Travail de Master / 2023

Computer Science - Cybersecurity

Sound Source Localization and Distance Estimation in Open Environment using Simulation and AI

Specification

22.02.2023 – Version 0.1

Denis Rosset

Supervisors: **Michael Mäder**
Beat Wolf

Principal: **ROSAS**

Versions table

Version	Publication Date	Author	Description
0.1	22.02.2023	Denis Rosset	Draft of specifications
0.2	27.02.2023	Denis Rosset	Objectives and task description

Table des matières

1   Introduction

Within the framework of the research project "NPR Teleoperation", the engineers of the HEIA-FR have developed the first concept in Switzerland of a remote-controlled automated vehicle. However, teleoperation only makes sense if the vehicle is automated. There can be no teleoperation without automation (economic factors) just as there can be no automation without teleoperation (legal, technical, and social factors). ROSAS then created the Autovete (Automatisation de véhicules téléopérés) project, financed by HEIA-FR, to build up vehicle automation expertise.

For a vehicle to be fully autonomous, the detection of other emergency vehicles is mandatory. To solve this issue, V2V (Vehicle-to-Vehicle) communication can be used but is not currently used by the emergency vehicles. So, to be able to detect such emergency vehicle, two signals need to be processed: the sound of the emergency siren and the blinking lights of the vehicle. This project focusses only on the sound source distance estimation and localization.

Big improvements in sound source localization with the help machine learning are being made[1] and can be used to reliably localize the origin of a sound using one or more microphone array (multiple microphones operating in tandem).

A non-negligible problem is that the number of real-world datasets with moving sources in open environment is limited or non-existing. A solution is to create the datasets in realistic sound propagation simulation.

To validate and use the model, it should also be tested to see how it react to adversarial attacks, understand how it can be used in a real environment and limit the attack vector, for finally improve the reliability of the system.

Figure 1 Perceptin: An Autonomous Vehicle

The project aims to:

- Create a realistic sound propagation simulation with moving sources in open environment

- Train the model on the created dataset

- Test the model in a real-world environmentActors

The following actors are part of the project:

- Denis Rosset, Computer Science student, MSE

---

[1] A SURVEY OF SOUND SOURCE LOCALIZATION WITH DEEP LEARNING METHODS (https://arxiv.org/pdf/2109.03465.pdf)

- Michael Mäder, Professor in Computer Science at HEIA-FR, Supervisor
- Beat Wolf, Professor in Computer Science at HEIA-FR, Supervisor

2 Objectives

2.1 Getting a dataset

Getting a dataset is the first and main part of the project. Since no datasets with sound of a vehicle and its relative position in open environment exists, the dataset needs to be created. The dataset of sounds and relative position can then be used to train a neural network. The dataset needs to have a real-life use case so it can be compared later in the project. To achieve that, a baseline needs to be established (i.e. syren sensor, sound recordings) to compare results. Sound spatialization simulation software can be used to generate data in 3D environment.

2.2 Creating a model

Creating a model to predict the position of the source of a sound is the second part of the project. The architecture of the model needs to be based on state-of-the-art solution to try getting the most out of the dataset. The dataset needs to be split into multiple smaller datasets (test, train, validate, etc.) to help understanding how the model perform after its training.

2.3 Testing the model

Testing the model in a real-world environment is the third part of the project. With the help of microphone arrays, the model should be tested with real-world data to see how it perform in real-world condition. The results can then be compared to the baseline to see if the model is performing as expected.

2.4 Validating the model

The third part of the project is to validate the model, the baseline established in the first part of the project is needed. It allows to compare what the system can do in real life condition.

To ensure the robustness of the system, an analysis of the potential of adversarial attacks against it needs to be done. Propositions of improvements of the model to make it better and more secure should be put down. It should also be tested to see how it react to environmental changes (i.e. more open environment, more echo in the street, difference in sound propagation).

2.5 Improvement of the system

Improving the model is the fifth part of the project. Once the model is tested and validated, it should be improved to increase its accuracy. Different techniques can be applied to the model to achieve better results (i.e. transfer learning, data augmentation, hyperparameter tuning, etc.)

3 Tasks

The tasks are divided in three main categories. Data generation, model creation and model validation.

3.1 Dataset generation

Task: Generate a dataset of sounds and its relative positions in open environment.

- Research and creation of datasets and sound simulation software
- Definition of the baseline
- Create 3D models of different environments
- Generate the dataset using the sound spatialization software
- Evaluate the dataset and make sure it is suitable for training

3.2   Model creation

Task: Create a model to localize the sound source.

- Research and study of existing models
- Design and implementation of the model
- Train the model on the generated dataset
- Evaluate the model and compare its accuracy

3.3   Model validation

Task: Validate the model against adversarial attacks and environmental changes

- Test the model against real-world data
- Evaluate the model against environmental changes
- Compare the results against the baseline
- Research and study of adversarial attack that could be applied to this project
- Test the model against the adversarial attack
- Find ways to improve the model and limit attack vector

3.4   Documentation

Task: Write a report that summarize the work done and present the results

- Write a report that summarize the work done
- Presentation of the results

4   Key dates

- 20.02.2023   Start of the Master Thesis
- 07.04.2023   Milestone 1 project review (first red line)
- 26.05.2023   Milestone 2 project review (second red line)
- 07.07.2023   Report deposit
- [14-25].08.2023 Presentation

5 Planning

Task	1	2	3	4	5	6	7	EASTER HOLIDAYS	8	9	10	11	12	13	14	15	16	17	18	19
Documentation	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
Specification and planning realisation	X	X	X																	
Report			X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
Dataset Generation	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
Research of existing methods of sound localization	X	X	X																	
Definition of the baseline		X	X	X					X	X										
Research and creation of baseline datasets and sound simulation software			X	X	X					X	X				X	X				
Creation 3D models of different environments				X	X	X				X	X				X	X				
Generate a dataset using the sound spatialization software					X	X				X	X				X	X				
Evaluate the dataset and make sure it is suitable for training					X	X				X					X	X	X			
Model Creation	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
Research and study of existing models					X	X					X	X			X	X	X			
Design and implementation of the model					X	X						X	X		X	X				
Train the model on the generated dataset						X	X					X	X			X	X			
Evaluate the model and compare its accuracy						X	X					X	X			X	X			
Model validation	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
Test the model against real-world data							X				X				X	X				
Evaluate the model against environmental changes							X		X		X						X			
Compare the results against the baseline							X		X		X		X					X	X	X
Research and study of adversarial attack that could be applied to this project												X	X					X		
Test the model against the adversarial attack													X	X				X		
Find ways to improve the model and limit attack vector														X				X	X	X