

## 第二屆系統黑客松參賽計畫書

題目：透過聲音來辨識電梯和馬達行為

報名編號：#16

組別：ARMmbed組

使用平台：NUCLEO-L476RG

隊伍名稱：KY-038

隊員：傅意茹、林怡亨

一、	摘要 .....	2
二、	作品設計動機及應用對象 .....	2
三、	作品結構及原理說明 .....	2
1	使用技術說明 .....	2
1.1	Fourier transform (傅立葉轉換) .....	2
1.2	MFCC (梅爾頻率倒譜系數) .....	2
2	作品結構 .....	4
2.1	系統開發環境 .....	4
2.1.1	硬體設備 .....	4
2.1.1.1	NUCLEO-L476RG .....	4
2.1.1.2	KY-038 .....	4
2.1.2	編譯環境 .....	4
2.1.2.1	ARMmbed 線上編譯器 .....	4
2.2	系統流程 .....	4
2.2.1	電風扇檢測 .....	4
2.2.2	電梯檢測 .....	5
四、	預期成果 .....	5
1	電風扇檢測 .....	5
2	電梯檢測 .....	5
五、	參考文獻 .....	5

## 一、摘要

為使機械檢測自動化以達到降低機械檢測成本的目的，本企劃預期使用 ARMmbed 平台搭配音訊接受設備，以收取電風扇以及電梯運轉的聲音，並使用傅立葉轉換及梅爾頻率倒譜等音訊辨識的技術加以分析，以判斷機械運轉狀況，以自動化的方式察覺機械異常行為。

## 二、作品設計動機及應用對象

因人力檢測機械所需成本高，且一般人缺乏相關知識，只能單方面聽取檢測人員說法，造成資訊不對稱；又現今人力檢測機械時，多以機械運轉發出之聲音為判斷基準之一，例如引擎運作時是否發出異音。故我們希望能夠透過收取並分析機械運轉聲音，以判斷機械行為，進而達到檢測機械是否正常運作的目的。

## 三、作品結構及原理說明

### 1 使用技術說明

#### 1.1 Fourier transform ( 傅立葉轉換 )

傅立葉轉換能夠將我們取得的原始訊號，自時域轉換到頻域。

如果我們的訊號可以表示成  $x[n]$ ,  $n = 0 \sim N-1$ ，那麼離散傅立葉轉換( DFT) 的公式如下：

$$X[k] = (1/N) \sum_{n=0}^{N-1} x[n] \cdot \exp(-j \cdot 2\pi \cdot n \cdot k / N), k=0, \dots, N-1$$

這些傅立葉係數  $X[k]$  所代表的資訊是  $k$  的函數，而  $k$  直接和頻率有正比關係，因此這些係數  $X[k]$  通稱為「頻譜」(Spectrum)。<sup>[2]</sup>

快速傅立葉轉換(FFT)會通過把DFT矩陣分解為稀疏因子之積來快速計算此類轉換。因此，它能夠將計算DFT的複雜度從只用DFT定義計算需要的 $O(n^2)$ ，降低到 $O(n \log n)$ ，其中 $n$ 為資料大小。<sup>[5]</sup>考量到平台的處理能力，我們採用複雜度較小的FFT將原始資料轉換至頻域，再進行聲音特徵的分析。

#### 1.2 MFCC ( 梅爾頻率倒譜系數 )

MFCC，全名Mel-scale Frequency Cepstral Coefficients。<sup>[3]</sup>為語音辨識 ( Speech Recognition ) 及語者辨識 ( Speaker Recognition ) 方面最常用到的語音特徵。此參數考慮到人耳對不同頻率的感受程度，因此特別適合用在語音辨識。

[5] [6] MFCC通常有以下之過程:

1 訊框化 ( Frame blocking ) :

1.1 為簡化連續變化的訊號，我們假設在一個短時間尺度內，音頻訊號不發生改變，將訊號以多個取樣點集成一個單位，稱為「訊框」。

1.2 取樣點數量N多取256或512。

2 漢明窗 ( Hamming window ) :

2.1 將每一個音框乘上漢明窗，以增加音框左端和右端的連續性。

2.2 假設訊框化訊號 $S(n)$ ,  $n = 0, \dots, N-1$ ，乘上漢明窗後為  $S'(n) =$

$S(n) * W(n)$ ，其中 $W(n)$ 為：

$$W(n) = 0.54 - 0.46 \cos(2n\pi/(N-1)), 0 \leq n \leq N-1$$
$$W(n) = 0, \text{ otherwise}$$

3 預強調 ( Pre-emphasis ) :

3.1 消除發聲過程中聲帶和嘴唇的效應，來補償語音信號受到發音系統所壓抑的高頻部分。

3.2 突顯在高頻的共振峰。

4 傅立葉轉換 ( Fourier transform )

5 梅爾濾波器(三角重疊窗口) :

