

2024年臺灣國際科學展覽會 優勝作品專輯

作品編號

190014

參展科別

電腦科學與資訊工程

作品名稱

噪音之眼：結合聲源定位與影像辨識的聲音照相系統

就讀學校

臺北市立第一女子高級中學

指導教師

陳伶志、陳怡芬

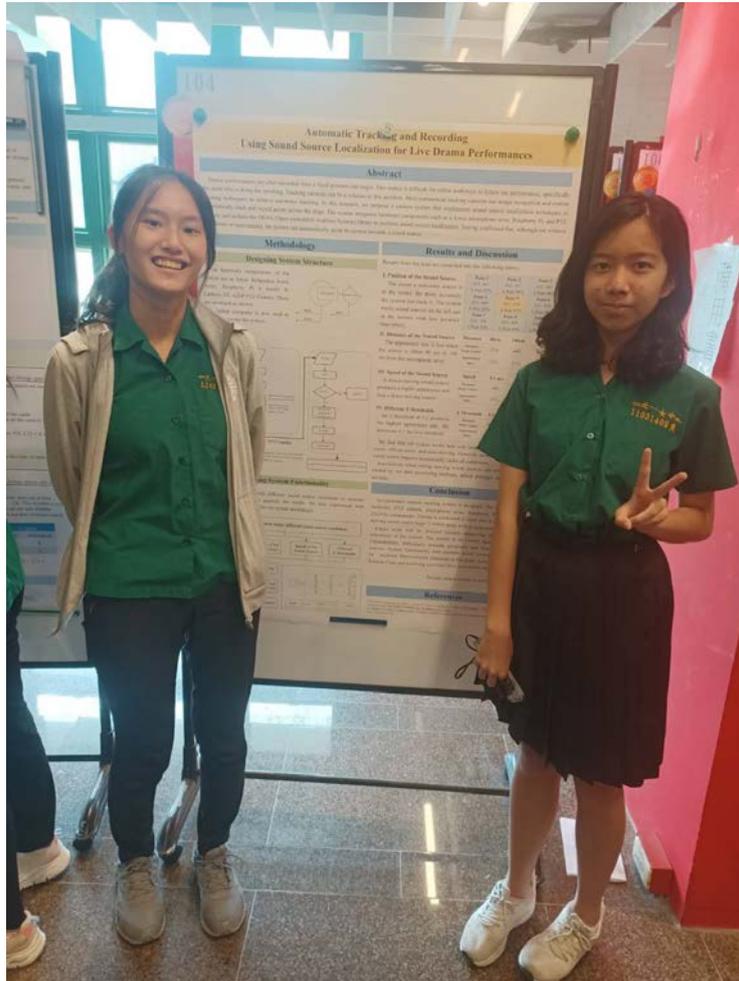
作者姓名

蔡予曦、林采穎

關鍵詞

聲源定位、影像辨識、聲音照相系統

作者簡介



我們是北一女中高三的蔡予曦和林采穎，非常高興能夠參加這一次的臺灣國際科展。我們在加入數理資優班後開始進行資訊領域的專題研究，在兩年半的研究歷程中，我們彼此合作、互相扶持，體驗做研究的過程，收穫了許多技術、知識、及寶貴的回憶。我們也十分感謝我們的指導老師陳怡芬及指導教授陳伶志一路上的教導、指引、及陪伴。

摘要

汽、機車呼嘯而過的排氣噪音經常造成周遭民眾的困擾，為杜絕這種狀況，環保署於2021年開始推動「聲音照相科技執法」，透過包含魚眼攝影機、車牌辨識攝影機等設備的整合，來紀錄行經車輛產生的噪音是否超過管制標準，不過此系統目前仍存在一些限制，如：當多台車輛同時經過時無法辨別確切的噪音來源，而需要人工判定。本研究旨在利用聲源定位原理結合影像辨識技術，實作出一套能夠追蹤聲源（噪音）並辨識聲音來源（車輛）的聲音照相系統，期日後能實際應用並解決道路上多車輛時無法辨識噪音來源的困擾。研究中提出一套能透過遠端控制程式，將收音裝置、攝影設備和資料分析軟體等各項軟硬體設備，串聯成自動化追蹤聲源的系統設計方案。

Abstract

The exhaust noise generated by passing modified vehicles often negatively impacts the quality of life of the surrounding community. To address this issue, the Environmental Protection Agency initiated the "Sound Photography Technology Enforcement" in 2021. This approach involves the integration of equipment such as fisheye cameras and license plate recognition cameras to record whether the noise generated by passing vehicles exceeds regulatory standards. However, this system currently faces certain limitations, such as its inability to precisely identify the source of noise when multiple vehicles pass simultaneously, necessitating manual assessment.

The objective of this study is to employ the principles of sound source localization combined with image recognition technology to develop a noise monitoring system capable of tracking and identifying the sound source (the vehicle). The aim is to address the issue of being unable to identify the noise source when multiple vehicles are present on the road. This study introduces a system design that can remotely control various hardware and software components, including microphone arrays, cameras, and data analysis software, to create an automated sound source tracking system.

壹、前言

一、研究動機與背景

我們觀察到改裝汽、機車呼嘯而過的排氣噪音經常造成周遭民眾的困擾及抱怨。進一步研究之後，我們發現依「道路交通管理處罰條例」第四十三條規定，汽車拆除消音器或以其他方式造成噪音，可處六千元以上、三萬六千元以下罰鍰[1]，而根據「機動車輛噪音管制標準」，超過標準值(限速小於 50km/h 之路段為 86 分貝，限速大於 50km/h 但小 70km/h 為 90 分貝)的車輛，能處 1800 元至 3600 元的罰鍰[2]。以往都是依靠人工舉報來執行，而在 2021 年，環保署開始推動「聲音照相科技執法」，來紀錄行經車輛產生的噪音是否超過管制標準，不過此系統目前仍存在一些限制，如：當多台車輛同時經過時無法辨別確切的噪音來源，而需要人工判定。因此，我們想要時做出能在畫面中有多台車輛辨別是哪一台發出噪音的系統，以增加執法的效率。

二、研究目的

本研究旨在利用聲源定位原理結合影像辨識及物件偵測追蹤技術，實作出一套能夠追蹤聲源（噪音）並辨識聲音來源（車輛）的聲音照相系統，期日後能實際應用並解決道路上多車輛時無法辨識噪音來源的困擾。

