Haute école d'ingénierie et d'architecture Fribourg
Hochschule für Technik und Architektur Freiburg

Travail de Master / 2023

Computer Science - Cybersecurity

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

Specification

03.02.2023 – Version 1.0

Denis Rosset

Supervisors: **Michael Mäder**
Beat Wolf

Principal: **ROSAS**

Hes·so
Haute Ecole Spécialisée de Suisse occidentale
Fachhochschule Westschweiz

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
0.3	28.02.2023	Denis Rosset	Planning and milestones
0.4	03.03.2023	Denis Rosset	Modification of the tasks and objectives paragraphs to better suit their purpose Rewriting of some paragraphs Modification of the planning + add of the concept of milestones and epics
0.5	13.03.2023	Denis Rosset	Added a use case for the detection of excessively noisy vehicle in the introduction and rewrote some sentences

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 yet integrated on emergency vehicles. So, to be able to detect such vehicle, two signals need to be processed: the sound of the emergency siren and the blinking lights of the vehicle. A first use case of this project focusses only on the sound source distance estimation and localization.

Figure 1 Perceptin: An Autonomous Vehicle

To understand if the sound source estimation and localisation could work for the emergency detection, a simpler use case has been created for this project. It is the detection of excessively noisy vehicle on the street. The goal is to measure the sound level of the passing vehicles and compare it with the legal limits. If a vehicle exceeds the limit, the system can record its license plate and report it to the authorities. This way, the system can help reduce noise pollution and improve road safety.

To implement this use case, the system requires a microphone array, a camera, and a processing unit. The microphone array captures the sound signals from different directions and sends them to the processing unit. The processing unit applies a sound source localization algorithm to estimate the direction and distance of the sound source. The camera captures the image of the vehicle and performs license plate recognition.

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. 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 against adversarial attacks, understand how it can be used in a real environment and limit the attack vector.

2 Actors

The following actors are part of the project:

- Denis Rosset, Computer Science student, MSE
- Michael Mäder, Professor in Computer Science at HEIA-FR, Supervisor
- Beat Wolf, Professor in Computer Science at HEIA-FR, Supervisor

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

3 Objectives

3.1 Objective n°1 Dataset according to the baseline

The first objective is to have a dataset constructed. It needs to be coherent with the baseline of the project and should help to create and understand the problem. The dataset should contain the target variable, the features, and the necessary pre-processing steps, such as missing data imputation, data normalization, feature engineering, etc.

3.2 Objective n°2 Model for better sound source localization and distance estimation

The project should use a neural network model to detect the origin of a sound using a microphone array. The neural network should be trained using the dataset created and should be able to accurately localize the sound source. The trained neural network model should be evaluated to see how it performs in a real environment. The model should also be evaluated to see how dependable it is in localizing sound sources and how it can be improved.

3.3 Objective n°3 The model should resist to attacks

The trained neural network model needs to be evaluated to see how it reacts to adversarial attacks. This should be done by testing the model on data that has been modified in some way, such as by adding or removing noise, or by modifying the sound source. The model could also be tested against various types of attacks, such as masking, time-warping, and frequency-shifting. The model should be able to accurately localize the sound source even when it is attacked. The model should also be as robust as possible to come back to a normal operating state once the attack is over.

4 Tasks

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

4.1 Dataset generation

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. siren sensor, sound recordings) to compare results. Sound spatialization simulation software can be used to generate data in 3D environment. The subtasks are:

- 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

4.2 Model creation

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. The subtasks are:

- 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

4.3 Model validation

Evaluating 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 performs in real-world conditions. The results can then be compared to the baseline to see if the model is performing as expected. 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 reacts to environmental changes (i.e. more open environment, more echo in the street, difference in sound propagation). The subtasks are:

- Evaluate 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 (and how it reacts after the attack)
- Find ways to improve the model and limit attack vector

4.4 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

5 Key dates

- 20.02.2023 (SP1) Start of the Master Thesis
- 15.03.2023 (SP4) **Milestone 1** Baseline definition
- 22.03.2023 (SP5) **Milestone 2** Dataset review
- 29.03.2023 (SP6) **Milestone 3** Model review
- 07.04.2023 (SP7) **EPIC 1 + Milestone 4** project review + model validation against baseline
- 28.04.2023 (SP9) **Milestone 5** Dataset review
- 05.05.2023 (SP11) **Milestone 6** Model review (SP 11)
- 17.05.2023 (SP12) **Milestone 7** Validation review
- 26.05.2023 (SP13) **EPIC 2 + Milestone 8** project review + adversarial attack potential
- 09.06.2023 (SP15) **Milestone 9** Dataset review
- 14.06.2023 (SP17) **Milestone 10** Model review
- 23.06.2023 (SP18) **Milestone 11** Model validation + adversarial resistance
- 07.07.2023 (SP19) **EPIC 3** Report deposit
- 14-25.08.2023 Presentation

6 Planning

Week	1	2	3	4	5	6	7		8	9	10	11	12	13	14	15	16	17	18	19
Documentation																				
Specification and planning realization																				
Report																				FR
Dataset Generation																				
Research of existing methods of sound localization																				
Definition of the baseline				1. WE																
Research and creation of baseline datasets and sound simulation software																				
Creation 3D models of different environments																				
Generate a dataset using the sound spatialization software								EASTER HOLIDAYS												
Evaluate the dataset and make sure it is suitable for training					2. WE					5. FR						9 . FR				
Model Creation																				
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. WE						6. FR						10. WE		
Model validation																				
Test the model against real-world data																				
Evaluate the model against environmental changes													7. WE							
Compare the results against the baseline							4. FR													
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														8. FR					11. FR	