

以頻率特徵辨識救護車鳴笛聲

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

連思斌
朝陽科技大學
資訊與通訊系
s9730602@cyut.edu.tw

摘要

聲音辨識是利用聲音訊號的特徵，獲得出聲音資訊中所含的意義。多數聲音的研究大多在於人類的語音或音樂，對於攸關民眾生命安全的救護車鳴笛聲辨識並不多。救護車鳴笛聲的功能在於提醒其他用路駕駛即時做出讓道的回應，但實際情況中，救護車鳴笛聲會因噪音破壞及聲音阻隔使駕駛無法聽到救護車聲音。本研究提出一個救護車鳴笛聲辨識方法，可用於偵測到救護車鳴笛聲時，提醒駕駛者使駕駛者有更充分的時間提早因應及準備。

救護車鳴笛聲是由高頻及低頻規律交互組成，本研究即利用此頻率特徵，用支持向量機來辨識救護車鳴笛聲。首先在訓練過程中是藉由單純救護車鳴笛聲進行訓練，取得分類模型。實際運作時，當實際收到聲音訊號後，即利用分類模型判斷，若聲音中含有救護車鳴笛聲的高頻或低頻，即判斷有救護車鳴笛聲存在。最後的實驗結果中顯示，在救護車鳴笛聲距離 25 公尺內，辨識準確率達 100%。

關鍵詞：救護車鳴笛聲、聲音辨識、支持向量機、高頻、低頻。

Abstract

In this paper, we use the Support Vector Machine (SVM) to identify the ambulance siren. The method is divided into the training processing and the identification processing. In training processing, the pure sound of ambulance siren is used to obtain the classification model. In identification processing, the ambulance siren signal can be recognized if the high frequency part or the low frequency part exists in the voice. In the experiments, we show the performances of the proposed scheme.

Keywords: Support Vector Machine, ambulance siren, voice recognition, high frequency, low frequency.

1. 簡介

在日常生活中，當駕駛者開車聽到救護車鳴笛聲時，為確保就援時間，第一反應都要禮讓救護車先行，但很多情況下會因其它因素影響，使駕駛者沒聽到救護車聲音，無法即時讓出車道而阻礙到救護車所需行徑的路線。這些因素可列舉有：一、上下班時段或是較熱鬧區域中汽、機車、公車及廣告宣傳車等車輛所發出的噪音。二、現今的汽車講求內裝氣密、安靜及舒適效果使車內不易聽到車外的聲音。三、開車時收聽音樂、廣播或接聽電話而沒聽到救護車聲音。

救護車鳴笛聲的功能在於提醒其他用路駕駛即時做出讓道的動作，但實際情況中，救護車鳴笛聲會因噪音破壞及聲音阻隔使駕駛無法聽到救護車聲音。也就是駕駛員是在沒有即時聽到救護車鳴笛聲的情況下未讓路給救護車先行。因此，若能夠讓駕駛者藉由救護車鳴笛聲辨識系統，於偵測到救護車鳴笛聲時，提醒車內的駕駛者「有救護車靠近」，如此一來，便可使駕駛者有更充分的時間提早因應及準備，以讓救護車能夠更迅速的到達目的地。

過去文獻對於救護車鳴笛聲的研究不多，由 Beritelli 等人於 2006 年提出了針對警車、拖吊車、消防車、救護車等工程車輛發出的緊急鳴笛聲辨識系統[1]。此研究中，作者利用頻域資訊，再由梅爾倒頻譜系數 (Mel Frequency Cepstral Coefficients, MFCCs)[2] 萃取出每段音框的特徵參數，並利用類神經網路 [3][4][5] 進行比對及辨識。救護車聲音的頻率相較於其它聲音的頻率，應是比較明確，因此將救護車的聲音以梅爾倒頻譜系數及類神經網路做分類及辨識，顯得執行了過多的運算，而影響處理時效。所以在這部份我們認為應該直接以濾波器濾出救護車鳴笛聲的頻率，再輔以分類器即可對救護車鳴笛聲做辨識。因此，在本論文中我們使用支持向量機 (Support Vector Machine) 分類方法來進行救護車鳴笛聲之辨識。支持向量機是由貝爾實驗室研究人員 Vapnik 與 AT&T 團隊於 1995 年所創立[6]。該方