三、文獻回顧

(一)聲音照相科技執法與其技術限制

環保署於 110 年 1 月 1 日起正式執行聲音照相科技執法，透過包含噪音計、風速計、魚眼攝影機、車牌辨識攝影機、控制箱體等設備的整合，來紀錄行經車輛產生的噪音是否超過噪音管制標準，高噪車只要行駛經過分貝超過標準就會被自動拍攝，並紀錄分貝值。[3]但因現行科學儀器技術仍有限制，當照片含兩輛車以上時，仍需以噪音歷時圖及影像進行辨別。若系統能搭配陣列式麥克風可分辨高噪音車輛，或經人工判斷有不當改裝排氣管或明確不當操作駕駛行為之情事時，可通知該等車輛到檢。[4]

(二)ODAS: Open embedeD Audition System

為了能夠準確地辨別噪音源，本研究中準確的聲源定位置關重要。我們使用 ODAS: Open embedeD Audition System [5] 來在 Raspberry Pi 上進行聲源定位。ODAS 之所以被選用，主要因其在計算上相對於其他聲源定位系統有更高的效率，考量到我們使用的 Raspberry Pi 的功能性，我們認為 ODAS 是計算音源方位的最佳選擇。

接下來將簡單敘述 ODAS 系統的定向原理。ODAS 系統使用 SRP-PHAT 演算法來估算聲源方向(Direction of Arrival, DOA)。SRP-PHAT (Steered Response Power with Phase Transform) [6] 是一種在遠場模型下使用濾波求和波束成形來估計聲源方向的演算法。

濾波求和波束成形中，每個訊號 X_m 經過延遲 Δ_m 及濾波 G_m 的處理，之後進行加總求和。每一組延遲對應單位球面上的一個離散點，計算出最大 SRP-PHAT 值的那組延遲所對應的點即為計算出聲源的方位。

我們使用的麥克風陣列有四個麥克風，因此有四個訊號。SRP-PHAT 的過程可表示為：

$$Y(\omega, \Delta_1, \Delta_2, \Delta_3, \Delta_4) \equiv \sum_{m=1}^4 G_m(\omega) X_m(\omega) e^{-j\omega\Delta_m} \quad \text{式 1}$$

$$P(\Delta_1, \Delta_2, \Delta_3, \Delta_4) \equiv \int_{-\infty}^{\infty} Y(\omega, \Delta_1, \Delta_2, \Delta_3, \Delta_4) Y^*(\omega, \Delta_1, \Delta_2, \Delta_3, \Delta_4) d\omega \quad \text{式 2}$$

其中濾波 G 即為 PHAT(Phase Transform)，會使音頻訊號僅保留相位資訊，可以降低混響環境對準確度的影響。

為了降低計算量，ODAS 採用以下策略。首先，在一個包含 162 個離散點的單位球面上，計算 SRP-PHAT 值，找到具有最大 SRP-PHAT 值的點。接著，在另一個包含 2562 個離散點的單位球面上，計算鄰近該點各點的 SRP-PHAT 值，並確定具有最大 SRP-PHAT 值的點，即被認為是聲源最有可能的方向。

(三)YOLO(You Only Look Once)

為了能在影像上追蹤噪音源，我們使用了機器學習訓練而成的物件偵測及追蹤系統。YOLO(You Only Look Once) [7] 是一種常用的影像辨識及物件偵測模型架構。YOLO 的運作原理是將圖像分成固定數量的網格，每個網格負責檢測其中的物件，預測物件的類別及位置。這種設計使得 YOLO 能夠一次監測多個物件，而不需要多次處理影像，故名 You Only Look Once。我們所使用的版本為 Ultralytics 的 YOLOv8-nano [8]。v8 為 YOLO 的最新版，考量到筆記型電腦的性能，選用耗能最小的 nano 模型。

