

# 應用於車載通訊之救護車鳴笛聲辨識方法

王文聖  
朝陽科技大學  
資訊與通訊系  
s9830606@cyut.edu.tw

廖俊鑑  
朝陽科技大學  
資訊與通訊系  
jjliaw@cyut.edu.tw

彭聖竺  
朝陽科技大學  
資訊與通訊系  
s10030606@cyut.edu.tw

## 摘要

救護車鳴笛聲容易受噪音破壞及聲音阻隔，致使駕駛無法聽到而讓道。本研究提出一個救護車鳴笛聲辨識方法，可用於車載通訊中。當偵測到救護車聲音時，提醒駕駛者。所使用的方法是根據救護車鳴笛聲頻率特性，分別利用頻率資訊及最長共同子序列方法，判別聲音訊號是否為救護車鳴笛聲。實驗結果顯示，系統於不同噪音環境下，在有救護車鳴笛聲下的辨識正確率為 92%，無救護車鳴笛聲時的辨識正確率為 98.5%。

**關鍵詞：**救護車鳴笛聲、聲音辨識、最長共同子序列。

## Abstract

The function of an ambulance siren is to remind other drivers on the road to make timely response to give way to the ambulance. The ambulance siren may be disrupted by noise and blocked by voices in the real situation. In this paper, we propose a process to recognize ambulance siren. The process can be used for the Telematics. The frequency characteristic, and Longest Common Subsequence (LCS) algorithm are used to recognized the ambulance siren. According to the experiment results, the accuracy rate is over 92.8%.

**Key words:** ambulance siren sound, sound recognition, Longest Common Subsequence.

## 1. 前言

依據 100 年全國消防機關緊急救護調查資料顯示，94 年救護車出勤次數為 678,989 次、95 年為 692,690 次，直至 100 年 1,003,974 次。近年來，緊急救護服務平均約以 10% 速率成長。救護車駕駛在執行勤務之時，需要就是快速且平安的到達目的地。當駕駛聽到救護車鳴笛聲時，為確保救援時間，第一反應都需禮讓

救護車先行。

現今生活中，駕駛者僅能依靠聽覺或視覺查覺週遭是否有救護車靠近。在許多情況下，救護車聲音會受環境因素影響，使駕駛者無法聽到。干擾因素可列舉有：

1. 上、下班時段或熱鬧區域中，汽車、機車、公車及廣告宣傳車等車輛產生的噪音。
2. 現今的汽車講求內裝氣密、安靜及舒適效果，在窗戶關閉情況下，不易聽到車外的聲音。
3. 駕駛者開車時，因收聽廣播、音樂或接聽電話，當察覺有救護車時，通常已距離自己很近，且當下無法做出讓道的回應。

因此，若能藉由鳴笛聲辨識技術輔助，並藉由車載通訊技術，提醒駕駛者有救護車靠近，或進而將資訊傳送前方車輛，便可使用周遭駕駛有更充分時間做迴避動作，讓救護車能迅速到達救援目的。

過去有學者提出對消防車、拖吊車、救護車、警車等工程車輛的鳴笛訊號分析的方法 [1]，主要是將聲音訊號轉換為頻域資訊，再利用梅爾倒頻譜係數 [2][3]，萃取聲音的特徵參數，並藉由類神經網路 [4][5]，進行訊號樣本的比對，達到識別效果。救護車鳴笛聲的頻率特性相較於其它聲音是明確的。因此，本研究直接使用濾波器濾出救護車鳴笛聲的高、低頻資訊，並透過序列比對的方法，即可對救護車鳴笛聲做辨識。

## 2. 救護車鳴笛聲辨識流程

### 2.1 聲音輸入

實驗中，我們使用麥克風，錄製時間長度 30 秒聲音訊號，取樣頻率為 8000Hz/秒。

### 2.2 音框切割

並利用音框切割，將聲音訊號做短時距處理。本論文每段音框為 0.1 秒。

### 2.3 快速傅立葉轉換

再將各個音框內的訊號，使用快速傅立葉

轉換轉換為頻域資料[10]。從轉換後的結果(如圖一)可以看出，救護車鳴笛訊號的高頻頻帶為 900Hz 至 1000Hz、低頻頻帶為 650Hz 至 750Hz，符合中華民國規範[11]：