法主要是將已建構好的資料類別做為分類的依據，並透過鬆弛變數(Slack Variable)與核心函數(Kernel Function)的方法，來解決線性不可分割的問題。因此在當有新的資料要做分類時，支持向量機即可明確的預測該資料點應當被分到哪一類別中，且支持向量機經計算後還能夠得到一個灰色地帶，將錯誤的資料全部集中於灰色地帶中，以便程式做捨棄的動作來達到最佳的辨識效果。因支持向量機在處理分類[7][8][9][10]與預測方面的問題上皆比其它傳統分類器的效果來的好[11][12][13]，因此在本研究中使用支持向量機方法來辨識出聲音訊號是否為救護車的高、低頻頻率，並藉由支持向量機所產生的灰色地帶，將其它非救護車高、低頻頻率的資料歸類於此區域中，做捨棄的動作，來得到救護車鳴笛聲辨識結果。

2. 方法

在本文中我們利用支持向量機的分類方法來辨識救護車的高、低頻訊號。方法的辨識過程分為訓練流程與辨識流程。

訓練流程圖如圖 1 所示。首先，我們對單純救護車聲音訊號進行收音，並將收到的聲音訊號利用切割音框的方法，將訊號切割為每段為 0.1 秒的音框，並對每段音框做傅利葉轉換[14]將時域資料轉換為頻域資料，並使用帶通濾波器濾出各音框中救護車鳴笛訊號所分佈的高頻頻率 900Hz 至 1000Hz 與低頻頻率 650Hz 至 750Hz 範圍，並以此高、低頻頻率做為救護車聲音訊號的特徵，再計算出救護車鳴笛聲分類模型。

辨識流程如圖 2 所示，當電腦收到外部聲音訊號時，也是將訊號透過切割音框、傅利葉轉換及帶通濾波步驟濾出訊號的高頻頻率 900Hz 至 1000Hz 與低頻頻率 650Hz 至 750Hz 範圍，並以此高、低頻頻率做為聲音訊號的特徵，再將此特徵藉由救護車鳴笛聲分類模型的計算，來判別出該訊號所分佈的頻率是否為救護車的高、低頻頻率。