貳、研究方法或過程

本研究透過以下兩項主要的任務階段以實作出一套能夠追蹤聲源（噪音）並辨識聲音來源（車輛）的聲音照相系統。

一、系統設計與實作

(一)硬體設備及配置

本研究實作系統硬體設備使用筆記型電腦一部、微型單板電腦一組、麥克風陣列一組、及 360 度攝影機一部，並以下圖示方式設置：

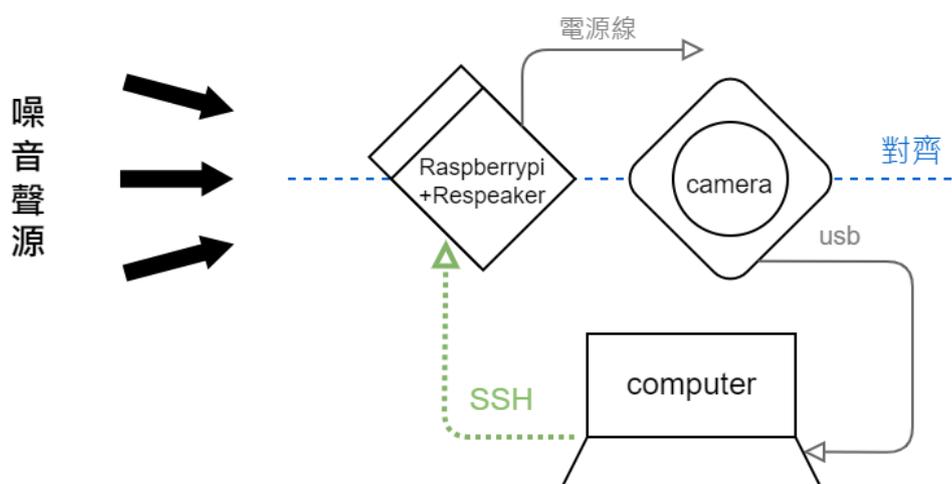


圖 1 系統硬體設備配置示意圖

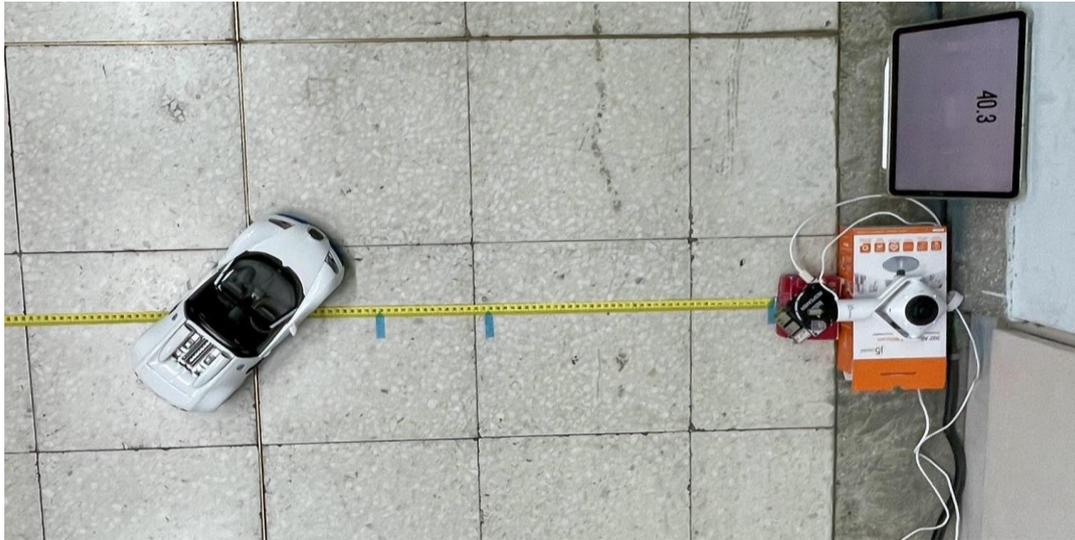


圖 2 系統硬體設備配置照片

其具體規格與功能說明如下：

1. 筆記型電腦 ACER Swift 3

配備 i5 處理器和 8GB 記憶體，主要用於執行控制程式以及進行 YOLO 影像辨識及追蹤。

2. 微型單板電腦 Raspberry Pi 4 Computer Model B 8GB RAM

體型小方便實驗中裝設。透過 SSH 連線由筆記型電腦遠端控制，其上裝置麥克風陣列以接收音頻輸入，並執行聲源定向。



圖 3 微型單板電腦 Raspberry Pi

3. 麥克風陣列 ReSpeaker 4-Mic Array for Raspberry Pi

配備四個麥克風的麥克風陣列。裝置在微型單板電腦上，用於接收四個音軌的聲音。

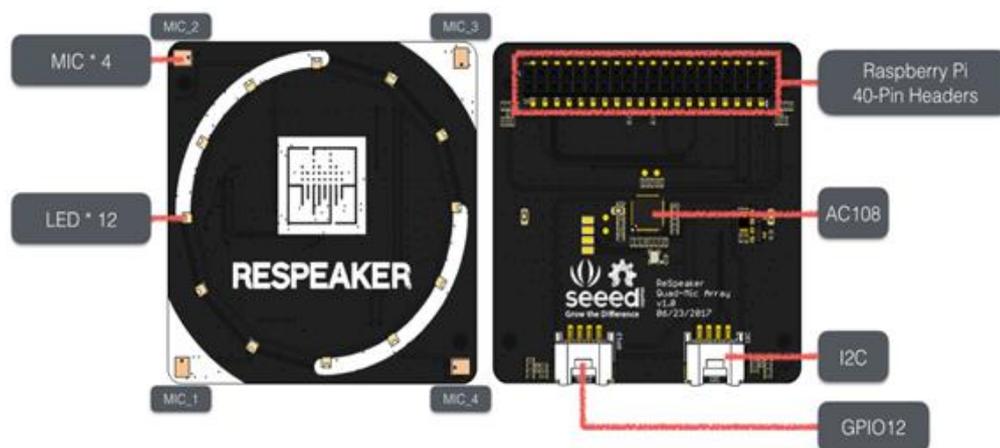


圖 4 麥克風陣列 ReSpeaker 4-Mic Array [9]

4. 360 度攝影機 j5 create JVCU360

以 USB 接頭與筆記型電腦連接。捕捉全方位 360 度影像，以進行影像辨識及追蹤。因追蹤目標相對實作系統的角度會有巨幅變化，故選用 360 度影像的攝影機。



圖 5 360 度攝影機 j5 create JVCU360

(二) 自動化控制程式設計

本系統的控制程式部分採用 Python 3.9 進行開發。為達到研究目的，本研究中設計的系統分為兩個階段及三個主要部分：同步進行聲源定位及物件偵測追蹤，接著整合兩邊資料進行資料處理並判斷有無重疊。程式流程的示意圖如下：

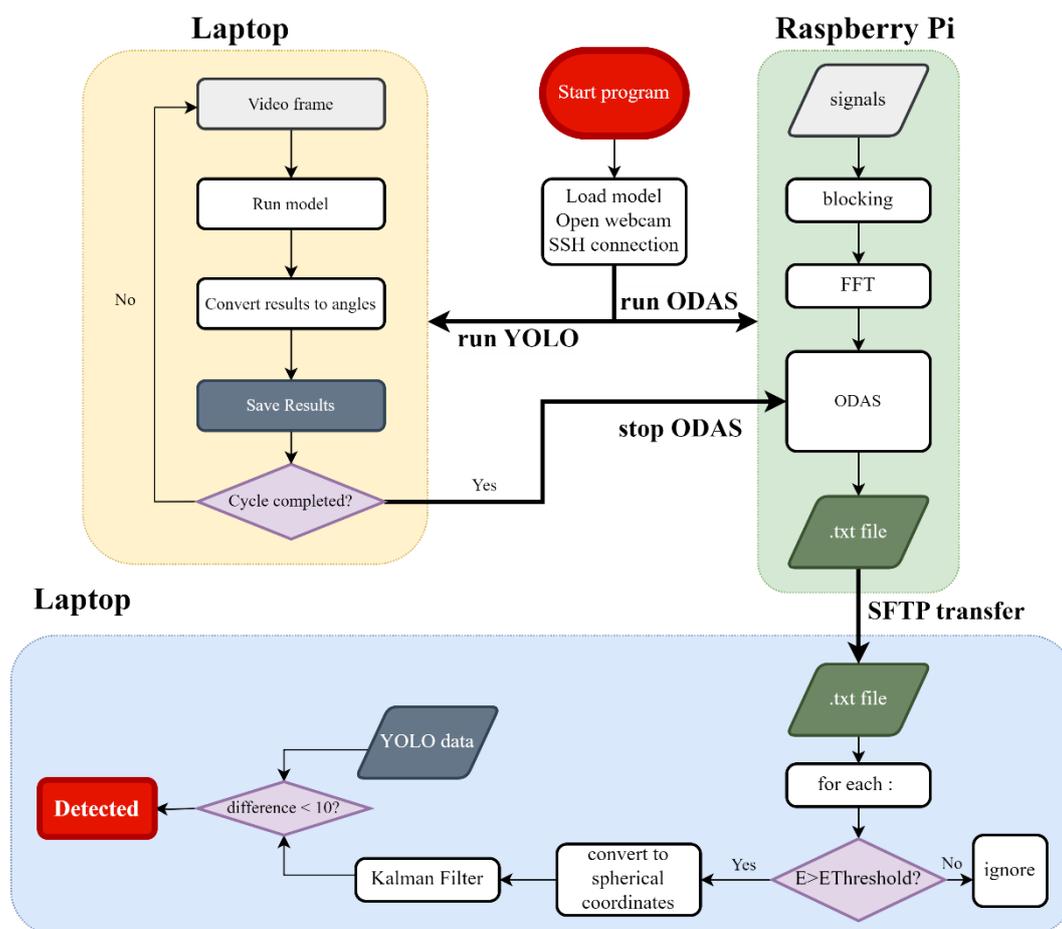


圖 6 控制程式內部流程圖

接下來將分前述的三大部分詳細解釋實作系統的運行流程及細節：

1. 聲源定位

(1) 建立筆記型電腦與 Raspberry Pi 的 SSH 連線

由於 ODAS 在 Raspberry Pi 上運行，我們必須確保我們的筆記型電腦能夠遠端登入並控制 Raspberry Pi。安全外殼協議 (SSH) [10] 是一種網路傳輸協議，用於建立客戶端 (在本研究中為筆記型電腦) 與伺服器 (在本研究中是 Raspberry Pi) 之間的連線。

