

(1) Write a Matlab or Python program that can convert a numbered musical notation (簡譜) into a music file (*.wav).

程式設計功能:

基本功能：給一簡譜、拍子，能輸出一個音樂檔(.wav)

進階功能：

1. 可以設定每個音階的振幅大小。
2. 每個音階發聲時，可以設定聲音會隨著時間 decay，衰減方式有兩種，一種為 linear decay，其斜率自動計算 $m = -\text{amplitude}/\text{total_time}$ ；另一種衰減方式為 exponential decay，並且可以設定衰減係數 alpha。
3. 可以設定每個音階其倍頻的數量要為多少，ex: f_0 、 $2*f_0$ 、 $3*f_0$。
4. 可以設定休止符。

透過以上設定，加入了進階功能後，我們的音樂聽起來會比較舒服，也比較像是真的音樂。

```
clear all;
%% 功能設定
score = [1, 1, 5, 5, 6, 6, 5]; % 1: Do, 2: Re, 3: Mi, ...
beat = [1, 1, 1, 1, 1, 1, 2]; % 拍子
name = 'twinkle';
rest = [1, 0, 0, 0, 0, 0, 0]; % 休止符，若要在某個音調後面加上一個四分休止符則value設為1，若不要休止符則設為0
chord_number = [3, 3, 3, 3, 3, 3, 3]; % 設定倍頻的數量 ex:  $f_0$ 、 $2*f_0$ 、 $3*f_0$ 、.....
amplitude = [1, 1, 1, 1, 1, 1, 1]; % 設定振幅
decay_type = 'linear'; % 可以設定聲音隨 "linear" 衰減，或者是 "exponential" 衰減
alpha = 0.2; % 若設定 "exponential" 衰減，則需要設定 衰減係數 來確定衰減的快慢

getmusic(score, beat, name, rest, chord_number, amplitude, decay_type, alpha)
```

(2)

R10945004

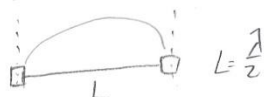
孫欽銘

(a)

當句子中兩個音節都只含母音時，就無法透過 amplitude 來判斷是一個音節還是兩個音節。

Ex: 注意 → 出 x' -
失 - 子 母 母

(b) 以共振的觀點



$$\text{共振: } L = \frac{\lambda}{2} n, \quad n \in \mathbb{N}$$

$$\lambda = \frac{2L}{n}, \quad f = \frac{v}{\lambda} = \frac{vn}{2L}$$

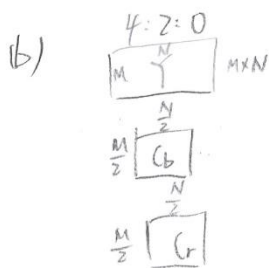
$$f_0 = \frac{v}{2L} \Rightarrow f = n f_0$$

所以除了基頻 f_0 之外， $n f_0$ 處也會有比較大的能量產生，因此會有和弦產生。

- (3) (a) ① 因為樂曲在頻域上有倍頻的現象, $f_0, 2f_0, 3f_0, \dots$
 ② 節拍上有一致性。
 ③ 旋律會重複

- (b) ① 因為卡通顏色變化較為固定, 因此只計算邊緣區域
 ② 在 edge zone 用線段 or 圓弧取代降低資料量

- (4) (a) 因為 YCbCr 比較符合人的眼睛對色彩的認識, 而這也與人的眼睛中各種視錐細胞的比例相關, 在 RGB space 中比較難以判斷何種 channel 較為重要, 而 YCbCr 我們可以知道 Y 是比較重要的 channel, 而可以對 CbCr 做壓縮。



$$MN + \frac{MN}{4} + \frac{MN}{4} = \frac{3}{2}MN \quad (4:2:0)$$

$$3MN \quad (4:4:4)$$

$$\therefore \text{compress ratio} = \frac{\frac{3}{2}MN}{3MN} = \frac{1}{2} \Rightarrow 50\%$$

(5)

a)

① 與 DFT 相比, DCT 運算的結果不會產生複數, 因此在運算上相對簡單, 所需資料量也較少。

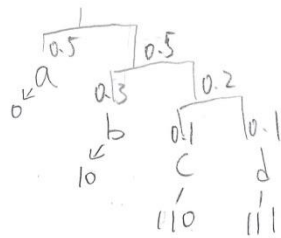
② 與 KLT 相比, DCT 的 kernel 為固定的, 而 KLT 需要因應不同輸入而額外新增一筆資料量儲存 K 矩陣, 所需運算空間會較大。

b) (ii) DC Difference

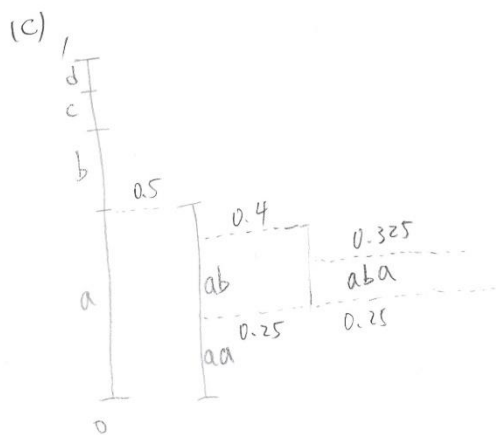
(iii) zigzag

(iv) the Huffman code

(b) (a)



$$\begin{aligned} \text{Entropy} &= 0.5 \ln 2 + 0.3 \ln \frac{10}{3} + 0.1 \ln 10 + 0.1 \ln 10 \\ &= 0.3466 + 0.3612 + 0.23 + 0.23 \\ &= 1.1678 \end{aligned}$$



$$1, a \rightarrow [0 + 0 \times (1-0), 1 + 0.5 \times (1-0)] = [0, 0.5)$$

$$2, ab \rightarrow [0 + 0.5 \times (0.5-0), 0 + 0.8 \times (0.5-0)] = [0.25, 0.4)$$

$$3, aba \rightarrow [0.25 + 0 \times (0.4-0.25), 0.25 + 0.5 \times (0.4-0.25)]$$

$$= [0.25, 0.325)$$

$$\text{lower} = 0.25, \text{upper} = 0.325$$

$$0.25 \leq C \cdot 2^{-b} < (C+1) \cdot 2^{-b} \leq 0.325$$

$$\frac{4}{16} \downarrow \quad \frac{5}{16} \downarrow$$

$$\therefore C=4, b=4, \text{用 4 個 bit 表示 4}$$

$$\therefore aba = 0100 \neq$$

$$\text{ceil}\left(N \times \frac{\text{entropy}}{\log K}\right) \leq b \leq \text{floor}\left(N \times \frac{\text{entropy}}{\log K} + \log K^2 + 1\right)$$

$$\Rightarrow \text{ceil}\left(\frac{100000 \times 1.6828245}{\ln 2}\right) \leq b \leq \text{floor}\left(\frac{100000 \times 1.6828245}{\ln 2} + 1 + 1\right)$$

$$\Rightarrow \text{ceil}(168549.5) \leq b \leq \text{floor}(168549.5)$$

$$168548 \leq b \leq 168549$$

7. ① 兩個聲音訊號頻率接近

② 兩個聲音訊號頻率一樣，只有 phase 不一樣

③ 兩個聲音訊號一樣，只是出現的時間點有差異

Extra 尾教4 question:

Ans: 因為人類的視覺暫留為 $\frac{1}{16}$ 秒，所以當一秒鐘內放入連續 16 張以上的圖片看起來就會是動畫。