

解放你的雙手 - 飲水機智慧偵測停水裝置

107011153 梁會啟

107020008 翁俊傑

107020007 龔柏丞

Abstract — 為了省下每次裝水時都需要按三個按鍵的麻煩，我們希望設計出一個能偵測水量並自動停水的裝置。配合高中所學的空氣柱共振原理，我們藉由將聲音進行 FFT(Fast Fourier Transform)、並設計演算法找出基頻的位置，最後成功實作出一個會即時顯示當時基頻與剩餘空氣柱長度的裝置；其適用於至少三種水瓶、也可用於有少量雜音的環境，具有一定程度的通用性。

I. INTRODUCTION

相信大部分的人都有這樣的困擾：每次到飲水機裝水時都要依序按下「連續出水」、「溫水」、「連續出水」這三個按鍵，且還要隨時盯著水杯的水是否快要滿出來了，甚是麻煩。

Unit 1 的實驗中，我們學到如何偵測一個波的 peak (峰值)，以及使用快速傅立葉轉換 FFT(Fast Fourier Transform)的技巧；結合高中物理所教的空氣柱共振原理：水瓶內空氣柱的長度即為當前聲音基頻之波長的四分之一，因此隨著水瓶裡的水位上升、空氣柱縮短，其所發出的聲響之基頻也會跟著上升，所以在這個疫情肆虐的背景下，我們便想到可以藉由偵測裝水時發出的聲響之基頻，來實作一個能自動停水的智慧停水裝置。

經過上網搜尋之後，我們發現關於這部分的研究並不多，因此本次 project 的實作方式與 code 的設計方式，皆為我們這組原創的想法。

我們首先將訊號讀取並切成許多 frame、並對每個 frame 個別進行第一單元所學到的 FFT，接著通過 LPF(Low-Pass Filter)濾除高頻雜訊，再來使用我們開發出來的演算法找到當下聲音的基頻，最後則是用 real-time 的方式將基頻與當前剩餘的空氣柱長度顯示於螢幕上，此及本次實驗大致的步驟。

II. DESIGN METHOD

A. 基礎原理

空氣柱現象日常隨處可見，許多樂器包括笛、簫、管風琴等，都是利用空氣柱原理，改變空氣柱長度已發出不同頻率的聲音。在飲水機裝水的時候，水柱沖擊瓶底產生聲音源，而由於瓶子的結構讓瓶內產生空氣柱，使聲音在瓶子內形成駐波，以管內介質（即空氣分子）的振動位移而言，開口端為波腹，閉口端則為波節，其間可以有也可以沒有波節，故當空氣柱形成駐波時，管長為四分之一個波長的奇數倍，如 Fig.1 之左圖所示。若管長 L ，聲速 v ，則空氣柱內的頻率為 $f = \frac{nv}{4L}$ 。其中當 $n = 1$ 時的聲音稱為基音，其頻率稱為基頻；當 $n > 1$ 時的聲音稱為泛音，其頻率稱為泛頻。

