Abstract

In this paper, a cascade framework is proposed to detect elevator buttons and recognize their labels from images for blind navigation. First, a pixel-level mask of elevator buttons is segmented based on deep neural networks. Then a fast scene text detector is applied to recognize the text labels in the image as well as to extract their spatial vectors. Finally, all the detected buttons and their associated labels are paired by combining the button mask and spatial vectors of labels based on their location distribution. The cascade framework is conducive to multitask but the accuracy may decrease task by task. To avoid the limitation of the intermediate task, a new schema is further introduced by pairing buttons with their labels to consider the region of button and label as a whole. First, the regions of button-label pairs are detected and then the label for each pair is recognized. To evaluate the proposed method, an elevator button detection dataset is collected including 1,000 images containing buttons captured from both inside and outside of elevators with annotations of button locations and labels and 500 images are captured in elevators but without button buttons which are used for negative images in the experiments. Preliminary results demonstrate the robustness and
Recognizing Elevator Buttons and Labels for Blind Navigation

effectiveness of the proposed method for elevator button detection and associated label recognition.

References

1. Saleh Alghamdi, Ron Van Schyndel, and Ahmed Alahmadi. Indoor navigational aid using active rfid and qr-code for sighted and blind people. In 2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pages 18–22. IEEE, 2013.
2. Pierre Baqu´e, Timur Bagautdinov, Franc¸ois Fleuret, and Pascal Fua. Principled parallel mean-field inference for discrete random fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5848–5857, 2016.
3. Jifeng Dai, Kaiming He, and Jian Sun. Instance-aware semantic segmentation via multi-task network cascades. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3150–3158, 2016.
4. Karen Duarte, Jose Cecilio, Jorge S´a Silva, and Pedro Furtado. Information and assisted navigation system for blind people. In Proceedings of the 8th International Conference on Sensing Technology, pages 470–473, 2014.
5. Robert Wall Emerson. Outdoor wayfinding and navigation for people who are blind: accessing the built environment. In International Conference on Universal Access in Human-Computer Interaction, pages 320–334. Springer, 2017.
6. Catherine Feng, Shiri Azenkot, and Maya Cakmak. Designing a robot guide for blind people in indoor environments. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, pages 107–108. ACM, 2015.
7. Wendy P Fernandcz, Yang Xian, and Yingli Tian. Imagebased barcode detection and recognition to assist visually impaired persons. In 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pages 1241–1245. IEEE, 2017.
8. Hugo Fernandes, Vitor Filipe, Paulo Costa, and Jo˜ao Barroso. Location based services for the blind supported by rfid technology. Procedia Computer Science, 27:2–8, 2014.
9. Ronghang Hu, Marcus Rohrbach, and Trevor Darrell. Segmentation from natural language expressions. In European Conference on Computer Vision, pages 108–124. Springer, 2016.
10. Minghui Liao, Baoguang Shi, Xiang Bai, Xinggang Wang, and Wenyu Liu. Textboxes: A fast text detector with a single deep neural network. In Thirty-First AAAI Conference on Artificial Intelligence, 2017.
11. Jingya Liu and Yingli Tian. Recognizing elevator buttons and labels for blind navigation. In 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pages 1236–1240. IEEE, 2017.
12. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In European conference on computer vision, pages 21–37. Springer, 2016.
13. Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015.
14. J Pablo Mu˜noz, Bing Li, Xuejian Rong, Jizhong Xiao, Yingli Tian, and Aries Arditi. An
assistive indoor navigation system for the visually impaired in multi-floor environments. In 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pages 7–12. IEEE, 2017.

15. Rai Munoz, Xuejian Rong, and Yingli Tian. Depth-aware indoor staircase detection and recognition for the visually impaired. In 2016 IEEE international conference on multimedia & expo workshops (ICMEW), pages 1–6. IEEE, 2016.

16. Department of Justice. Ada standards for accessible design. Title III regulation, 28, 2010.

17. Romedi Passini and Guyltne Proulx. Wayfinding without vision: An experiment with congenitally totally blind people. Environment and Behavior, 20(2):227–252, 1988.

18. Ant´onio Pereira, Nelson Nunes, Daniel Vieira, Nuno Costa, Hugo Fernandes, and Joao Barroso. Blind guide: an ultrasound sensor-based body area network for guiding blind people. Procedia Computer Science, 67:403–408, 2015.

19. Lorenzo Picinali, Amandine Afonso, Michel Denis, and Brian FG Katz. Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. International Journal of Human- Computer Studies, 72(4):393–407, 2014.

20. Xuejian Rong, Bing Li, J Pablo Munoz, Jizhong Xiao, Aries Arditi, and Yingli Tian. Guided text spotting for assistive blind navigation in unfamiliar indoor environments. In International Symposium on Visual Computing, pages 11–22. Springer, 2016.

21. Juan Manuel S´aez, Francisco Escolano, and Miguel Angel Lozano. Aerial obstacle detection with 3-d mobile devices. IEEE journal of biomedical and health informatics, 19(1):74–80, 2014.

22. Tobias Schwarzke, Martin Lauer, Manuel Schwaab, Michailas Romanovas, Sandra B¨ohm, and Thomas J¨urgensohn. A camera-based mobility aid for visually impaired people. KIK ´unstliche Intelligenz, 30(1):29–36, 2016.

23. Baoguang Shi, Xiang Bai, and Cong Yao. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE transactions on pattern analysis and machine intelligence, 39(11):2298–2304, 2016.

24. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

25. Long Tian, Chucai Yi, and Yingli Tian. Detecting good quality frames in videos from mobile camera for blind navigation. Journal of Computer Vision and Image Processing, 5(10), 2015.

26. YingLi Tian, Xiaodong Yang, Chucai Yi, and Aries Arditi. Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments. Machine vision and applications, 24(3):521–535, 2013.

27. Kristof Van Beeck, Toon Goedem´e, and Tinne Tuytelaars. Towards an automatic blind spot camera: robust real-time pedestrian tracking from a moving camera. In Proceedings of the twelfth IAPR conference on machine vision applications, pages 528–531. MVA Organization; Japan, 2011.

28. ShuihuaWang, Chucai Yi, and Yingli Tian. Signage detection and recognition for blind persons to access unfamiliar environments. Journal of Computer Vision and Image Processing, 2(2), 2012.

29. Cang Ye, Soonhac Hong, Xiangfei Qian, and Wei Wu. Corobotic cane: A new robotic navigation aid for the visually impaired. IEEE Systems, Man, and Cybernetics Magazine, 2(2):33–42, 2016.
30. Cang Ye and Xiangfei Qian. 3-d object recognition of a robotic navigation aid for the visually impaired. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(2):441–450, 2017.

31. Dong Zhang, Dah-Jye Lee, and Brandon Taylor. Seeing eye phone: a smart phone-based indoor localization and guidance system for the visually impaired. Machine vision and applications, 25(3):811–822, 2014.

32. Shuai Zheng, Sadeep Jayasumana, Bernardino Romera- Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr. Conditional random fields as recurrent neural networks. In Proceedings of the IEEE international conference on computer vision, pages 1529–1537, 2015.

Index Terms

Computer Science

Information Systems

Keywords

Object detection, Semantic segmentation, Computer vision, Deep learning