我們使用 SSH 從筆記型電腦通過遠端控制 Raspberry Pi 來執行 ODAS，並利用 Paramiko 模組來在 Python 程式中執行連線。為了建立連線，我們首先需要將 Raspberry Pi 連接到無線網絡並獲取其 IP 地址。

(2)執行並停止 ODAS 軟體

我們透過先前建立的 SSH 連線，在程式中從電腦對 Raspberry Pi 下達指令，使 Raspberry Pi 運行 ODAS。同時，我們使用在 Raspberry Pi 上裝置的麥克風陣列收音，作為供 ODAS 分析音源方向的音頻訊號。在運行 ODAS 時，除了提供音頻訊號之外，還須輸入一個配置檔（config 檔），檔案內包括 ReSpeaker 的資料及定向結果的輸出格式等設定。

ODAS 分析音頻訊號後會輸出 SSL(Sound Source Localization)資料，為對音檔中的每一個 timeStamp 的四個音源可能的方向。一秒的音頻資料大約會生成 172 個 timeStamp。資料的格式為多個 JSON 資料，具體如下圖所示。

```
{
  "timeStamp": 1,
  "src": [
    { "x": 0.000, "y": 0.000, "z": 1.000, "E": 0.000 },
    { "x": 0.000, "y": 0.000, "z": 1.000, "E": 0.000 },
    { "x": 0.000, "y": 0.000, "z": 1.000, "E": 0.000 },
    { "x": 0.000, "y": 0.000, "z": 1.000, "E": 0.000 }
  ]
}
{
  "timeStamp": 2,
  "src": [
    { "x": 0.972, "y": -0.104, "z": 0.209, "E": 0.154 },
    { "x": 0.067, "y": 0.669, "z": 0.740, "E": 0.156 },
    { "x": 0.133, "y": 0.589, "z": 0.797, "E": 0.137 },
    { "x": 0.201, "y": -0.957, "z": 0.209, "E": 0.109 }
  ]
}
{
  "timeStamp": 3,
  "src": [
    { "x": 0.161, "y": -0.221, "z": 0.962, "E": 0.224 },
    { "x": -0.951, "y": -0.138, "z": 0.276, "E": 0.101 },
    { "x": -0.758, "y": 0.637, "z": 0.140, "E": 0.097 },
    { "x": -0.773, "y": -0.092, "z": 0.628, "E": 0.078 }
  ]
}
```

圖 7 SSL 資料擷圖

四行各代表一個聲源可能的方向。要停止運行 ODAS 時，我們讓這個檔案以普通文字檔（.txt）的格式輸出。並找出運行中 ODAS 的進程 ID 以停止運行 ODAS。

(3)SSL 資料檔 SFTP 傳輸至筆記型電腦

從 ODAS 得出 SSL 資料後，我們透過 SFTP 協定將資料從 Raspberry Pi 傳輸到筆記型電腦。SFTP (SSH 檔案傳輸協定, SSH File Transfer Protocol) 是 SSH 的擴充，提供檔案存取與傳輸的功能。在本研究中與 SSH 同樣透過 Paramiko 模組執行。

2. 物件偵測及追蹤

(1)在影像中的每一幀上進行物件偵測及追蹤

首先，我們使用 OpenCV 從 360 度攝影機上讀入影像，並將該影像輸入 YOLOv8-nano 模型進行處理。我們限制模型只偵測特定的物件種類 (車子)。



圖 8 YOLO 物件偵測示例

(2)將結果轉換為方位並儲存

我們將所追蹤物件在影像上的 x 座標轉換成角度並與當下時間儲存在陣列中，以便之後資料整和處理。

3. 資料整合及結果輸出

在 YOLO 已運行先前設定的次數後，發出指令停止 ODAS 並接收回傳的資料檔，用以進行資料的整合及結果輸出。

(1)分析 SSL 資料檔得出音源方位角及俯仰角

我們將傳輸到電腦的資料檔讀入程式，並利用 JSON 格式括號的特性切割不同的 timeStamp。一個 timeStamp 中包含了 4 行資料，本研究只取 E 值通常最大的第一行。一行中有四種資料 x, y, z, E。其中 x, y, z 為可能聲源方向在單位球面上的笛卡爾座標，E 值則代表了 ODAS 對那個可能聲源方向的強度，值在 0 和 1 之間。我們以 E 的值篩選資料，只取 E 大於 0.2 的值，並將笛卡爾座標轉換成球座標。

(2)將 SSL 資料套入卡爾曼濾波器

我們採用卡爾曼濾波器[11]來對具有不可避免的雜訊及誤差的數據進行處理。卡爾曼濾波器能夠適應系統和測量誤差的變化，有效的在雜訊和不確定性下估計系統狀態，被廣泛應用於許多領域。

卡爾曼濾波器會在時刻 k-1 儲存系統狀態的角度和角速度，用符號 x 表示，並通過乘以 F 矩陣來預測時刻 k 的系統狀態。接著在獲取測量值 z 後使用通過變異數計算得出的卡爾曼增益 K 將 z 與 x 加權平均，從而得出當前系統狀態的估計值。示意圖如下：