因此，隨著水瓶內的水位漸漸上升，其所產生的基頻頻率將會越高，其示意圖顯示於 Fig.1 之右圖，而我們希望實際上錄到的裝水聲頻率也可以接近這樣的結果。

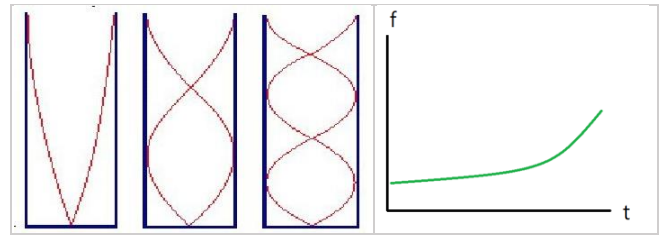


Fig. 1. (左)閉管空氣柱之駐波；(右)理想上基頻隨空氣柱縮短的變化

B. 前置作業

在了解到空氣柱共振的基本原理後，我們錄了一段飲水機裝水的聲音，然後透過 python 中的 librosa 對此段音訊做分析，以驗證理論公式的正確性。

首先，為了瞭解一段音訊隨著時間變化，其頻率變化的過程，我們對這段音訊做短時距傅立葉變換 STFT(Short-time Fourier transform)。它是傅立葉變換的一種變形，將長時間訊號分成數個較短的等長訊號，然後再分別計算每個較短訊號段的傅立葉轉換，如此便可以知道每一小段時間的頻率分布，以描繪出整段訊號接近連續的頻率變化。STFT 中有個重要的參數 n_fft ，代表取一小段時間的資料數，預設是 2048，也就是音訊每 2048 個資料點做一次 FFT，而頻譜就會是以 2048 個單位代表 0 到 $fs/2$ 赫茲的頻率範圍。因此，時域和頻域的解析度會是呈現反比關係，若 n_fft 數值取較大，則頻域解析度越好，但就會感覺時間上不夠連續，反之亦然。由於我們希望做完 STFT 後能夠有較好的頻率解析度，以提高找基頻的準確度，因此將 n_fft 調高為 8192，且我們錄音裝置的取樣率為 48000Hz，這樣系統約每 0.17 秒就會計算一次基頻頻率，也足夠用於 real time 的偵測與及時停水。

下方 Fig.2 左圖是原音訊經過 STFT 後的時頻譜，其橫軸是幀數(frame)，單位為時間，縱軸是取過 \log_2 的頻段，單位為赫茲。而右圖是將左圖中每個 frame 取平均加上兩個標準差當作 threshold，過濾後的時頻譜圖，以凸顯基頻。由 Fig.2 可以看出飲水機在裝水時，的確會有類似空氣柱共振的現象發生，且利用常見的錄音設備如手機或電腦也都可以錄到十分清晰的基頻聲，因此證實我們的確可以用這樣的方法進行智慧偵測停水。

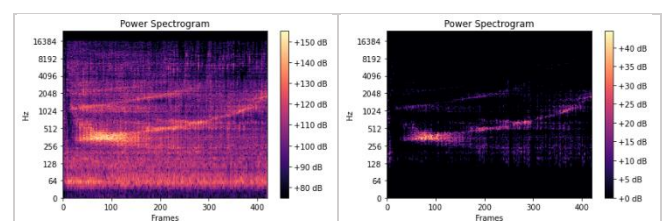


Fig. 2. 原音訊經過 STFT 後的時頻譜(spectrogram)

接著，有了音訊的時頻譜後，我們希望能透過簡單的方式直接找出每段時間(frame)下的基頻數值。由於 Fig.2 中基頻的位置幾乎都在振幅較大的區域(顏色較亮)，因此我們決定先以找每一 frame 的 peak 值來得出基頻頻率。下方 Fig.3 是將 Fig.2 中左圖取每一 frame 的最大值所得到的結果圖，由於瓶深 22.5cm、錄音時的溫度約為 20°C，因此根據閉管空氣柱的頻率公式，其基頻約為 380Hz，而取完 peak 後的基頻預測值為約為 360Hz；在水位離瓶口剩約 5cm 時，根據公式此時的基頻約為 1715Hz，而取完 peak 後的基頻預測值約為 1780Hz。可見大致上結果都符合理論的基頻值，不過受到雜訊的影響，有某幾個時間點的峰值並不在正確的基頻位置，因此，單純透過找最大值無法完全解決找基頻的問題，我們需要找出更複雜更準確的演算法。

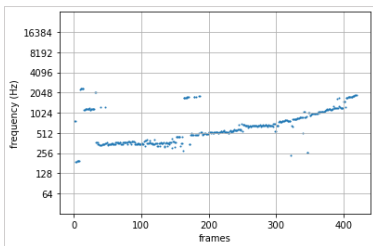


Fig. 3. 原音訊經過 STFT 後再取 peak 的散布圖

C. 演算法介紹

為了解決在前置作業中，取 peak 容易被雜訊給影響的情況，我們透過訊號加強與濾波來得到較理想的音訊訊號，再透過限縮範圍與交叉驗證的方式來加強找基頻的準確性，另外也透過連續峰值偵測的算法來降低系統對雜訊的敏感度。

