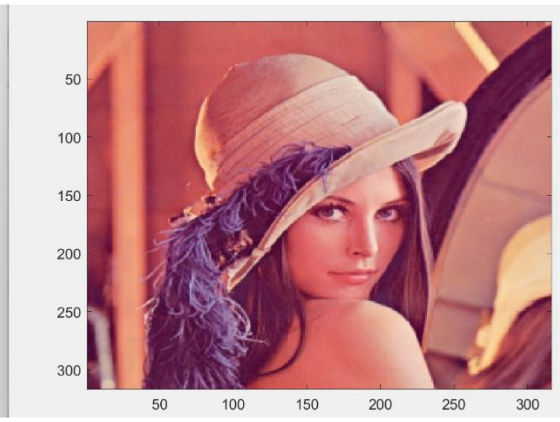
ADSP HW3 R12631055 林東甫

1.



2. ① 基本上都是實數，避免了一般作FT時會出現複數的情況。

② 因為 $Y[m] = \log\left\{\sum_{k=f_{m-1}}^{f_{m+1}} |X[k]|^2 B_m[k]\right\}$ 其中是Summation，因此為零的機率較低得多。

③ $f_1, f_2, f_3 \ldots$ 是以等比級數增加：$f_m = \alpha^m f_0$，而人類耳朵對頻率差異的感受是以比率差為主而非算數差，因此較為match.

④ Since $X[K] = X^*[-k]$, $|X[k]| = |X[-k]|$，因此改用cosine來取代 ~~inverse~~ IFT 也可以，除了上述避免了複數輸出，也可以減少計算量。DCT: (discrete cos transform)

3. For the inverse cepstrum $D_*^{-1}[\cdot]$

$$x[n] \longleftarrow \boxed{\text{IFT}} \xleftarrow{X(F)} \boxed{\exp} \xleftarrow{\hat{X}(F)} \boxed{\text{FT}} \longleftarrow \hat{x}[n]$$

以 inverse Z transform　　　　Z transform 取代

$\hat{x}[2] = 1 \quad \hat{x}[n] = 0 \ \forall n \neq 2 \quad \therefore \overset{\text{z-tran.:}}{\hat{X}(z)} = \hat{x}[2] z^{-2} = z^{-2}$

取exp. $X(z) = e^{z^{-2}}$, z^{-1}-trans.: by Taylor series $f(t) = f(t_0) + \sum_{n=1}^{\infty} \frac{f^{(n)}(t_0)}{n!}(t - t_0)^n$

$x[n] = \delta[n] + \sum_{n=1}^{\infty} \frac{\delta[n-2n]}{n!}$, $a_1 = -2$ $\quad x[n] = \begin{cases} \frac{(2)^n}{n}, & n > 0 \\ 0, & \text{otherwise} \end{cases}$

~~Ex：學號末5：~~

人耳聽得到 20～20,000 Hz，是以 80dB(以上)為標準

4.(a) $f = \frac{v}{\lambda} = \frac{340}{x} = 250$, $x = 1.36$

$\lambda : x = 2d = 1.36$, $d = 0.68$ Ans: 68 cm

(b)

$f_{Do} = 250$ $f_{La} = 2^{\frac{9}{12}} f_{Do} = 420.4482076\cdots$

$\lambda_{La} = 0.80866\cdots$, $d_{La} = 0.40433\cdots$

Ans: 40.4 cm

5.(a) ①音樂訊號的頻率分佈比較固定,只會在固定音高以及其倍頻有能量

②音樂訊號的拍子是固定的,代表固定的時間間隔才有訊號

③在單一音符內的頻率是固定不變的

以上使音樂訊號有一致性,容易壓縮

(b)

① Cartoon/mark image 的同個區域內的顏色大多一致

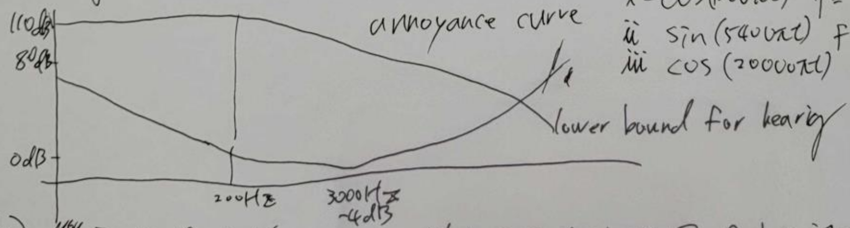
② Edge 通常是較簡單的線條,如圓、直線…。

6.

i $-\cos(1200\pi t)$ $f = 600Hz$

ii $\sin(5400\pi t)$ $f = 2700Hz$

iii $\cos(20000\pi t)$ $f = 10000Hz$



(a) 響度主要由振幅決定。若各方面條件相同,則三訊號之響度也應相同,惟人耳對 3000 Hz 最敏感,ii $\sin(5400\pi t)$ 可能最易聽到。

(b) 同上述,人耳對 3000 Hz 最敏感,最易聽見,~~故~~ 如 ii $\sin(5400\pi t)$,然而根據 annoyance curve,對 200 Hz 以下之高 dB 容忍度最高,因此對 i $-\cos(1200\pi t)$ 的高 dB 有較高的容忍度。

7. (a) ① DCT 相較 DFT，能量更為集中 (at zero)

② DCT output 為實數，DFT則為複數，需額外記錄虛數部，不利壓縮

③ DCT independent of input 相較於 KLT 依賴於輸入

(b) ① 在進行 JPEG 規範定義時，8×8 是經由實驗最佳化所得的結果。

② 8×8 相較 whole image 有較低計算複雜度，也有效降低 buffer size。

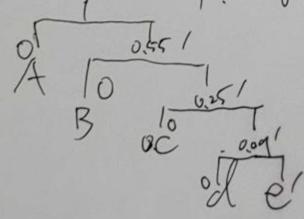
③ 在切分 8×8 後反映了 image 不同區域的特性，區分區塊內的一致性。

8. (a) $entropy = P(x='a')\cdot\log\left(\frac{1}{P(x='a')}\right) + P(x='b')\cdot\log\left(\frac{1}{P(x='b')}\right) + P(x='c')\cdot\log\left(\frac{1}{P(x='c')}\right) + P(x='d')\cdot\log\left(\frac{1}{P(x='d')}\right) + P(x='e')\log\left(\frac{1}{P(x='e')}\right)$

$= 0.45\log\frac{1}{0.45} + 0.3\log\frac{1}{0.3} + 0.16\log\frac{1}{0.16} + 0.06\log\frac{1}{0.06} + 0.03\log\frac{1}{0.03}$

$= 1.28773479$

(b)



A : 0 , B : 10 , C : 110

d : 1110 , e : 1111

(c) a : 1 b : 2 c : 3 d : 4 e : 4

$0.45\cdot 1 + 0.3\cdot 2 + 0.16\cdot 3 + 0.06\cdot 4 + 0.03\cdot 4$

$= 0.45 + 0.6 + 0.48 + 0.24 + \cancel{0.15}\ 0.12 = \cancel{1.92}\ \underline{1.89}$