MASTER RESEARCH INTERNSHIP PROPOSAL

Year 2018

Laboratoire L3i

Internship Topic:

Applications of Scene Text Detection/Recognition for the Visually Impaired and/or Tourist Guidance

Executive Summary of the Proposed Topic:

Text appears everywhere in our natural surrounding environments such as in traffic signs, license plates, advertisement billboards, business cards, building signs, labels on posted parcels and on name plates. People who suffer from visual impairment face challenges reading the text in all those situations, whether reading such text is a necessity for their everyday life or simply for navigating and/or enjoying the world around them. **We propose to develop an app for mobile and/or wearable devices to assist the visually impaired by narrating the text in their surrounding natural scenes.**

This internship would produce a complete working pipeline of a mobile scene text accessibility system for the visually impaired. The main modules of this pipeline of a smartphone app are: video capture, preprocessing and frame selection, text detection, text recognition (OCR), text-to-speech and/or augmented reality. The app will be tested on Android smartphones and smart-glasses.

Key Words:

Scene text detection, scene text recognition, OCR, text accessibility for the visually impaired, Android, wearable devices, mobile-captured images

Context of the Internship Subject:

The research topic of this internship will rely on- and take forward the research already being conducted in the L3i on the problems of scene text detection & recognition and video capture on mobile devices.

The researchers proposing this internship have experience working on the different modules of the proposed pipeline, and they are actively working on the topics of scene text detection, mobile app development for imaging applications, OCR and video capture since few years.

With this internship, we target to put our research work into a useful application for the visually impaired people. There are also other possible applications in tourist guidance and in e-education. Figure 1 shows how such an app would help people to read/hear scene text clearly in their surrounding environment.

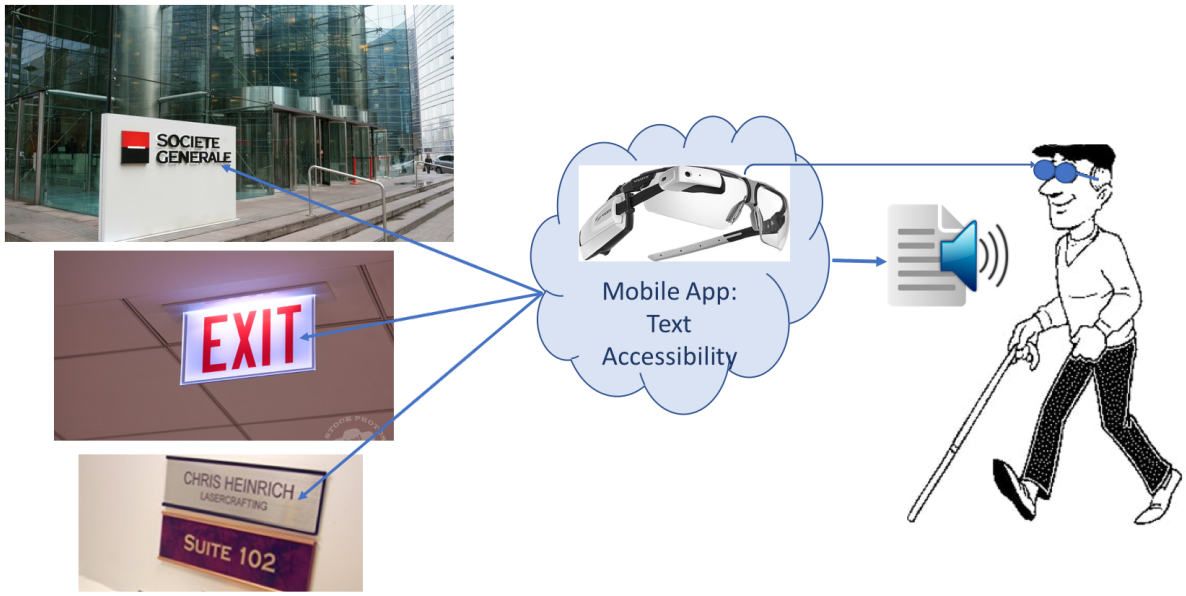

Figure 1: Text accessibility in natural scenes for the visually impaired.