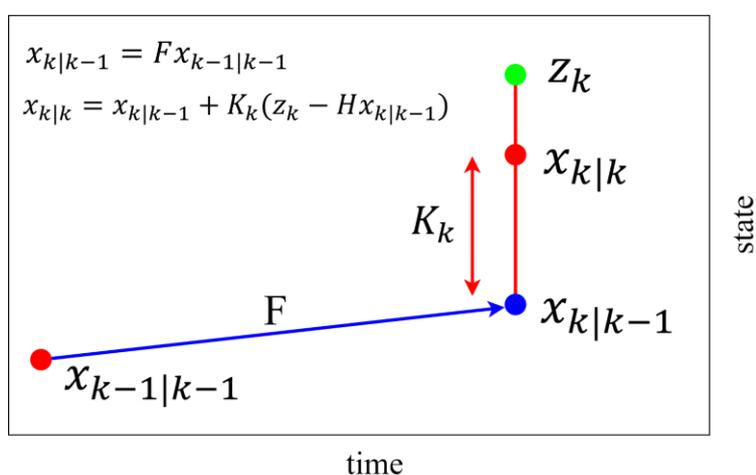


圖 9 卡爾曼濾波器原理示意圖

(3)比較 SSL 資料及影像辨識資料，並輸出結果

在迴圈中，對 ODAS 和 YOLO 的檢測結果進行對比。若其時間戳接近，則進一步比較檢測到的目標的方向角度，若兩者之間的差異小於 10 度，則顯示著檢測到了為噪音源的物件（汽車），輸出相應的時間戳及檢測物件 ID。

```
time: 1698719225.8812575
detected id: 2 detected time: 0 1698719225.8812575
1
time: 1698719226.1934137
detected id: 2 detected time: 1 1698719226.1934137
2
time: 1698719226.4050002
detected id: 2 detected time: 2 1698719226.4050002
3
time: 1698719226.6969235
time: 1698719226.916143
detected id: 2 detected time: 4 1698719226.916143
4
time: 1698719227.197956
time: 1698719227.4016168
time: 1698719227.7156963
time: 1698719227.9189956
```

圖 10 輸出結果擷圖

二、系統實測

在結合軟體及硬體設備，利用程式串聯製作出一套追蹤噪音車輛的自動化系統後，我們接著透過模擬三種道路情境，包括不同的車輛移動狀況和發出的噪音分貝數，啟動系統並記錄 ODAS 和 YOLO 的資料重合率以實測此系統的性能，最終分析數據並討論結果。實測流程如下圖所示。

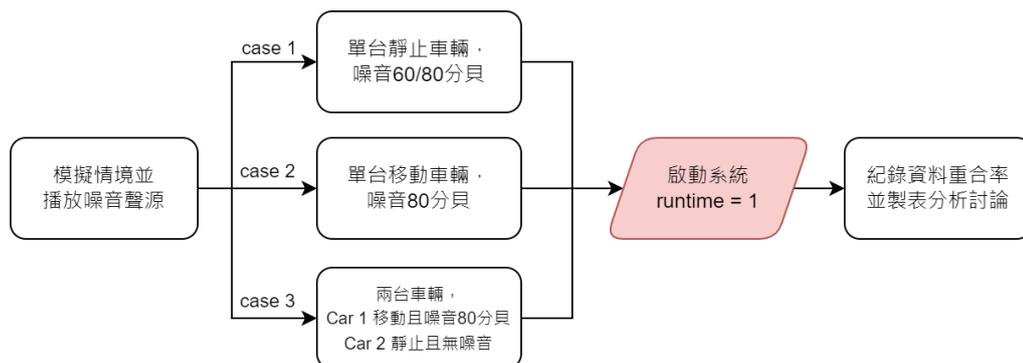


圖 11 實測流程圖

(一) 模擬環境並撥放噪音

為模擬道路上的實際情況，我們利用兩台模型車和噪音播放器（手機），並改變車輛的移動狀態、發出的噪音分貝數以及與 Respeaker 的距離，建構出以下三種情境：

1. 單台靜止車輛發出 60 及 80 分貝的噪音

為了模擬車輛在靜止狀態下發出噪音的情境，我們將噪音播放器置於靜止車輛上，依序將該台車輛放置在三個不同位置，分別距離 Respeaker 40、55、70 公分，並在每個距離下分別播放 60 分貝及 80 分貝的模擬噪音，各重複兩次，示意圖如下。

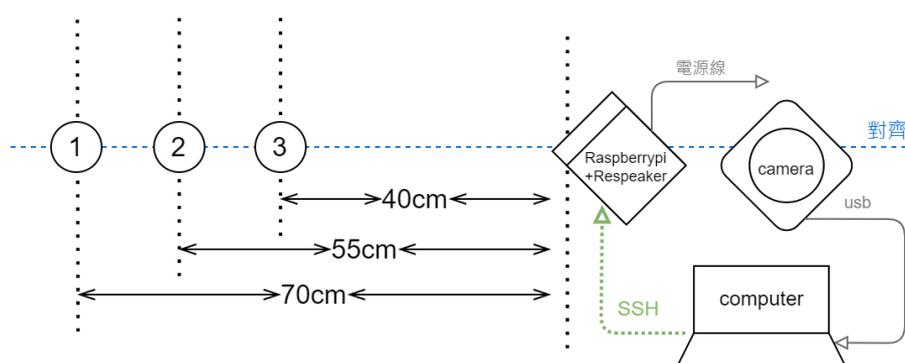


圖 12 第一種情境(單台靜止車輛)實驗示意圖

2. 單台移動車輛發出 80 分貝的噪音

為了模擬單一車輛在行進中發出噪音的情境，我們將噪音播放器置於移動車輛上，依序使該車輛沿三條路徑直線移動，在對齊線上分別距離 Respeaker 40、55、70 公分，並在每個距離下播放 80 分貝的模擬噪音，各重複兩次，示意圖如下。

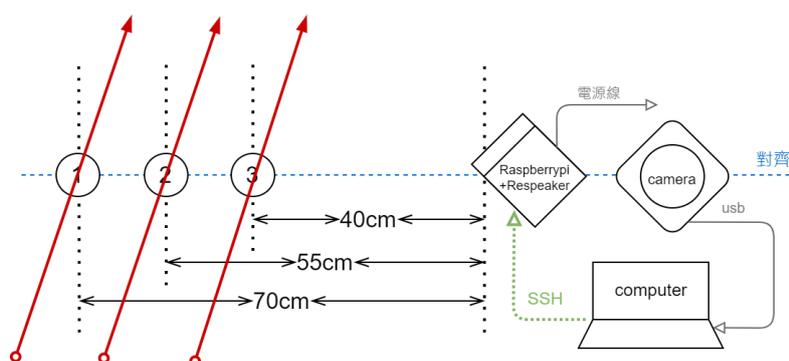


圖 13 第二種情境(單台移動車輛)實驗示意圖

3. 兩台車輛，第一台移動且噪音 80 分貝，第二台靜止且無噪音

為了單一車輛在車陣中發出噪音的情境，我們將噪音播放器置於一台移動車輛上，並放置另一台靜止且無噪音的車輛於固定位置（距離 Respeaker 55 公分處），依序使發出噪音的車輛沿三條路徑直線移動，在對齊線上分別距離 Respeaker 40、55、70 公分，並在每個距離下播放 80 分貝的模擬噪音，各重複兩次，示意圖如下。