1. 訊號加強與濾波

原始音訊存在許多雜訊，因此經過傅立葉轉換後，其頻率成分也較為雜亂，如 Fig.4(a)所示。我們先對此頻譜訊號轉成分貝(dB)以強化頻率的功率差異，如 Fig.4(b)的綠線所示，接著為了消除頻率訊號中的鋸齒狀雜訊，我們設計了一個 5 階的 Butterworth 低通濾波器，其中截止頻率(cut-off frequency)設為 250Hz，而經過濾波器後的頻率訊號顯示於 Fig.4(b)的紅線，這樣較為平滑的頻譜曲線有利於更準確地找峰值。

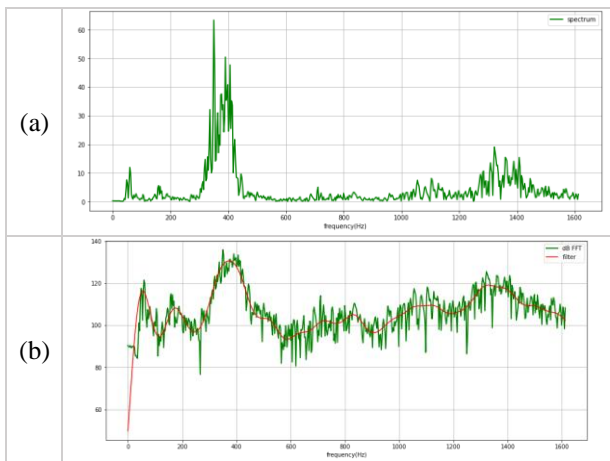


Fig. 4. (a) 原音訊經 FFT 後的頻譜圖。(b) 縱軸轉為分貝(dB)與經過低通濾波器後的頻譜圖。

2. 限縮範圍

在前置作業中，我們將每一 frame 做 FFT，得到的頻譜訊號再整段找最大值當作基頻頻率，然而，一般的水瓶基頻頻段不會有這麼大的範圍。一般情況下水瓶高度介於 10cm 到 30cm 之間、室內環境溫度介於 10°C 到 30°C 之間，透過閉管空氣柱頻率公式可知，可能的基頻範圍會介於 250Hz 到 800Hz 之間，因此，我們找 peak 的範圍其實不需要用到整段頻譜(即 0Hz 到 24000Hz 之間)，其他非可能基頻範圍內的頻率訊號都可以當作環境雜訊。

3. 連續峰值偵測

儘管限縮找 peak 的範圍增加了偵測基頻的正確性，仍然會有許多環境雜訊的頻率可能落在 250Hz 到 800Hz 之間，也容易導致基頻的誤判。環境雜訊之頻率值是隨機分布，因此我們透過計算峰值頻率的連續次數來判定是否為裝水的基頻，其中基頻連續的定義是現在的 peak 與前一次 peak 之間的差距小於正負 5 的頻率單位(約為正負 30Hz)，因為錄音當下或是取 peak 時可能產生誤差；而我們設定連續次數為 8 個 frame(相當於 1.36 秒)，此參數會隨著飲水機所在位置而定，若周遭吵雜，則連續測數需要調高以降低系統對雜訊的敏感度，若周遭較安靜，則可以調低以增加系統偵測速度。

4. 交叉驗證

當連續峰值未達一定次數時，系統會處於初始狀態，也就是沒有在裝水；當連續峰值達一定次數後，代表目前高機率是在裝水，因此進入偵測基頻狀態。

由於瓶內水位隨著時間漸漸變高，基頻頻率也會是連續的升高，兩次 peak 之間應該會是接近連續的關係，若兩次之間的頻率差太多，則可以推斷此 peak 並非是我們所要的基頻頻率，因此我們透過設定下一次 peak 值可能的範圍來找連續的 peak。由於裝水時頻率會愈來愈高，且在水位接近瓶口處的頻率變化會十分劇烈(根據公式，瓶深 L 愈接近零，則頻率 f 愈大)，下一次 peak 會比前一次的大許多，因此需要設定較大的高頻範圍；另外前一次 peak 可以因為錄音時或是取 peak 時的誤差，因此也需要設定低頻範圍。經過幾次測試後，我們發現往上頻率範圍設定 15(約 90Hz)、往下頻率範圍設定 5(約 30Hz)會是最佳值。

