Abstract

There is no low-cost aid for visually impaired people despite several advances in technology. This paper presents a mobile head-mounted device to detect and convert text in natural scenes to speech. The major components of the device are a Raspberry Pi, a high definition webcam, earphones and a portable power bank. The Raspberry Pi is connected to the webcam which captures the image. A text detection algorithm using Class Specific Extremal Regions (CSERs) is implemented to detect the text in complex natural scenes. The segmented image is passed to the Tesseract OCR engine for text detection. The identified text is converted to audio using the espeak Python module in the Raspberry Pi. Thus, a visually impaired person can use this device to hear all the text in his surroundings like the name of a shop, public notices, billboards, road directions, etc.

References

http://www.who.int/mediacentre/factsheets/fs282/en/
17. Thierry DutoitTTS research team, TCTS Lab:An Introduction to text-to-speech synthesis - TCTS Lab
18. Neumann L., Matas J.: Real-Time Scene Text Localization and Recognition, CVPR 2012 (Providence, Rhode Island, USA)

Index Terms

Computer Science
Artificial Intelligence
Keywords

Class-Specific Extremal Region, Head-mounted device, MSER(Maximally Stable Extremal Regions), Raspberry Pi, Tesseract OCR, Probabilistic Hough Lines Transformation