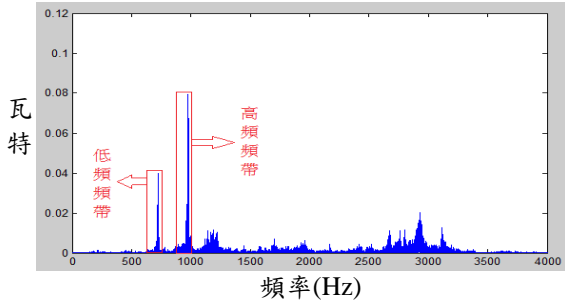


圖 1 救護車聲音訊號頻譜圖

## 2.4 帶通濾波

根據救護車鳴笛聲頻率特性，使用帶通濾波器，濾出救護車聲音訊號高、低頻頻帶。並令高頻頻帶強度最大值為  $fH_{max}$ ；低頻頻帶強度最大值為  $fL_{max}$ ，如圖 2：

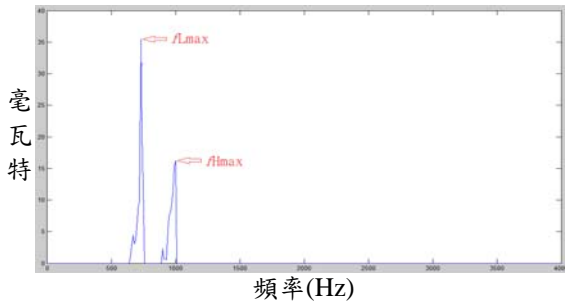


圖 2 濾波後救護車聲音訊號頻譜圖

## 2.5 特徵分析

經由 Eq.(1)，判別各個音框特徵為高頻(H)或低頻(L)。

$$\text{音框特徵} \begin{cases} H, & \text{if } (fH_{max} \geq fL_{max}) \\ L, & \text{else} \end{cases} \quad (1)$$

## 2.6 最長共同子序列運算

本研究使用最長共同子序列方法(Longest Common Subsequence, LCS)[6][7]及動態規劃法(Dynamic Programming, DP)[8][9]，求解聲音序列與救護車聲序列的輸出值，藉此判斷聲音訊號是否為救護車鳴笛訊號。

我們給定序列  $A=[C_1, Y_2, U_3, T_4]$ 、序列  $B=[I_1, C_2, E_3, C_4, Y_5, C_6, T_7]$ ，下標符號為兩序

列的序列長度。並且令  $LCS[i, j]$  為序列 A 與序列 B 各個字元，則 LCS 的輸出值為

$$LCS[i, j] = \begin{cases} LCS[i-1, j-1]+1 & , \text{if } A_i = B_j \\ \max\{LCS[i, j-1], LCS[i-1, j]\} & , \text{if } A_i \neq B_j \end{cases} \quad (2)$$

如表一所舉列，在運算過程中， $A_i$  是由序列 A 第 1 個字元依序與序列 B 第 1 個字元至第最後一個字元運算，再由序列 A 第 2 個字元與序列 B 第 1 個字元至第最後一個字元做運算。依此方式，運算至序列 A 第最後一個字元與序列 B 第 4 個字元，即可求得兩序列的輸出分數。因此，這整個運算過程共需計算上  $7 \times 4$  個字元。所以可直接使用動態規劃法，計算出各比對位址的輸出分數。最後由  $LCS[7, 4]$  位址，可得知兩序列經運算後的輸出值。

表 1 動態規劃法求解兩序列的輸出分數

$B_j$	0	1	2	3	4	5	6	7
$A_i$		I	C	E	C	Y	C	T
0	0	0	0	0	0	0	0	0
1	C	0	0	1	1	1	1	1
2	Y	0	0	1	1	1	2	2
3	U	0	0	1	1	1	2	2
4	T	0	0	1	1	1	2	3

表 2 救護車聲音高、低頻排列組合

1	LLLLHHHHHH	6	HHHHHLLLLLH
2	LLLHHHHHHHL	7	HHHHLLLLLHH
3	LLHHHHHHHLL	8	HHHLLLLLHHH
4	LHHHHHHHLLL	9	HHLLLLLHHHH
5	HHHHHHHLLLL	10	HLLLLLHHHHH

並當音框經過特徵分析判斷後，我們再將音框特徵以 1 秒的時間長度(10 個特徵)劃分為一組聲音序列，且每間隔 0.1 秒，即劃分一組與救護車序列做 LCS 運算。

並依據救護車的規範[11]，我們可清楚得知，救護車鳴笛聲在時間 1 秒間是由連續的 4 個低頻及連續的 6 個高頻所組成，但因開始計算的時間不同，可能有如表 2 的排列可能。在此，我們將取表 2 中第 1 組序列做為與聲音序列做 LCS 運算的救護車序列。

當直接將 LCS 方法，直接使用於救護車

聲音時，在幾種特例情況，會因聲音序列取樣排列的不同，而有判斷錯誤的情況發生，當中也包含了第 2 組至第 10 組救護車序列。因此，在聲音序列做 LCS 運算前，我們將透過兩條件做修正：