Fig.5 是結合上述的演算法後，對原音訊進行的基頻頻率預測圖，一開始幾個 frame 是環境雜訊，因此 peak 值會隨機分布，約在第 30 個 frame 才開始裝水，而我們也及時偵測到。而對比 Fig.3 只取最大值的方法，可以看出透過我們設計的演算法，能有效排除環境雜訊所造成的錯誤峰值，並且幾乎正確預測基頻。

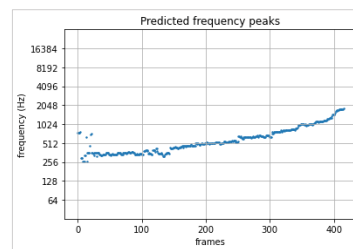


Fig. 5. 基頻頻率預測圖

D. Real-Time 系統

我們的智慧偵測停水系統可以實現即時偵測、即時顯示、可重複使用、不占用記憶體等功能。部分實測結果顯示於 Fig.6，說明如下：

- 當偵測到頻率大於 1200Hz 時(自行設定)，會顯示當前頻率(freq)、剩餘水深(rest)以及停止紅線(stop)，告訴飲水機與使用者將會停水。
- 隨著時間繼續，停止線會跟著往後，其與真實停止時間之相對位置不變。而系統會重新回到初始狀態，繼續偵測下一次裝水。
- 當下一次裝水快滿時(頻率達到 1200Hz)，紅色停止線會更新至最新的停止時間。
- 當顯示長度超過 300 個 frame 時，會將最舊的 peak 值刪除，所有值往前移一格，以維持顯示之長度、密度，並且節省電腦所需記憶體。

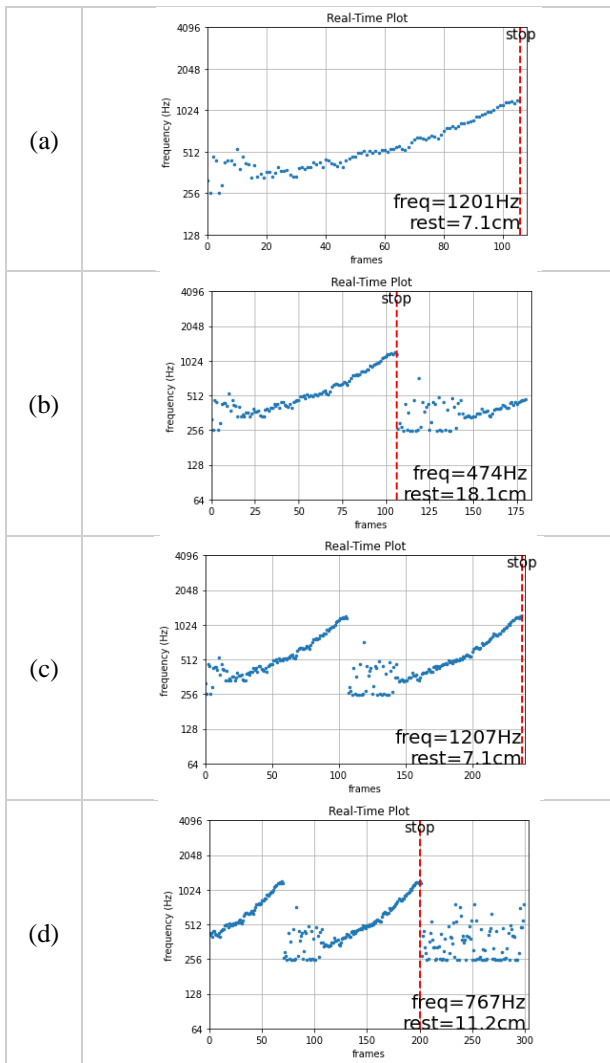


Fig. 6. 實測結果圖(擷取)