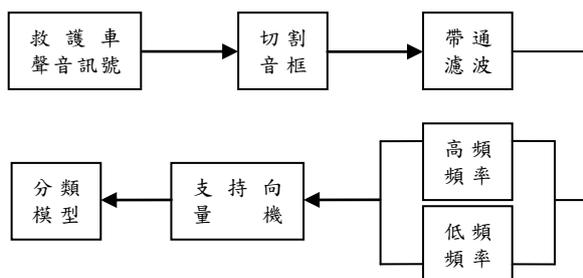


圖 1：支持向量機訓練流程

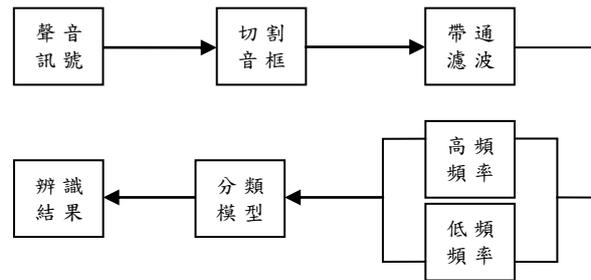


圖 2：支持向量機辨識流程

2.1 收音

圖 3 為支持向量機在訓練流程中電腦收到的救護車聲音訊號，訊號長度為十秒。

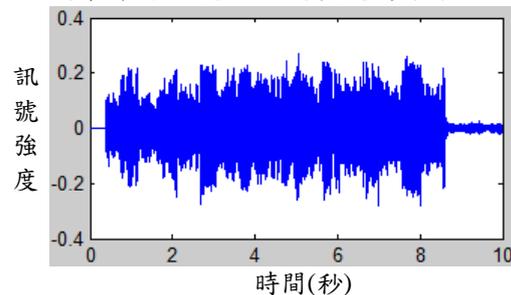


圖 3：救護車聲音訊號

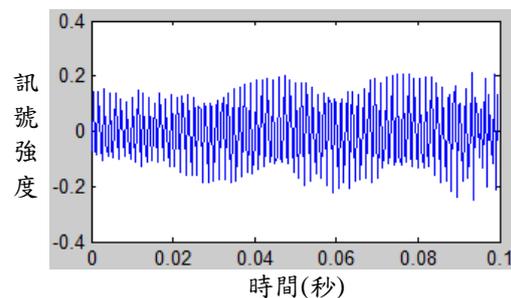


圖 4：切割後之救護車聲音訊號

2.2 切割音框

將收到的聲音訊號以 0.1 秒的訊號長度切割為一段一段的音框，如圖 4 所示。

2.3 傅利葉轉換

將切割後各音框資料利用傅利葉轉換，將時域資料轉換為頻域資料，如圖 5 所示：

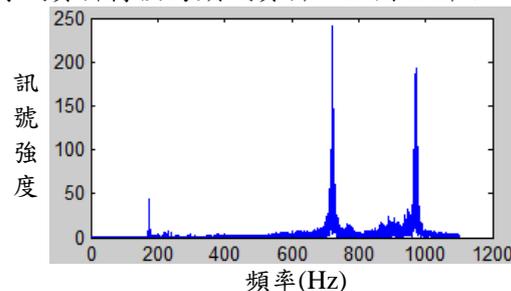


圖 5：救護車聲音訊號頻譜圖

2.4 帶通濾波器

因救護車鳴笛時，為了讓其他駕駛者注意到有救護車要經過，所以在鳴笛聲的設計上高頻頻率 900Hz 至 1000Hz 與低頻頻率 650Hz 至 750Hz 兩頻帶中訊號的強度會特別明顯。因此我們使用帶通濾波器來濾出救護車鳴笛訊號中各音框的高頻頻率與低頻頻率範圍，並以此高、低頻頻率做為救護車聲音訊號的特徵。

圖 6 左邊欄為救護車聲音訊號一秒內低頻頻率的訊號強度變化(L)，右邊欄為救護車聲音訊號一秒內高頻頻率的訊號強度變化(H)。

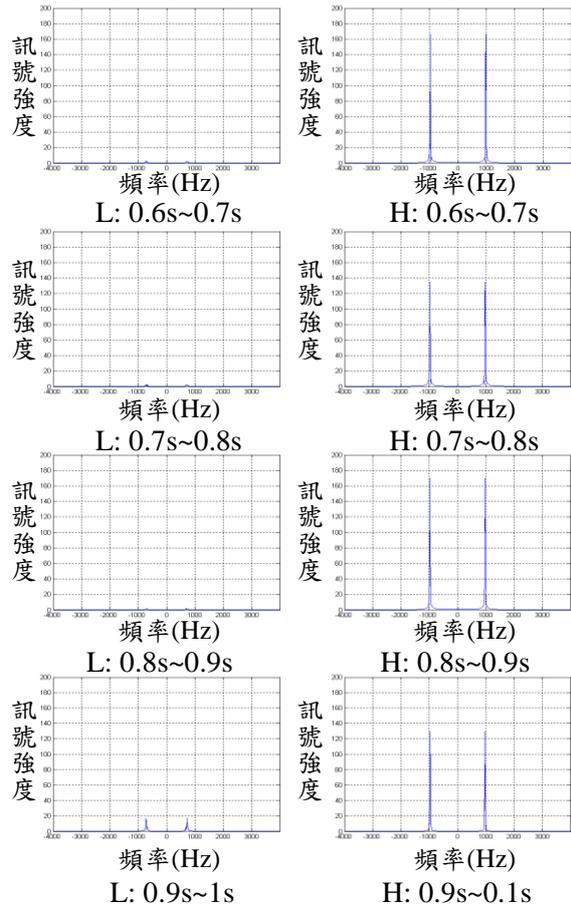
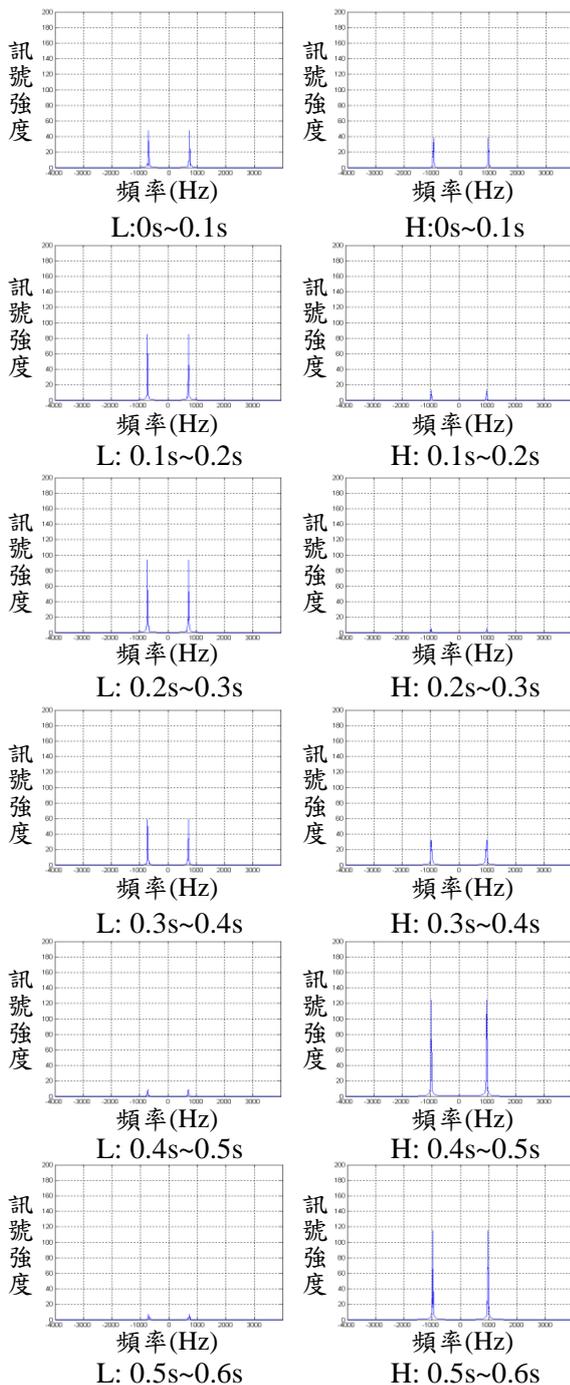


