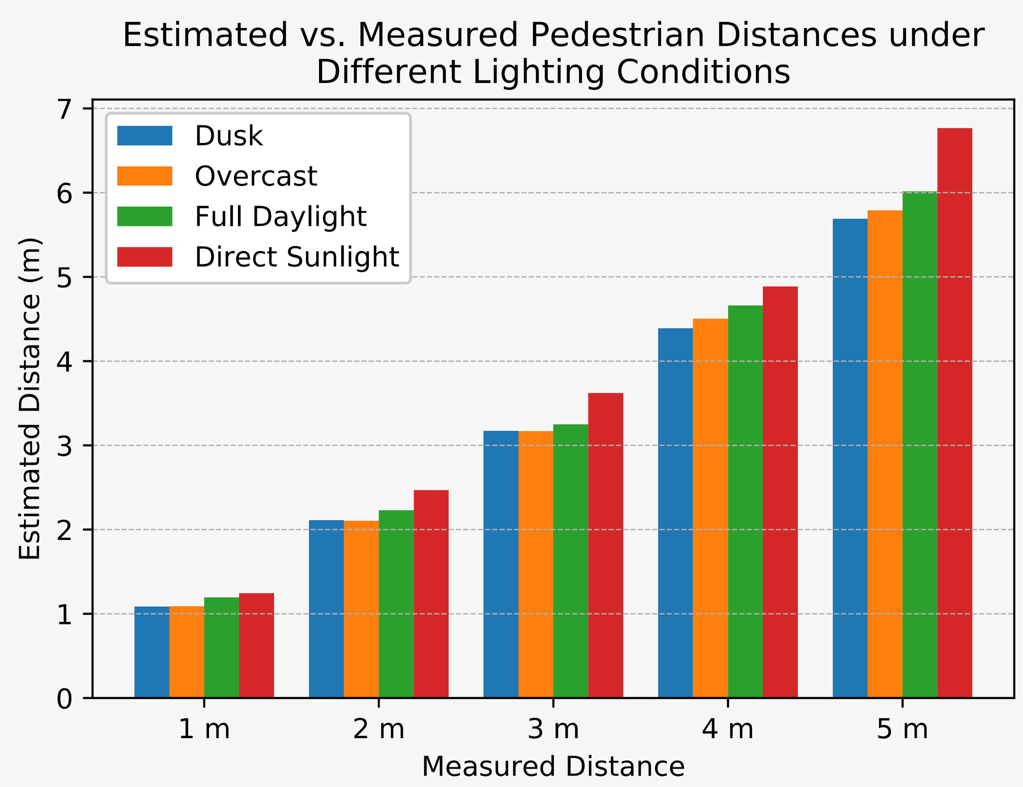
- The Supervisely Persons dataset contains 5722 images with 6884 fine instance-level annotations. These were split 70-30% for training and testing.

Figure 2: The coarse mask annotations included with the COCO dataset (left) compared to the fine mask annotations included with the Supervisely Persons dataset (right).

Results



- For pedestrian segmentation, our model achieves an **AP_{50} score of 94.6%** on the Supervisely Persons dataset.
- Only at distances of 1 m and 2 m in direct sunlight did the average fill rate fall below 90%.



- Our approach performs well across the full range of outdoor lighting conditions and distances, achieving an **average distance estimate accuracy of 87.7%**.
- **Depth estimation degrades with distance and brightness**.
- The degradation with distance can be explained by the reduced disparity between the left and right images at greater distances from the camera.
- The degradation with brightness can be explained by increase in infrared interference.

Conclusions & Future Work

- We present a new **method for detecting pedestrians and estimating their distances** using RGB-D data, based on Mask R-CNN [7] and the depth information captured using an infrared stereo Intel RealSense D435 camera.
- Unlike previous methods tested in only controlled indoor environments [8, 9, 10, 11], our approach **performs well across the full range outdoor lighting** conditions and distances.
- Our method shows promise for use in **automated and assistive driving technologies**, and for **monitoring dynamic safety envelopes** around industrial, agricultural or construction equipment.
- Avenues for future work include:
  - Harnessing **depth data for segmentation**
  - Investigating methods for **increasing distance estimation accuracy** under bright conditions and at greater distances.
  - Exploring different **methods for depth estimate aggregation**.

References

[1] Ministry of Transport, “Pedestrian Crashes,” 2018. [Online]. Available: https://www.transport.govt.nz/mot-resources/new-road-safety-resources/pedestrians/

[2] WorkSafe New Zealand, “WorkSafe Fatalities Detail,” 2019. [Online]. Available: https://worksafe.govt.nz/data-and-research/ws-data/fatalities/

[3] H. Hirschmuller, “Stereo Processing by Semiglobal Matching and Mutual Information,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 328–341, Feb. 2008.

[4] Intel Corporation, “Depth Post-Processing for Intel RealSense D400 Depth Cameras,” 2019. [Online]. Available: https://dev.intelrealsense.com/docs/depth-post-processing

[5] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollr, and C. L. Zitnick, “Microsoft COCO: Common Objects in Context,” in Computer Vision ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing,2014, pp. 740–755.

[6] Supervisely, “Supervisely - Web platform for computer vision. Annotation, training and deploy,” 2018. [Online]. Available: https://supervise.ly/

[7] K. He, G. Gkioxari, P. Dollr, and R. Girshick, “Mask R-CNN,” in 2017 IEEE International Conference on Computer Vision (ICCV), Oct. 2017, pp. 2980–2988.

[8] L. Xia, C. Chen, and J. K. Aggarwal, “Human detection using depth information by Kinect,” in CVPR 2011 WORKSHOPS, Jun. 2011, pp. 15–22.

[9] L. Spinello and K. O. Arras, “People detection in RGB-D data,” in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sep. 2011, pp. 3838–3843.

[10] U. Sharma and R. Green, “Anti-Collision System for Pedestrian Safety,” Computer Vision Lab, University of Canterbury, Tech. Rep., 2017.

[11] J. Nimmo and R. Green, “Pedestrian Avoidance in Construction Sites,” Computer Vision Lab, University of Canterbury, Tech. Rep., 2017.

[12] A. Vit and G. Shani, “Comparing RGB-D Sensors for Close Range Outdoor Agricultural Phenotyping,” Sensors (Basel, Switzerland), vol. 18, no. 12, p. 4413, Dec. 2018.

More Information

SCAN ME

github.com/adamtupper