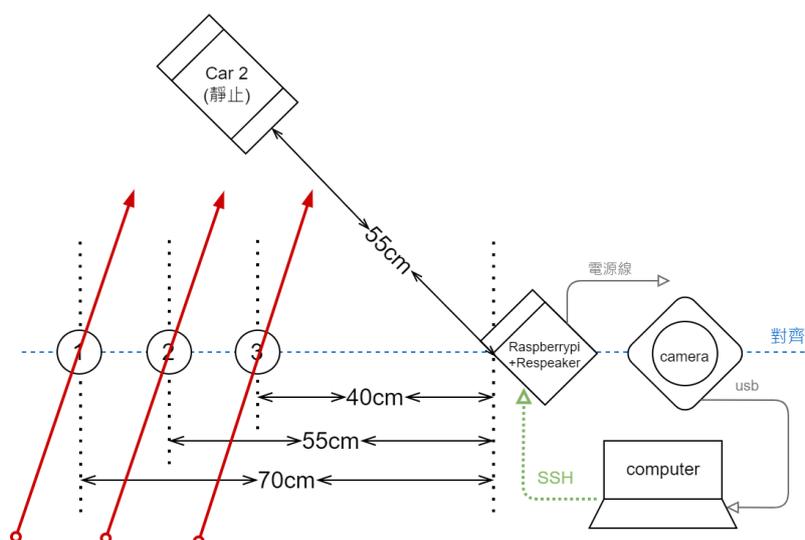


圖 14 第三種情境(兩台車輛)實驗示意圖

(二)紀錄資料重合率並製表分析

在模擬情境並啟動系統後，系統會輸出 ODAS 及 YOLO 分別偵測的角度對比後誤差 ≤ 10 度的時間戳個數，將該個數除以總時間戳個數即可得出資料的重合率，最後我們將數據平均後和原始數據圖表一併計入表格（參、研究結果中表 1、表 2）並分析討論，以比較在不同情況下的系統性能。

參、研究結果與討論

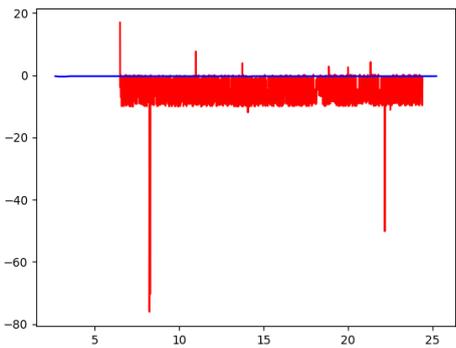
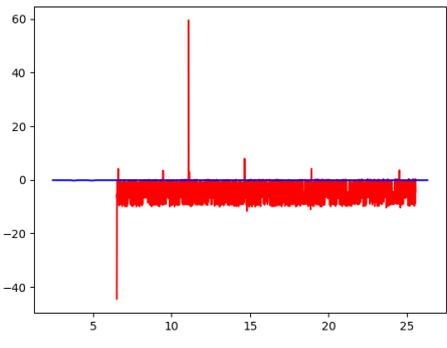
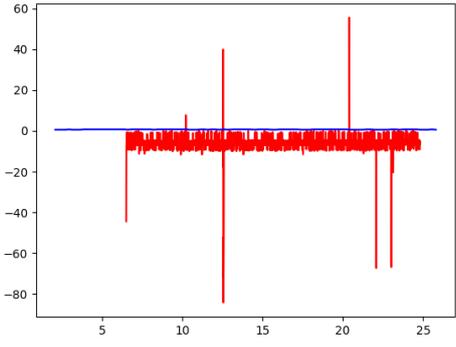
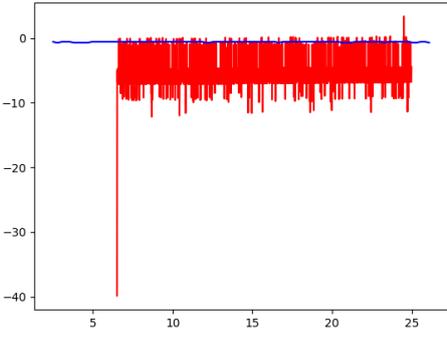
一、研究結果表格

在實測結束後，我們將三種情境的系統輸出資料重合率分為以下兩個表格紀錄，分別比較單台靜止噪音車輛在不同距離和噪音分貝下的系統表現，以及其餘車輛的有無對單台移動噪音車輛情況下的系統表現。

(一)單台靜止噪音車輛在不同距離和噪音分貝下的系統表現

表 1 為單台靜止車輛在不同噪音分貝數和距離下之資料重合率比較，在單台靜止噪音車輛距離 Respeaker 40 公分時，系統對於 60 分貝和 80 分貝的噪音分別具有 79.0% 和 81.5% 資料重合率，由原始輸出角度對時間的關係圖表可觀察到 ODAS 偵測噪音後所輸出角度（紅線）和 YOLO 偵測車輛後所輸出角度（藍線）趨勢相近，由此可知兩組資料重合狀況良好。以此類推下表（表 1）中單台靜止噪音車輛在距離 Respeaker 55 公分和 70 公分時的數據意義。

表中最左欄括弧內的數字為依照一般車輛（380-430cm）和遙控模型車的車輛長度比等比例縮放後的實際距離，例如：當遙控模型車距離 Respeaker 40 cm，粗略計算遙控車（車長 31cm）與真實車輛（車長 405cm）的比例為 1:13.1，依此比例放大距離後，可推估應用到實際情況時，真實車輛與噪音照相系統的距離約為 5.24 公尺。