經過實測後，我們的偵測停水系統的確可以及時量測基頻頻率，並推算出理論剩餘水位，不過此為理論上的剩餘水深，並不一定和實際的一樣。

III. RESULTS AND DISCUSSION

對本組設計的基頻偵測自動停水系統進行兩項實驗測試，分別是通用性測試與訊雜比測試。

A. 通用性測試

為了測驗本次實驗所設計偵測方法的通用性，共對四種常見裝水容器進行測試。測試方式採柱水時以手機在瓶口處收音後，再進行分析。四種測試容器分別為：

- 塑膠水壺(圓柱體，瓶內直徑 6.5cm，瓶高 22.5cm)。
- 金屬保溫瓶(圓柱體，瓶內直徑 5.0cm，瓶高 22.0cm)。
- 寶特瓶(方柱體，寬度 5.7cm，瓶高 23cm)。
- 馬克杯(圓柱體，杯內直徑 8.5cm，杯高 8.8cm)。

TABLE I. 通用性測試結果

(左圖：實際錄音的時頻譜圖。右圖：黑線代表以空氣柱公式換算理想的基頻變化；藍點為頻譜每個 FRAME 中的最大值；紅點為利用本組設計的演算法所計算出的連續基頻變化。)

瓶種	Spectrogram	Result
(a)		
(b)		
(c)		
(d)		

四種容器中 A、B、C 的左圖(頻譜圖)都可以發現肉眼可見的基頻連續變化，D 則不明顯。比較右圖中紅線(本組設計的演算法所計算出的連續基頻變化)與左圖中的基頻變化曲線相符合，表示在 A、B、C 三種容器下，演算法都能夠有效找到基頻的連續變化，而在 D 容器可以發現 frame 大於 125 後基頻處沒有明顯的最大值，因此演算法找到的基頻也較不準確。

比較實際的基頻曲線與理論的基頻變化，容器 A、B 與理論值相近，能夠由基頻準確回推瓶內剩餘高度並在即將滿水時顯示停水。而容器 C(寶特瓶)的實際基頻曲線與理論值的差別較大，實際基頻小於理論值約 200Hz。詳細原因本組尚未細探，但推測可能與容器材質有關。以本次實驗的四種容器相互比較，寶特瓶的材質最具彈性，其次是塑膠水壺，而金屬的保溫瓶與陶瓷的馬克杯屬於彈性較低的材質。因此可以推測容器材質彈性越低，基頻越接近理論值。然而，即使容器 C 實際基頻與理論基頻有 200Hz 的差距，在接近滿水時基頻仍然會達到 1200Hz，因此若將停水頻率設為 1200Hz，仍然可以在即將滿水之前顯示停止。

容器 D(馬克杯)是本次實驗中偵測效果最差的，其中包含了兩個問題。第一，由於馬克杯杯口較寬且高度較低，因此柱水時發出的聲音不如其他容器大，在水量快滿時收音效果不佳。第二，由於馬克杯的高度只有 8.8cm。在達到停水頻率 1200Hz 時，杯內剩餘高度約為 6cm。6cm 對其他容器是即將滿水的高度，但對馬克杯而言杯內此時只裝了 2.8cm 的水量，此時自動停水是不合理的。針對這項問題的改善方式:利用初始注水時的基頻換算杯子原始高度，以比例的方式來決定停水頻率。例如 8.8cm 的馬克杯，當水裝到 80%時即停止。此時杯內剩餘高度為 1.7cm，換算停止頻率約為 5000Hz。

綜合以上幾點，目前基頻偵測自動停水系統能夠成功應用在長條的柱形容器，且適用於不同形狀與材質。對於杯型的容器，雖然仍可偵測到基頻，但存在訊雜比過高以及停水時杯內剩餘空間比例過大的問題。