Description of the Internship Subject:

The main goal of this internship is to develop an app for mobile and/or wearable devices to assist visually impaired people by narrating the text appearing in their surrounding natural scenes.
Visually impaired people face challenges when they want to read text in everyday life situations such as: labels on posted parcels, building signs, name plates in a medical cabinet, street signs or advertisements. Whether reading such text is a necessity for their everyday life or simply for navigating and/or enjoying the world around them, our proposed app would assist the visually impaired by narrating the text that appears in their surrounding natural scenes.
This internship would produce a complete working pipeline of a mobile scene text accessibility system for the visually impaired. The main modules of this pipeline of a smartphone app are: video capture, preprocessing and frame selection, text detection, text recognition (OCR), text-to-speech and/or augmented reality. The app will be tested on Android smartphones and smart-glasses.
The initial planning of the internship work is as follows:

- Month-1: Implementing the pipeline on smartphone + literature review on scene text detection
- Month-2: Implementing the pipeline on smartphone + literature review on scene text detection
- Month-3: Novel contribution on scene text detection
- Month-4: Novel contribution on scene text detection
- Month-5: Demo finalization + writing the internship report
- Month-6: (Optional) Writing an international workshop/conference paper

Note: In the pipeline, the modules of "text recognition" and "text-to-speech" will be integrated as off-the-shelf ready functions/system either from state-of-the-art or from previous internships.

Required Skills:

- Very good programming skills (in a high-level language such as java or python)
- Development in Android
- Able to read scientific literature in English

Preferable skills:

- Knowledge of image processing
- Programming in OpenCV

Complementary Information:

Supervisors: Nibal NAYEF (60%), Muzzamil LUQMAN (30%), Guillaume CHIRON (10%)
Laboratory: L3i – University of La Rochelle
Collaboration framework: ANR project "AUDINM"
Start Date: January or February 2018
Internship duration: 5 - 6 months
Salary: ~550 Euros/month

References:

- Nibal Nayef, Jean-Marc Ogier, « Semantic text detection in born-digital images via fully convolutional networks", ICDAR 2017.
- Nibal Nayef, Fei Yin, Imen Bizid, Hyunsoo Choi, Yuan Feng, Dimosthenis Karatzas, Zhenbo Luo, Umapada Pal, Christophe Rigaud, Joseph Chazalon, Wafa Khlif, Muhammad Muzzamil Luqman, Jean-Christophe Burie, Cheng-lin Liu and Jean-Marc Ogier, "ICDAR2017 Robust Reading Challenge on Multi-lingual Scene Text Detection and Script Identification – RRC-MLT", ICDAR 2017
- Joseph Chazalon, Petra Gomez-Krämer, Jean-Christophe Burie, Mickaël Coustaty, Sébastien Eskenazi, Muzzamil Luqman, Nibal Nayef, Marçal Rusiñol, Nicolas Sidère and Jean-Marc Ogier. SmartDoc 2017 Video Capture: Mobile Document Acquisition in Video Mode, ICDAR-OST 2017.
- JC. Burie, J. Chazalon, M. Coustaty, S. Eskenazi, M.M. Luqman, M. Mehri, N. Nayef, JM. Ogier, S. Prum and M. Rusinol, "ICDAR2015 Competition on Smartphone Document Capture and OCR (SmartDoc)", ICDAR 2015.
- Guillaume Chiron, "Système complet d'acquisition vidéo, de suivi de trajectoires et de modélisation comportementale pour des environnements 3D naturellement encombrés : application à la surveillance apicole », Thesis 2014.
- https://www.youtube.com/watch?v=gtGAcqjebyI&list=PL_SHh-VkO8VI3X5T6-YdeAFetKjTK_tDN&index=11
- http://soundpaper.com/markets/visually-impaired/
- https://www.hackster.io/tushar-chugh/smart-cap-vision-for-the-visually-impaired-ca0ea1

Contacts:

Emails: mluqma01@univ-lr.fr , nibal.nayef@univ-lr.fr