圖 6：救護車訊號低頻及高頻頻譜變化對照圖

在這些音框的頻譜圖中，可完整的看到一次救護車鳴笛訊號的變化，其訊號的強度依序從低頻開始，持續時間為 0.4 秒，接著轉換為高頻，持續時間為 0.6 秒，高、低頻兩者交替進行。接著我們以此為救護車鳴笛聲的特徵，將音框內的資料輸入至以下的支持向量機中做計算。

2.5 支持向量機

支持向量機主要概念是在一空間中建構一個最佳的分類超平面來區隔兩類或多類別的資料。以一個兩類別資料為例(如圖 7 所示)，在這特徵空間中的資料點以方程式(1)表示，其中 \mathbf{x}_i 表示為該資料點的資料向量， y_i 為該資料點的資料類別。我們以 ± 1 兩類別來表示救護車的高頻頻率與低頻頻率。

$$\begin{cases} \mathbf{x}_i, y_i, & i = 1, \dots, n \\ y_i \in \{+1, -1\} \end{cases} \quad (1)$$

接著透過訓練資料集，求出能將兩類別區隔開

的分類超平面。則此分類超平面方程式可表示成：

$$f(x) = w^T \mathbf{x}_i + b \quad (2)$$

其中 w 為垂直於超平面之法向量 (Normal Vector) 及 b 為資料點的偏移量。對於要將資料分為兩類別，則要將資料點帶入決策函式做計算。當輸入一筆測試資料時，就可依據決策函數的值來分類。將經決策函數計算後小於 0 的資料歸類於 $f(x) < 0$ 類別，而大於 0 的資料歸類於 $f(x) > 0$ 類別。如此一來，即可根據決策函式計算後的正負值判別該資料點屬於哪一類別。

兩類別中距離分類超平面所行成超平面之最近數個資料點，稱為 Support Vector。那在線性可分的情況下，就會存在一個最佳分類超平面 [15] 使訓練資料集能夠完全分開，最佳分類超平面可表示成：

$$f(x) = w^T \mathbf{x}_i + b = 0 \quad (3)$$

得知最佳分類超平面與兩超平面間的關係後，從所有資料點中任選一個為 \mathbf{x} ，而最佳分類超平面與 \mathbf{x} 距離最近且垂直的點為 x' ，因此我們設至 \mathbf{x} 至 x' 的距離為 d ，也就是 \mathbf{x} 至最佳分類超平面間垂直的距離，其關係為

$$\mathbf{x} = x' + d \cdot \frac{w}{\|w\|} \quad (4)$$

根據方程式 (4)， \mathbf{x} 為 x' 加上與 x' 同方向且長度為 d 的向量，並經由移項後便可得到

$$x' = \mathbf{x} - d \cdot \frac{w}{\|w\|} \quad (5)$$

將 x' 代入方程式 (3) 中，即可得

$$d = \frac{|w^T \mathbf{x} + b|}{\|w\|} \quad (6)$$

當資料點 \mathbf{x} 為 Support Vector 時， d 的兩倍就是我們所希望得到的 Margin，且稱此函數的邊界為 Support Hyper-plane，可表示成：

$$\begin{aligned} w^T \mathbf{x}_i + b &\geq +1 \text{ for } y_i = +1 \\ w^T \mathbf{x}_i + b &\leq -1 \text{ for } y_i = -1 \end{aligned} \quad (7)$$

合併可得

$$y_i (w^T \mathbf{x}_i + b) - 1 \geq 0 \quad (8)$$

接著使用 Lagrange Multiplier 來求得最佳化分類超平面，也就是

$$\begin{aligned} \text{Maximize : } W(\alpha) &= \sum_{i=1, j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \\ \text{subject to : } \alpha &\geq 0, \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = 0 \end{aligned} \quad (9)$$