B. 訊雜比測試

測試方式為錄製一段裝水音檔後，加入白雜訊 AWGN(Additive white Gaussian noise)，測量偵測系統可容忍的最高訊號雜訊比(Signal-to-noise ratio)，其中，訊雜比的公式如下：

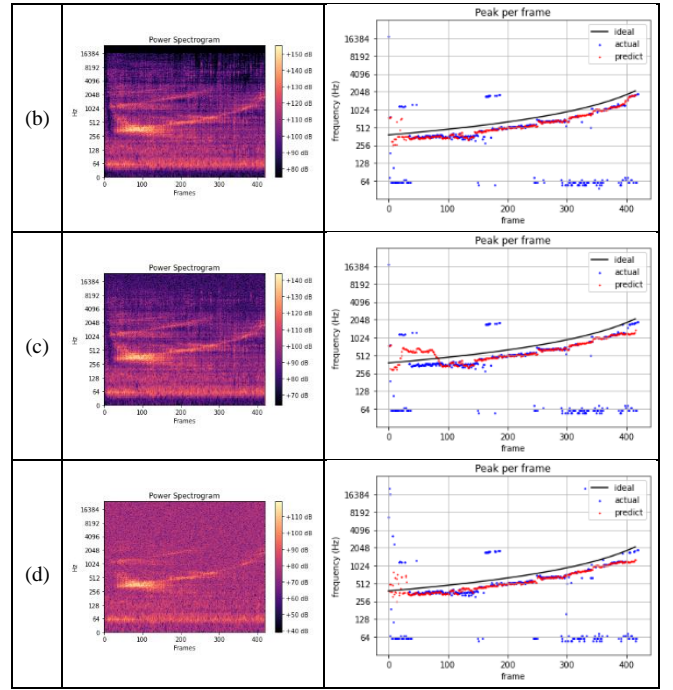
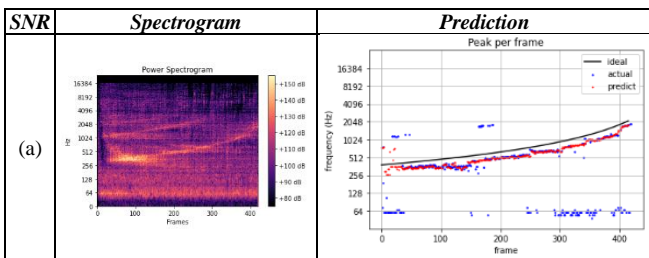
$$SNR (dB) = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right)$$

AWGN 指的是一種功率函數是常數，且振幅服從高斯分布的雜訊訊號，其特性在於在任何頻率上皆有相同的功率密度，因此在訊號處理領域中十分有用。以下四種音訊的訊雜比分別為：

- (a) 原始音檔 (未加雜訊)
- (b) 加 SNR=40 的雜訊 (雜訊之功率為原訊號的 10^{-4} 倍)
- (c) 加 SNR=20 的雜訊 (雜訊之功率為原訊號的 10^{-2} 倍)
- (d) 加 SNR=0 的雜訊 (雜訊之功率與原訊號相等)

TABLE II. 訊雜比測試結果

(左圖：實際錄音的時頻譜圖。右圖：黑線代表以空氣柱公式換算理想的基頻變化；藍點為頻譜每個 FRAME 中的最大值；紅點為利用本組設計的演算法所計算出的連續基頻變化。)



由測試結果可以看出，當 SNR 低於 40 時，其音訊的時頻譜以及利用我們找峰值的演算法所得出的預測曲線，都幾乎和不加雜訊時的一樣。而當 SNR 達到 20 以上的時候會發現，其音訊的時頻譜明顯較原本的雜亂，基頻的部分也較不清楚，且儘管在裝水的前期(也就是頻率低於 1000Hz)仍然可以正確預測頻率，但是在裝水的後期因為頻率上升幅度過快，再加上雜訊的影響，導致連續變化的基頻無法被正確偵測。