1. 計算聲音序列 L 與 H 變換次數，需為 1 或 2 次。
2. 計算聲音序列 L 與 H 最大連續數，需介於 4 至 7 個。

其中，會將 L 與 H 的最大連續數取到 7，是因為救護車聲音高頻與低頻做變換的時候，高頻與低頻的強度是比較接近的，換句話說，實際上會出現連續 3 個低頻、連續 7 個高頻或連續 5 個低頻、連續 5 個高頻等序列組合。我們將上述可能出現的序列組合，整理為上述兩條件。若聲音序列符合兩條件，即判定為救護車序列，輸出值設定為 10。若有一項條件不符，則以原來的聲音序列與救護車序列做 LCS 運算。

### 3. 實驗分析

#### 3.1 實驗環境

我們於實際環境下錄製並模擬兩種救護車聲音可能遭受破壞的噪音情境，並將不同噪音情境做為辨識系統收音時背景雜訊，藉此計算系統於噪音雜訊干擾下，量測到辨識準確率。實驗環境設置及實驗聲音大小(分貝)，分別如表 3 及表 4：

表 3：實驗環境設置

實驗環境	取樣日期 取樣時間	取樣地點	環境描述
i	2012-06-18 P.M.07:34	台中市大里區東湖路	下雨的雨滴聲
ii	2012-06-17 P.M.08:54	台中市大里區國光路	倒車的警示聲

表 4：實驗環境分貝表

實驗環境	有救護車鳴笛聲		無救護車鳴笛聲	
	分貝 最大值	分貝 最小值	分貝 最大值	分貝 最小值
i	100(dB)	87(dB)	100(dB)	86(dB)
ii	100(dB)	86(dB)	85(dB)	70(dB)

#### 3.2 實驗結果

我們以混淆矩陣(Confusion Matrix)分析法例出實驗結果(如表 5)。其中正確率計算方式，分別列於 Eq.(3)至 Eq.(6)。辨識結果分別如圖 3 至圖 6 所示；辨識正確率整理列於表 6：

表 5：混淆矩陣分析表

	分類為有救護車鳴笛聲	分類為無救護車鳴笛聲
有救護車鳴笛聲	正確率(TPR) True Positive Rate	錯誤率(FNR) False Negative Rate
無救護車鳴笛聲	錯誤率(FPR) False Positive Rate	正確率(TNR) True Negative Rate

- 有救護車鳴笛聲

C = 有救護車鳴笛聲區間輸出個數

D = 判斷錯誤區間輸出個數

E = 有救護車鳴笛聲區間輸出分數

$$TPR = \frac{E = 10 \text{ 的總數量}}{C + D \text{ 的總數量}} \times 100\% \quad (3)$$

$$FNR = \frac{E \leq 9 \text{ 的總數量}}{C + D \text{ 的總數量}} \times 100\% \quad (4)$$