最後利用 Fletcher 學者於 1987 年從 Wolft 的理論提出二次規劃將問題 [16] 轉化為滿足方程式

(9) 的 α_i 集合，且 $w = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$ 中 $\alpha_i = 0$ 的樣本對分類並無作用，只有當 $\alpha_i > 0$ 的樣本才對分類起作用，稱為支持向量 (Support Vector)，如圖 8 中的白色資料點。若有 n 個支持向量，即可將 w 改寫為 $w = \sum_{i=1}^n \alpha_k y_{ki} x_k$ ，最後分類函數為

$$f(x) = \text{sgn} \left(\sum_{i=1}^n y_i \alpha_i (x \cdot \mathbf{x}_i) + b \right) \quad (10)$$

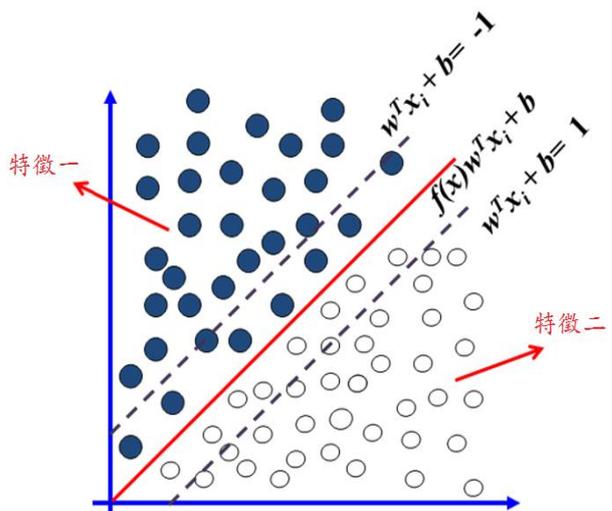


圖 7：二維空間資料點示意圖

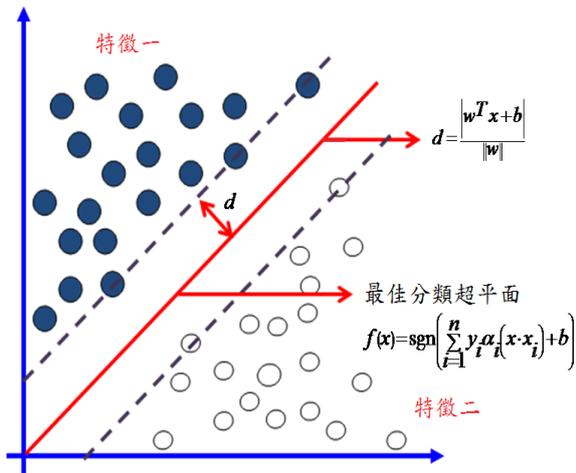


圖 8：分類最佳化示意圖

3. 實驗

實驗中我們於實際情況下錄製三種不同背景環境下靜止中的救護車鳴笛聲，第一種實驗環境為在救護車旁錄音，以取得單純的救護車鳴笛聲，做為實驗中支持向量機訓練時所用的資料；第二種實驗環境為在道路上錄製單純的汽、機車聲音，並無救護車鳴笛聲；而第三種實驗環境為汽、機車與救護車鳴笛聲同時存在時的聲音，其收音距離分別為 5 公尺、10 公尺、25 公尺、30 公尺、40 公尺及 50 公尺。

圖 9 至圖 16 分別為依據上述限定條件所得到的實驗結果圖，圖 9 為乾淨救護車鳴笛聲訓練結果圖；圖 10 為單純汽、機車聲音，無救護車鳴笛聲辨識結果圖，圖 11 至圖 16 分別為距離 5 公尺、10 公尺、25 公尺、30 公尺、40 公尺及 50 公尺經支持向量機計算後實驗結果圖。