這樣的結果會對自動停水的功能造成十分嚴重的影響，因為頻率上不去會讓此系統以為水位尚未接近瓶口，因而繼續出水，使自動停水的功能異常。因此，此系統還不適合應用在周遭雜訊較大的地方，不過當收音麥克風置於瓶口附近，大多數情況下的日常噪音並不影響偵測結果。

除以上兩點測試比較外，這項技術還存在一項實施上存在的困難。在進行實驗時，本組測試了多台飲水機，發現並不是每台都能夠成功偵測到基頻。原因在於飲水機的出水量以及水流型態。若飲水機出水時的水柱穩定且水量不大時，此時瓶內的水擾動較低，瓶內發出聲響不大，偵測效果很差。

若要實際運用此偵測技術，飲水機出水是需要被控制的，且需特別設計。必須在注水時讓瓶內的水有一定的擾動以製造聲響(例如讓飲水機出水時參雜些微氣泡使水流不穩定，或使用多道水柱製造擾動)，同時避免聲響過大影響使用者。

IV. CONCLUSIONS

結合了 Unit 1 的實驗技巧以及高中物理的空氣柱共振原理，我們成功設計出一套演算法來偵測當前聲音的基頻位置，並藉此實作出一個能自動偵測水位並且自動停水的裝置。

在偵測波峰位置的部分，我們首先將訊號讀取進來，並將每 8192 個 data point 切成一個 frame、接著對每個 frame 個別進行第一單元學到的 FFT 得到其頻譜；對頻譜取 log 後，將其通過一個 LPF 以濾除高頻的雜訊，最後得到一條平緩的曲線，我們希望該曲線的波峰位置即為基頻的位置。

然而由於雜訊的影響，有時找到的波峰位置並非基頻的位置，而是奇怪的頻率；根據空氣柱共振的原理，因為空氣柱的縮短是連續的，所以頻率的上升亦是連續的，因此我們設計了一種演算法，在前一個基頻的位置的前後設定一小段區間、在此區間內找波峰並將之視為目前的基頻位置；藉由此演算法，我們有效地排除了雜訊的影響，能成功在水快到達瓶口時發出停水的通知。

總結我們的實驗成果，除了成功找出基頻位置、將其以即時的方式顯示出來之外，我們也於水快滿時發出提醒；而經過我們的實際測試，此裝置不僅能適用於有少量雜訊的環境：如教室走廊、桌球室等，它也同時適用於至少三款不同的水瓶，具有基礎的通用性質。

在疫情嚴重的此時，這項實作具有減少接觸、達到防疫效果的優點，然而這只是其中的一個意義所在；我們認為這個裝置更大的意義在於，它是一個以自動化替代人力的開端：本次實驗中，我們省下了按按鈕的人力成本，但是其應用不止於此，以家中用來煮開水的水壺為

例，其在水煮沸時會發出鳴笛聲，因此我們也可利用偵測基頻的方式，配合使用者的需求，設計出一個能自動熄火的智慧裝置。本次實驗的成果可以運用在許多將自動化替代人力的程序之上。

最後是關於未來可能的延伸方向。進行測試的過程中，我們發現此裝置仍然會有基頻位置偵測錯誤的情況，因此如何更有效地排除雜訊、提高位置偵測的準確率，是一個未來能繼續鑽研的方面。其次是音訊放大的部分，不論是在裝置內加入訊號放大器、或者是調控飲水機的出水量，若能達到訊號放大的效果，便能更有利於訊號的處理。最後是通用性的問題，雖然從理論公式的角度出發，偵測空氣柱長短的做法能適用於任何水瓶，然而現實畢竟無法滿足理論條件，瓶口形狀、瓶身材質、距離出水口的高度等因素都會影響最終的結果，如何建立一個更具有通用性的偵測方式，也是一個值得繼續深究的方向。

REFERENCES

- [1] Klapuri, A. (2003). Multiple fundamental frequency estimation based on harmonicity and spectral. IEEE Trans. Speech and Audio Proc., 11(6), 804-816.