靜止	60 分貝	80 分貝
距離 40cm (5.2m)	79.0% 	81.5% 
距離 55cm (7.2m)	67.5% 	84.0% 

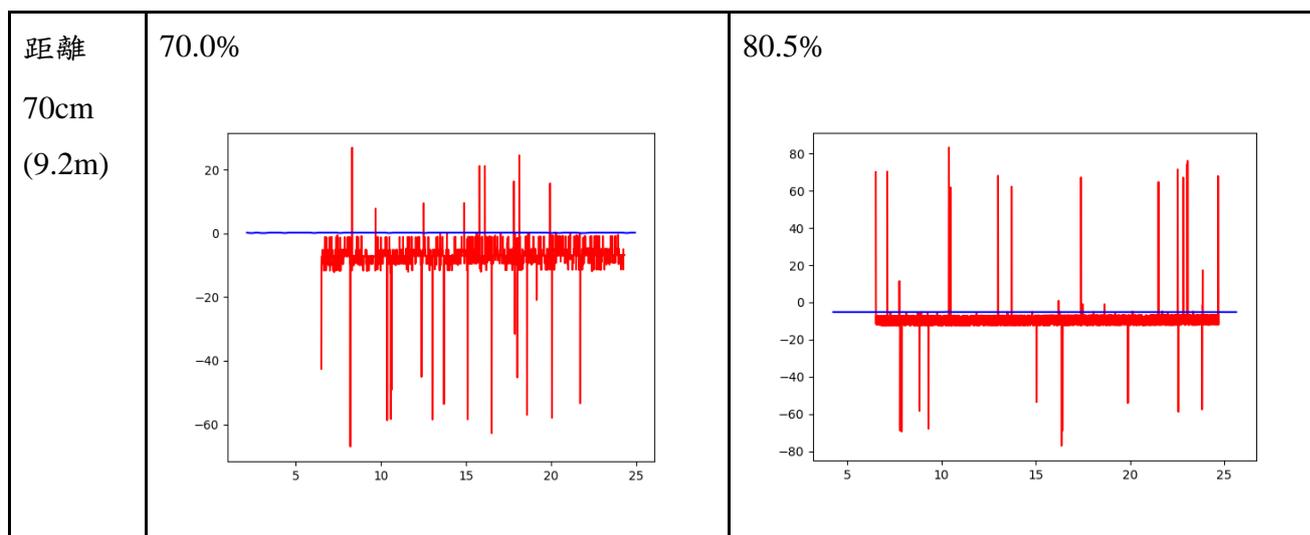


表 1 單台靜止車輛在不同噪音分貝數和距離下之資料重合率比較

(二) 其餘車輛的有無對單台移動噪音車輛情況下的系統表現

表 2 為單台移動車輛在不同情境(是否有其他車輛)和距離下之資料重合率比較，在移動噪音車輛發出 80 分貝的噪音且距離 Respeaker 40 公分時，系統對於是否有其餘車輛干擾的資料重合率皆為 16.0%，由原始輸出角度對時間的關係圖表可觀察到 ODAS 偵測噪音後所輸出角度(紅線)和 YOLO 偵測車輛後所輸出角度(藍線)在中段(車輛移動時)趨勢相近，但在前段和後段時，可觀察到 YOLO 輸出資料(藍線)消失，代表 YOLO 沒有在畫面中辨識到該車輛。以此類推下表(表 2)中單台移動噪音車輛在距離 Respeaker 55 公分和 70 公分時的數據意義。

表中最左欄括弧內的數字為依照一般車輛(380-430cm)和遙控模型車的車輛長度比例縮放後的實際距離，例如:當移動的遙控模型車距離 Respeaker 40 cm，粗略計算遙控車(車長 38cm)與真實車輛(車長 405cm)的比例為 1:10.6，依此比例放大距離後，可推估應用到實際情況時，真實車輛與噪音照相系統的距離約為 4.2 公尺。以此類推距離 Respeaker 55 公分及 70 公分時對應的實際距離。(本實驗中的移動車輛和靜止車輛為相異兩台車，靜止車輛的車長為 31cm，移動車輛的車長為 38 公分，因此兩表格中相同距離等比例縮放後的實際距離亦相異。)

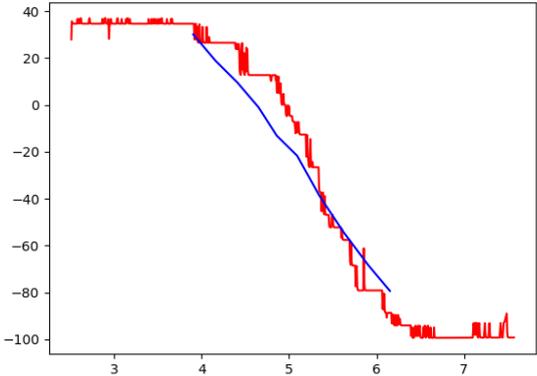
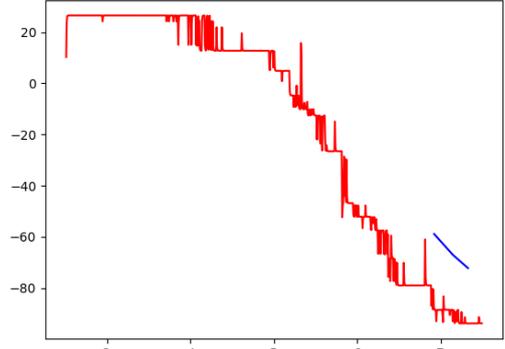
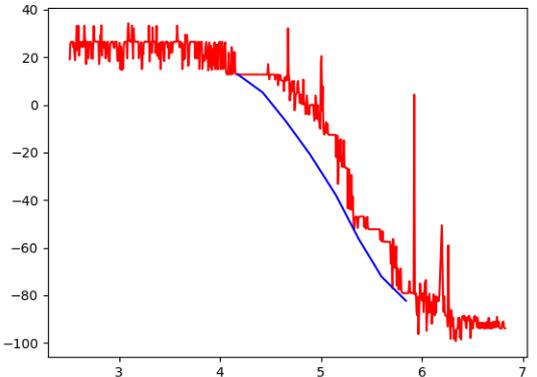
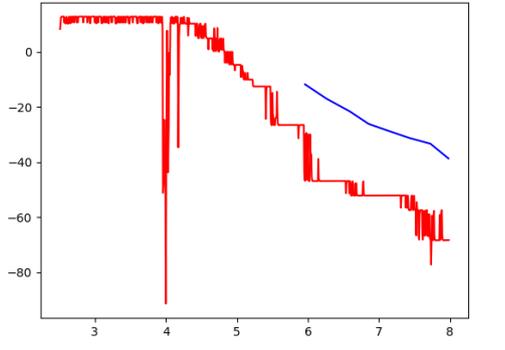
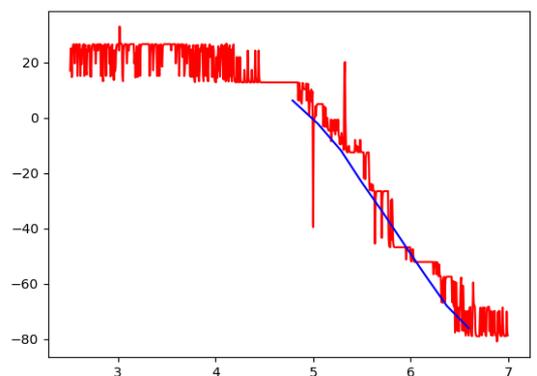
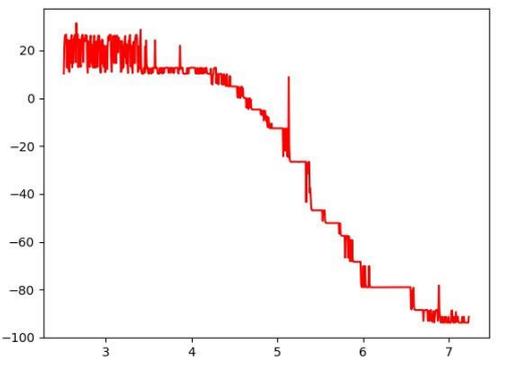
移動	一台移動	一台移動、一台靜止無噪音
40cm (4.2m)	16.0% 	16.0% 
55cm (5.8m)	26.0% 	14.0% 
70cm (8.2m)	30.0% 	4.0% 

