
Introduction

這次的作業是需要將 hw5 找到的 high peaks 做 trajectory formation，將臨近的幀作對應的頻率的峰值相連，也被稱為峰值匹配。而對應的頻率一般意味彼此在各自幀內的頻率最接近。如果將峰值相連就等於去掉“幀”這個離散的存在。

Method

Trajectory formation 的做法我是參考論文 [1] 的 FRAME-TO-FRAME PEAK MATCHING 部分。假設每個 frame 的 frequencies 被定義為 ω_n^k ， k 是 frame 的數量， n 則是當前 frame 中 high peak 位置。設定一個迴圈將所有的 frame 透過設定一個 threshold 來檢查是兩個 frame 之間的高 peak 是否符合限制，如果超過代表當前的 ω_n^k 為“dead”且會移除記錄的位置。

$$|\omega_n^k - \omega_m^{k+1}| \geq \Delta \quad (1)$$

如果符合下公式 2 則會記錄起來：

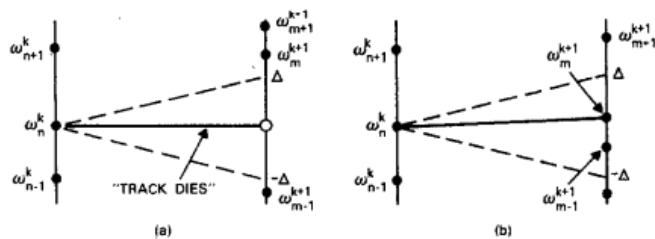
$$|\omega_n^k - \omega_m^{k+1}| < |\omega_n^k - \omega_i^{k+1}| < \Delta \quad (2)$$

接下來，會利用公式 3 檢查 ω_n^k 有沒有比較好的組合，如果沒有則確認 ω_n^k 與 ω_m^{k+1} 相連。如果有則將 ω_n^k 與 ω_m^{k+1} 連結消除並往 ω_{n+1}^k 搜尋。

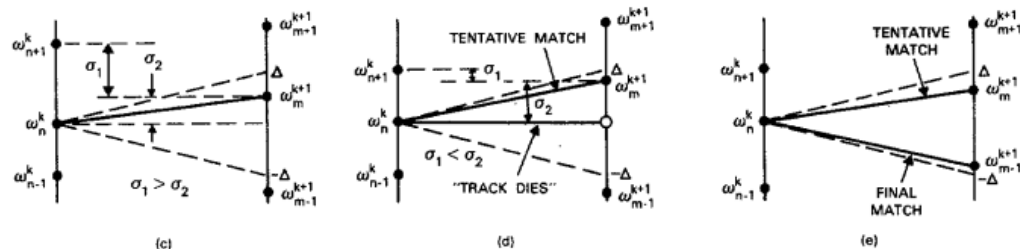
$$|\omega_m^{k+1} - \omega_n^k| < |\omega_m^{k+1} - \omega_{i+1}^k| \text{ for } i > n \quad (3)$$

如果在 ω_m^{k+1} 沒有找到對應的 peak 則會在 ω^k 給於一個 peak，則稱之“born”，詳細過程如下圖

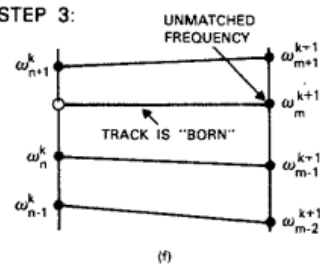
STEP 1:



STEP 2:



STEP 3:



Birth-death frequency tracker.

但我在實作時沒有在時間內完成，很可惜沒有辦法展示結果。

Extra

如果只使用 magnitude-only synthesis，合成得到的聲音品質其實並不是很好。我有找到三個方法能夠提高品質的做法

1. 提高 resolution

我將一個 frame 的長度減少且 frame 的總數量提升。

2. Linear phase to zero phase

這是參考網絡上[2]提供的方法，此方法如下圖。使用這個方法可以很有效的改善合成品質但作者並沒有解釋這個方法的原理。

```

for m=1:nFrames
    %tt = (m-1)*R+1:(m-1)*R+M;
    tt = 1:samPerFra;
    xFrame = zeros(N, 1);
    xFrame(tt) = xz((m-1)*hopSamples + tt);
    xFrame(tt) = xFrame(tt) .* w;
    %linear phase to zero phase
    xFrame = [xFrame(floor(samPerFra/2)+1:end);
              xFrame(1:floor(samPerFra/2))];
    X = fft(xFrame, N);

```

Linear phase to zero phase

3. 使用 phase 做合成

雖然人耳聽不出正弦波與餘弦波的差異，但某些聲音其實是相對敏感，下圖是參考論文[3]的比較，可以看到出不帶相位的合成方法與原始的波形有所落差。

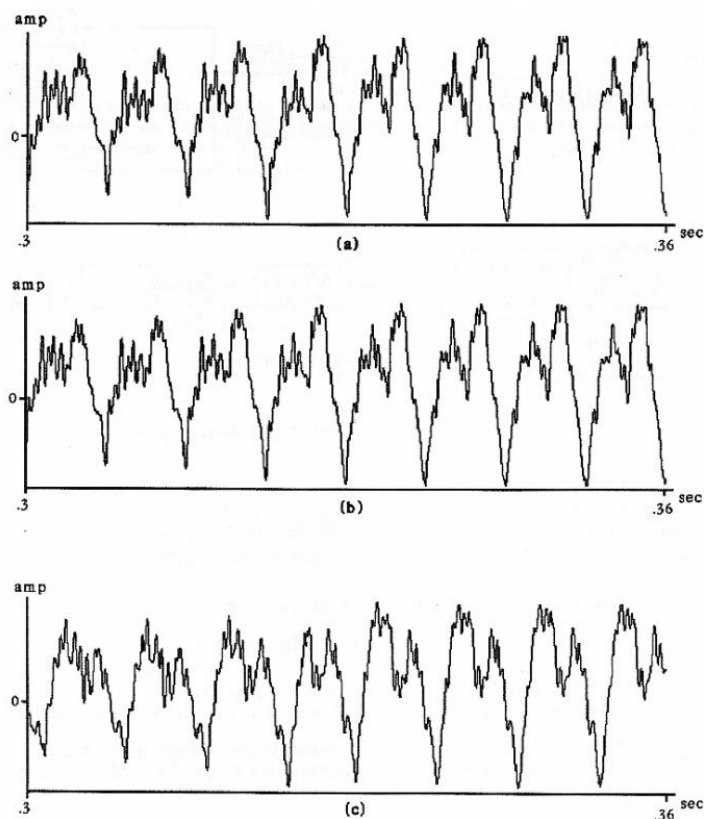


Figure 3.7: Sinusoidal synthesis example: (a) original cello sound, (b) synthesis using phase information, (c) synthesis without phase information.

Reference

1. [McAulay, R., & Quatieri, T. \(1986\). Speech analysis/Synthesis based on a sinusoidal representation. IEEE Transactions on Acoustics, Speech, and Signal Processing, 34\(4\), 744–754. doi:10.1109/tassp.1986.1164910](#)
2. [能用三角函數表示聲音嗎——正弦模型綜述](#)
3. [Serra, X. 1989. “A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic Plus Stochastic Decomposition.” PhD thesis, Stanford University.](#)