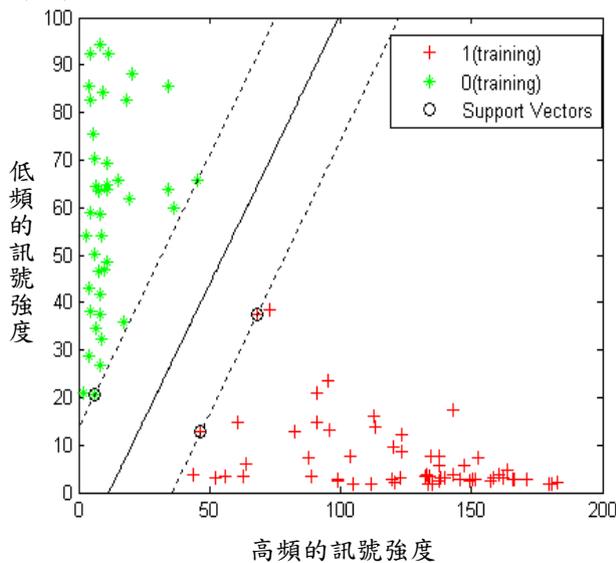


圖 9：乾淨救護車鳴笛聲訓練結果

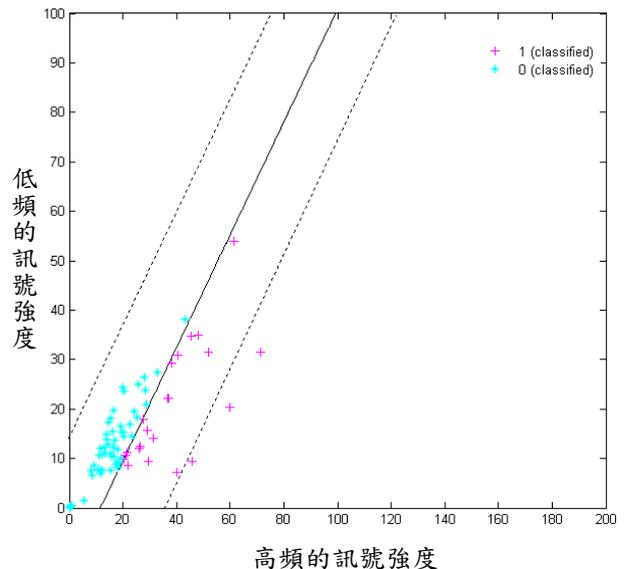


圖 10：單純汽、機車聲音辨識結果圖

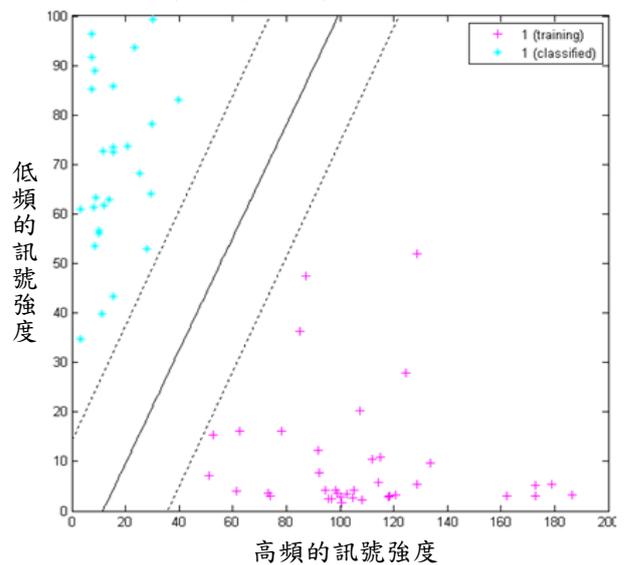


圖 11：5 公尺之救護車鳴笛聲辨識結果

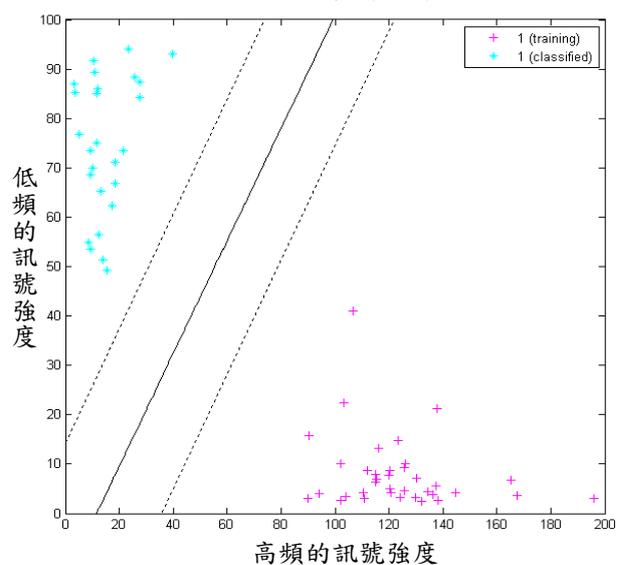


圖 12：10 公尺之救護車鳴笛聲辨識結果

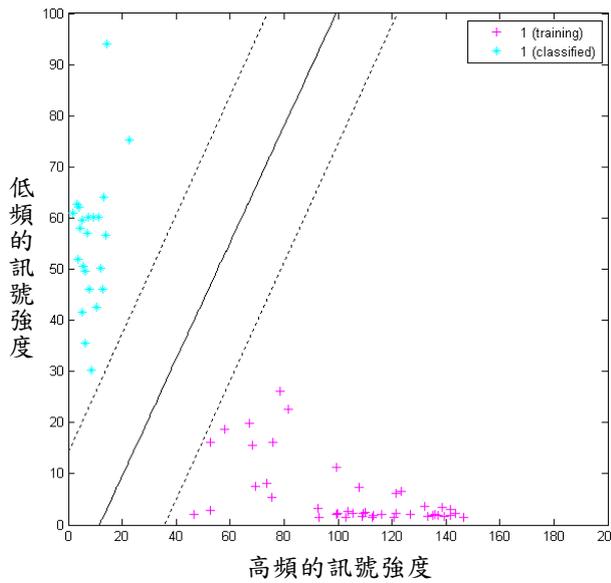


圖 13：25 公尺之救護車鳴笛聲辨識結果

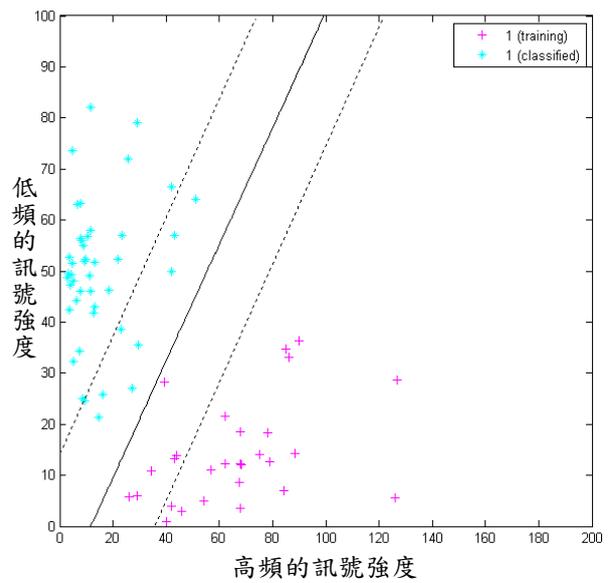


圖 16：50 公尺之救護車鳴笛聲辨識結果

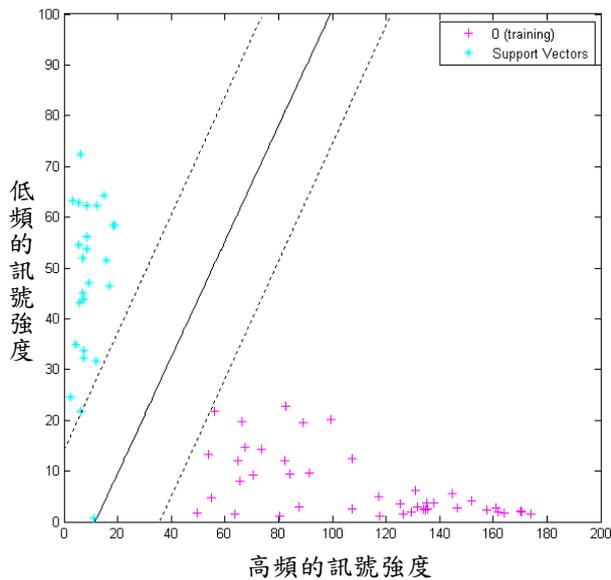


圖 14：30 公尺之救護車鳴笛聲辨識結果

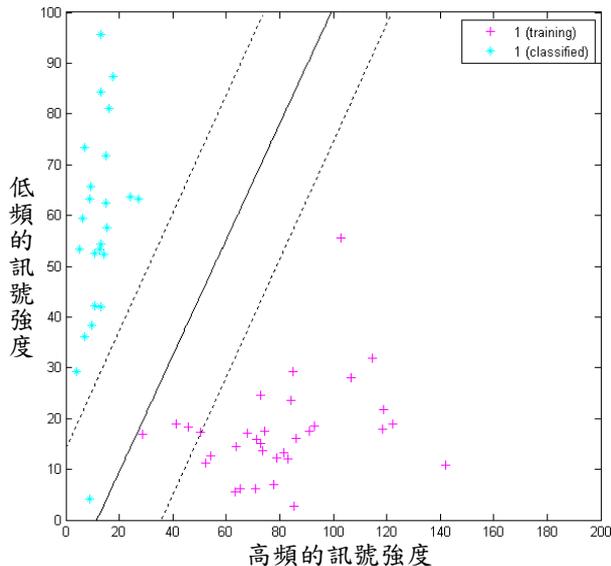


圖 15：40 公尺之救護車鳴笛聲辨識結果

表 1 為我們實驗中依據不同的收音距離所測得的辨識結果。其中救護車鳴笛聲辨識準確率的計算方式為

$$\frac{\text{辨識正確數量}}{(\text{辨識正確數量} + \text{辨識錯誤數量})} \times 100\%$$

表 1：不同距離下的救護車辨識準確率

距離(公尺)	準確率(%)
5	100
10	100
25	100
30	99
40	95
50	86

4. 結論

本研究提出利用支持向量機方法來辨識聲音中是否存在有救護車所發出的鳴笛聲。辨識方法分為訓練過程及識別過程。在訓練過程中，我們藉由擷取一段乾淨且清晰的救護車鳴笛聲來進行訓練，並將聲音訊號切割成長度為 0.1 秒的音框，再擷取出我們所需的高、低頻成份，做為支持向量機訓練的樣本。再依支持

向量機的方法產生分類平面。在實際接收道路上的聲音訊號後，同樣經提取音框的高頻及低頻資訊，判別是否存在救護車的高頻或低頻成份，確認是否有無救護車鳴笛聲的出現。實驗結果中顯示，在救護車鳴笛聲距離 25 公尺內，辨識準確率達 100%。當救護車鳴笛聲距離 50 公尺時，辨識準確率為 86%。

參考文獻

- [1]. Beritelli, F., Casale, S., Russo, A., and Serrano, S., "An Automatic Emergency Signal Recognition System for The Hearing Impaired," *Proceedings of 12th Digital Signal Processing Workshop and 4th Signal Processing Education Workshop*, Jackson Lake Lodge Grand Teton National Park, Wyoming, U.S.A., pp. 179-182, Sept. 2006.