- 無救護車鳴笛聲

F = 無救護車鳴笛聲區間輸出個數

G = 無救護車鳴笛聲區間輸出分數

$$TNR = \frac{G \leq 9 \text{ 的總數量}}{F \text{ 的總數量}} \times 100\% \quad (5)$$

$$FPR = \frac{G = 10 \text{ 的總數量}}{F \text{ 的總數量}} \times 100\% \quad (6)$$

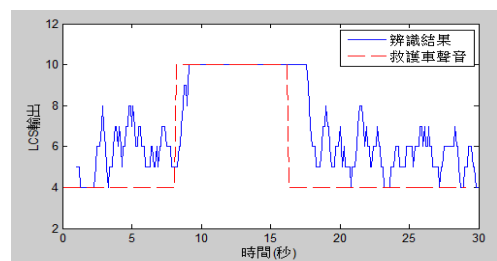


圖 3：有救護車鳴笛聲於環境 i 輸出辨識圖

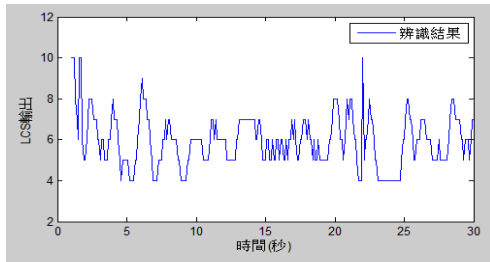


圖 4：無救護車鳴笛聲於環境 i 輸出辨識圖

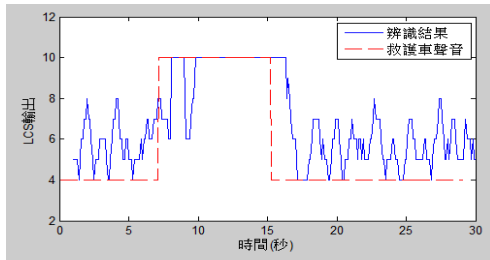


圖 5：有救護車鳴笛聲於環境 ii 輸出辨識圖

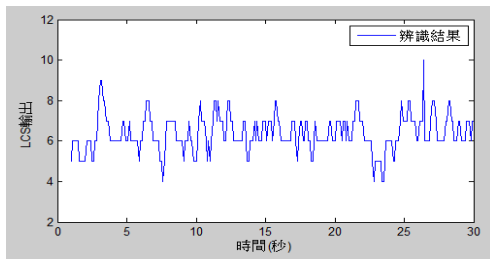


圖 6：無救護車鳴笛聲於環境 ii 輸出辨識圖

表 6 辨識正確率

實驗環境	有救護車鳴笛聲		無救護車鳴笛聲	
	TPR(%)	FNR(%)	TNR(%)	FPR(%)
i	94	6	98	2
ii	90	10	99	1
平均	92	8	98.5	1.5

#### 4. 結論

本論文提出應用最長共同子序列方法辨識救護車鳴笛聲。當聲音輸入後，切割為每段 0.1 秒聲音訊號，接著再將各音框資料經特徵分析，分類為高頻或低頻特徵。音框特徵以 10 個為一組的方式與救護車序列做最長共同子序列運算，藉由其輸出值以判別聲音訊號是否為救護車鳴笛聲。由實驗結果顯示，當有救護車鳴笛聲時，辨識正確率為 92%，無救護車鳴笛聲存在時，辨識正確率為 98.5%。

由最長共同子序列運算後，我們可求得各個聲音序列輸出值，未來可加入模糊(Fuzzy)[12]判斷方法，利用 Fuzz 的優點，進一步判斷聲音訊號為救護車鳴笛聲、可能為救護車鳴笛聲、非救護車鳴笛聲三種結果，以提高系統的辨識正確率及實用性。

除此之外，系統在辨識上需花費 1 秒響應時間，才能得到辨識結果，因此，若能縮減辨識響應時間，且模擬更多救護車實際出現情況，例如：多種噪音混雜的情境、考量救護車移動時，產生的都普勒效應等因素，相信可替駕駛帶來許多行車上的便利性。

#### 參考文獻

- [1] F. Beritelli, S. Casale, A. Russo, and S. Serrano, "An Automatic Emergency Signal Recognition System for The Hearing Impaired," in *Proc. of 12th Digital Signal Processing Workshop and 4th Signal Processing Education Workshop*, pp. 179-182, 2006.
- [2] B. Milner, and X. Shao, "Clean speech reconstruction from MFCC vectors and fundamental frequency using an integrated front-end," *Speech Communication*, Volume. 48, Issue 6, pp. 697-715, 2006.
- [3] S. Imai, "Cepstral analysis synthesis on the mel frequency scale," *Proc. of ICASSP*, volume. 8, pp. 93-96, 1983.
- [4] F. F. Li, and T. J. Cox, "A neural networks model for speech intelligibility quantification," *Applied Soft Computing*, Volume. 7, Issue 1, pp. 145-155, 2007.
- [5] N. V. Viadimir, "An Overview of Statistical Learning Theory," *Journal of IEEE Transcation on Neural Networks*, Volume. 10, No. 5, pp. 988-999, 1999.
- [6] D. S. Hirschberg, "Algorithms for the longest common subsequence problem," *J. ACM*, Volume. 24, No. 4, pp. 664-675, 1977.
- [7] R. W. Irving and C. B. Fraser, "Two algorithms for the longest common subsequence of three (or more) strings," *Lecture Notes in Computer Science*, Springer Verlag, pp. 214-229, 1992.
- [8] C. Li, "An approach for solving the constrained longest common subsequence problem," Master Thesis, Department of Computer Science and Engineering, National Sun Yat-Sen University, 2003.
- [9] Y. L. Santis, A. De, A. Lisa, N.L. Kim, and S. K, "A simple algorithm for the constrained sequence problems," *Information Processing Letters*, Volume. 90, pp. 175-179, 2004.
- [10] 王宣富, 陳維, 鄔文杰 (2008)。數位訊號處理。台北：高立圖書有限公司。
- [11] 內政部所主管災害緊急應變警報訊號之種類、內容、樣式、方法及其發布時機公告，

內政部，台內消字第 1010821910 號，中華  
民國 101 年 4 月 23 日。

[12]L. A. Zadeh, “Fuzzy sets”, *Information  
Control*, Volume. 8, pp. 338-353, 1965.