表 2 單台移動車輛在不同情境(是否有其他車輛)和距離下之資料重合率比較

二、討論

(一)單台靜止噪音車輛在不同噪音分貝數和距離下對系統輸出資料重合率的影響

由表 1 我們可以觀察到當靜止噪音車輛距離麥克風陣列 (Respeaker) 40 公分時，系統對於 80 分貝和 60 分貝噪音源的資料重合率差距不顯著，因此可推測在距離較近時，噪音的分貝數對系統性能影響較小，但隨著距離拉遠至 55、70 公分，系統對於 80 分貝的噪音源的資料重合率相較 60 分貝的噪音源明顯較高，由此可推知在距離稍遠時，噪音的分貝數對系統性能影響較大。

(二)單台噪音車輛在移動與靜止狀態下對系統輸出資料重合率的影響

比較表 1 和表 2 可以發現在同一距離、噪音分貝下，系統對於移動噪音車輛的表現遠不如靜止噪音車輛，若比較表格中的原始實驗數據圖表 (ODAS (紅線) 及 YOLO (藍線) 輸出的角度對時間的圖表) 卻可以觀察到聲源在中段的移動過程 (大約 4~6 秒之間) 兩線趨勢相近、重合率良好，代表移動中的噪音有被穩定的追蹤，但 YOLO 角度資料的輸出 (藍線) 卻在第 4 秒以前和最後 1 秒消失，也就是車輛啟動前和停止後位移為 0 的時間點，我們推測有以下原因：

1. 遙控車不夠擬真和置於其上的噪音播放器體積過大導致 YOLO 辨識車輛困難。
2. 遙控車的起始和終止位置距離攝影系統較遠導致 YOLO 辨識物件困難。
3. 遙控車在靠近起始點和終止位置時，面對攝影系統的角度偏向正面/正背面導致 YOLO 辨識車輛困難。

(三)其餘車輛的有無對系統輸出資料重合率的影響

由表 2 可以觀察到移動噪音車輛在距離 Respeaker 較近時 (40 公分) 其餘車輛的有無對系統輸出資料重合率看不出影響，可推測若對應到真實車輛，此系統可在約 4.2 公尺左右時，從兩台車輛中辨識出發出噪音的那一台，且系統的表現能與辨識單台移動噪音車輛時相差無幾。但在較遠的距離時，YOLO 物件追蹤的表現較為不穩定。

(四)等比例縮放

若不考慮其他因素，由括弧內換算數值可推知此系統大約可在距離靜止真實車輛 5.2 至 9.2 公尺處良好的偵測到其噪音。

肆、結論與應用

一、結論

本研究結合了 YOLO 的物件偵測及追蹤與 ODAS 音源方位分析，成功實作出用以追蹤噪音車輛的自動化系統。我們將 ReSpeaker 麥克風陣列安裝在微型電腦 Raspberry Pi 上，透過 SSH 連線在筆電端使用 Python 控制 ODAS 軟體分析聲源位置，並同時執行 360 度攝影機回傳影像上的 YOLO 的物件偵測及追蹤，實現了兼具聲音上與影像上的噪音追蹤系統。

二、應用

除了應用在執法上外，本研究中的系統還能應用在任何需要整合追蹤影像和聲源的使用情境，例如錄製話劇表演及生態觀測。

三、未來展望

未來我們下一階段的目標包括：

1. 加入分貝計，以檢測發出的噪音是否超過法規標準。
2. 優化使用者介面，包裝成方便操作的應用程式。
3. 紀錄分貝值超標時的影音資訊以提供有力證據

伍、參考文獻

[1] “[道路管理處罰條例 第 43 條](http://law.moj.gov.tw/LawClass/LawSingle.aspx?pcode=K0040012&flno=43)。” 全國法規資料庫, 3 May 2023, law.moj.gov.tw/LawClass/LawSingle.aspx?pcode=K0040012&flno=43. Accessed 31 Oct. 2023.

- [2] “行駛中機動車輛噪音管制標準及裁罰金額為何?” 臺北市政府環境保護局, 18 Oct. 2023,
www.dep.gov.taipei/News_Content.aspx?n=ACEFA960B5A4ACD7&s=52D375B6FAD9C CA1. Accessed 31 Oct. 2023.
- [3] “何謂聲音照相系統?” 臺北市政府環境保護局, 18 Oct. 2023,
www.dep.gov.taipei/News_Content.aspx?n=ACEFA960B5A4ACD7&sms=87415A8B9CE8 1B16&s=D7279E294B48077C. Accessed 30 Oct. 2023.
- [4] 行政院環境保護署. 聲音照相科技執法實務運作指引. 30 June 2023, p. 26.
- [5] DiBiase, J. H. (2000). A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays. Brown University.
- [6] Ylonen, T., & Lonvick, C. (2006). The secure shell (SSH) protocol architecture (No. rfc4251).
- [7] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
- [8] Jocher, Glenn. “YOLOv8 Documentation.” *Docs.ultralytics.com*, 18 May 2020, docs.ultralytics.com/.
- [9] Seed Studio. (2023, January 11). Respeaker 4-MIC array for Raspberry Pi: Seed studio wiki. Seed Studio Wiki RSS.
https://wiki.seedstudio.com/ReSpeaker_4_Mic_Array_for_Raspberry_Pi/
- [10] Grondin, F., Létourneau, D., Godin, C., ... & Michaud, F. (2019). ODAS: Open embedded audition system. *Frontiers in Robotics and AI*, 125.
- [11] Labbe, R. (2014). Kalman and bayesian filters in python.

【評語】 190014

噪音之眼有前瞻性及實用性，值得持續努力發展。本成果美中不足的是：實驗場景使用小道具，可能無法與實際情況收集的資料有密切關聯性。請一併考量雲端解決方案的可行性。希望在不久的將來能夠看到本系統的完成並且實際使用於道路現場。