- [2]. Milner, B. and Shao, X., "Clean speech reconstruction from MFCC vectors and fundamental frequency using an integrated front-end," *Speech Communication*, Vol. 48, No. 6, pp. 697-715, 2006.
- [3]. Jain, A. K., Mao, J., and Mohiuddin, K. M., "Artificial neural networks: a tutorial," *Computer*, Vol. 29, No. 3, pp. 31-44, 1996.
- [4]. Li, F. and Cox, T. J., "A neural networks model for speech intelligibility quantification," *Applied Soft Computing*, Vol. 7, No. 1, pp.145-155, 2007.
- [5]. Viadimir, N. V., "An Overview of Statistical Learning Theory," *Journal of IEEE Transcation on Neural Networks*, Vol. 10, No. 5, pp. 988-999, 1999.
- [6]. Vapnik, V., Golowich, S., and Smola, A. J., "Support vector method for function approximation, regression estimation, and signal processing," *Proceedings of the 9th Neural Information Processing Systems*, Denver, Colorado, U.S.A., pp. 281-287, Dec.1997.
- [7]. Burges, C. J. C., "A tutorial on support vector machines for pattern recognition," *Journal of Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 121-167, 1998.
- [8]. Fung, G., Mangasarian, O. L., and Shavlik, J., "Knowledge-based support vector machine classifiers," *Proceedings of the 14th Neural Information Processing*, Cornell University, New York, U.S.A., pp. 213-221, Dec. 2002.