5.1 將訊框頻譜乘以一組三角帶通濾波器，取得梅爾刻度 ( Mel scale )，並從每一個梅爾刻度提取對數能量 ( Log Energy )。一組數量通常取為20，這 20 個三角帶通濾波器在「梅爾頻率」 ( Mel Frequency ) 上是平均分佈的，而梅爾頻率和一般頻率  $f$  的關係式如下：

$$\text{mel}(f) = 1125 * \ln(1 + f/700)$$

5.2 梅爾頻率代表一般人耳對於頻率的感受度，由此也可以看出人耳對於頻率  $f$  的感受是呈對數變化的，相較高頻部分，人耳在低頻部分的感受較為敏銳。

5.3 三角帶通濾波器有兩個主要目的：

5.3.1 對頻譜進行平滑化，並消除諧波的作用，突顯原先語音的共振峰。因此以 MFCC 為特徵的語音辨識系統，不會受到輸入語音的音調不同而有所影響。

5.3.2 降低資料量。

6 對上面獲得的結果進行離散傅立葉反轉換，轉換到倒頻譜域。

7 MFCC即這個倒頻譜圖的幅度(amplitudes)。一般使用12個係數，與訊框能量疊加得13維的係數。

## 2 作品結構

### 2.1 系統開發環境

#### 2.1.1 硬體設備

2.1.1.1 NUCLEO-L476RG

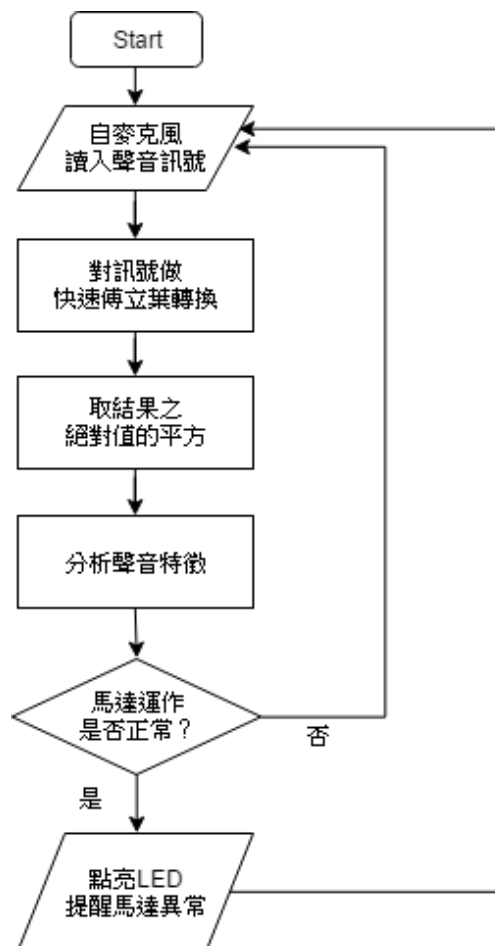
2.1.1.2 KY-038

#### 2.1.2 編譯環境

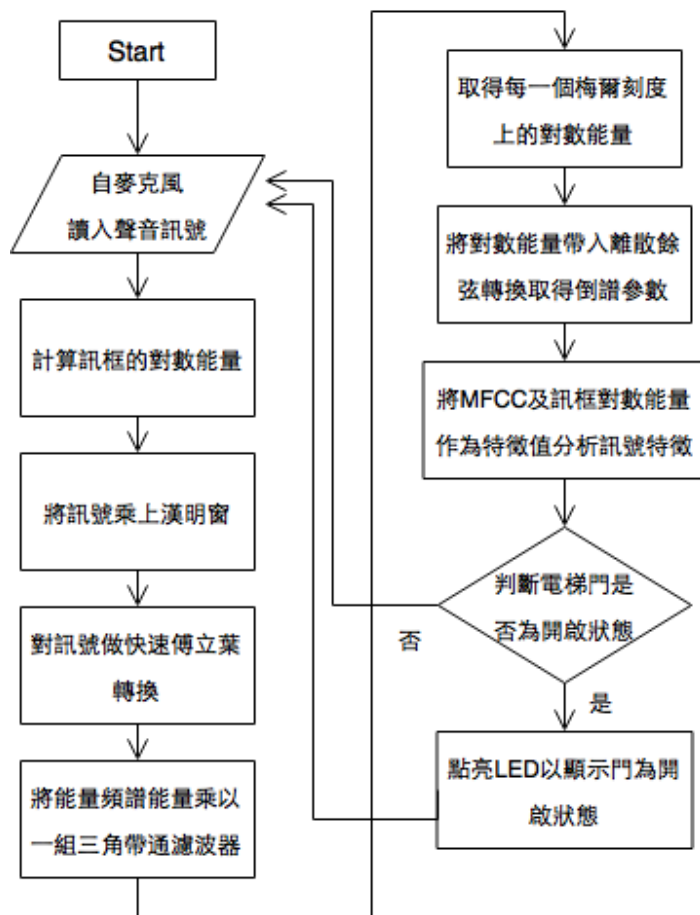
2.1.2.1 ARMmbed線上編譯器

### 2.2 系統流程

#### 2.2.1 電風扇檢測



### 2.2.2 電梯檢測



## 四、 預期成果

### 1 電風扇檢測

透過收取電風扇運轉的聲音，以判斷是否在運轉過程中產生異音，以及目前的風扇轉速。

### 2 電梯檢測

透過收取電梯內部的聲音，判斷目前電梯開關門的狀況。

## 五、 參考文獻

[1] 陳韋戎。聲音特徵的取得及辨識。台灣大學農業機械工程學系四年級，未出版，台北。

[2] 未具名。演算法筆記-Audio。2017/2/7，取自  
<http://www.csie.ntnu.edu.tw/~u91029/Audio.html>。

- [3] 張智星。音訊處理與辨識。2017/2/7，取自  
<http://mirlab.org/jang/books/audioSignalProcessing/>。
- [4] ARM Ltd.。ARMmbed。2017/2/7，取自  
<https://www.mbed.com/zh-cn/>。
- [5] 維基百科。快速傅立葉轉換。2017/2/7，取自  
<https://www.mbed.com/zh-cn/>
- [6] 維基百科。梅爾頻率倒譜係數。2017/2/14，取自  
<https://zh.wikipedia.org/wiki/梅尔频率倒谱系数>