- [9]. Joachims, T., "Text categorization with support vector machines: learning with many relevant features," *Proceedings of 10th European Conference on Machine Learning*, Chemnitz, Germany, pp. 137-142, Apr. 1998.
- [10]. Hearst, M. A., "Support vector machines," *IEEE Intelligent Systems*, Vol. 13, No. 4, pp. 18-28, 1998.
- [11]. Crammer, K. and Singer, Y., "On the learnability and design of output codes for multiclass problems," *Proceedings of Computational Learning Theory*, Stanford University, California, U.S.A., pp. 35-46, Jun. 2000.
- [12]. Joachims, T., "A statistical learning model of text classification with support vector machines," *Proceedings of 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, U.S.A., pp. 128-136, Sept. 2001.
- [13]. Joachims, T., **Learning to Classify Text Using Support Vector Machines**. Kluwer academic Publishers, 2001.
- [14]. Cooley, J. W. and Tukey, J. W., "An algorithm for the machine calculation of complex Fourier series," *Mathematics Computation*, Vol. 19, pp.297-301, 1965.
- [15]. Klaus-Robert, M., Sebastian, M., Gunnar, R., Koji, T. and Bernhard, S., "A Introduction to Kernel-Based Learning Algorithms," *Journal of IEEE Transaction on Neural Network*, Vol. 12, No. 2, pp. 181-201, 2001.
- [16]. Fletcher R., "A general quadratic programming algorithm," *Journal of the Institute of Mathematics and its Applications*, Vol. 7, pp. 76-91, 1971.