姓名:葉冠宏　學號:R11943113 ADSP HW3

1.

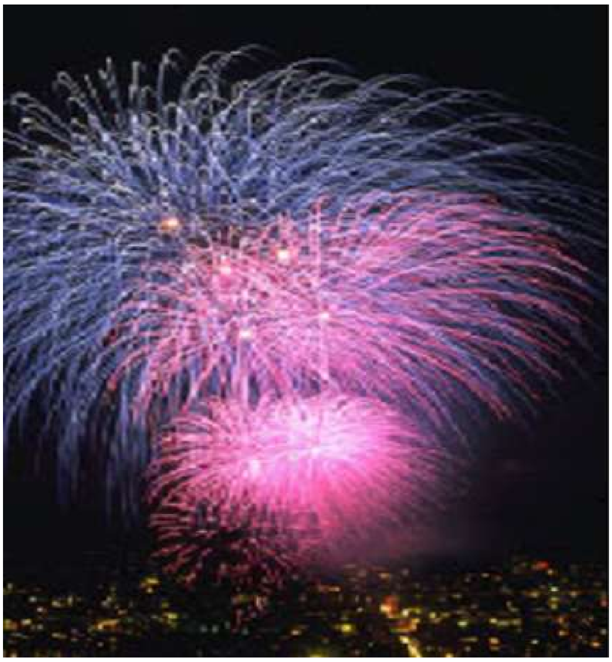
原始影像:

轉成 YCbCr 後 y CHANNEL 影像: 可以觀察到 Y channel 呈現原始影像大部分輪廓。

壓縮成 4:2:0 後 CB CHANNEL 影像:

壓縮成 4:2:0 後 CR CHANNEL 影像:

重建後影像:可以觀察到和原始影像差異不大



2.

(a) $y[n] = x[n] + 0.3x[n-15] + 0.2x[n-25]$

$\Rightarrow$ let $p[n] = \delta[n] + 0.3\delta[n-15] + 0.2\delta[n-25]$

$\cdot \log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$

$P(z) = 1 + 0.3z^{-15} + 0.2z^{-25}$

$\hat{P}(z) = \log(1 + 0.3z^{-15} + 0.2z^{-25})$，令 $x = 0.3z^{-15} + 0.2z^{-25}$

$\Rightarrow \hat{P}(z) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left[0.3z^{-15} + 0.2z^{-25}\right]^k$

$= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{m=0}^{k} C_m^k (0.3z^{-15})^m (0.2z^{-25})^{k-m}$

$= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{m=0}^{k} C_m^k\, 0.3^m \times 0.2^{k-m} \times z^{(10m-25k)}$

$\Rightarrow \hat{p}[n] = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{m=0}^{k} C_m^k\, 0.3^m \times 0.2^{k-m}\, \delta[n-(25k-10m)]$ #

(b) 可以在 $n = 25k - 10m$，$k \in (1,\infty)$，$m \in (0,k)$ 時

把雜訊 delay 去做移除 #

3、

(a)

$\cos(300\pi t)$: $f = 150$ ⇒ ≒ 50dB

$-\sin(1200\pi t)$: $f = 600$ ⇒ ≒ 10dB

$\sin(6000\pi t)$: $f = 3000$ ⇒ ≒ -40dB

⇒ (i) 可以聽得最大聲 #

(b)

$\boxed{C} = f \times \lambda$ ⇒ 頻率越大，波長越短 ~~短~~，長，可以傳得遠

⇒ (i) 可以傳得最遠 #

4、

(a)

Do 和 Mi 差 4 個半音，∴ 頻率為 $240 \times 2^{\frac{4}{12}}$

Do 和 So 差 7 個半音，∴ 頻率為 $240 \times 2^{\frac{7}{12}}$ #

(b)

$f = \frac{170}{L} n$，$L = \frac{170}{f} n$

⇒ Mi 時 length 為 $\frac{170}{240 \times 2^{\frac{4}{12}}} n$

So 時 length 為 $\frac{170}{240 \times 2^{\frac{7}{12}}} n$

#

5.

(1) ∵ energy in the f-domain is constrained at n fo Hz，∴我們可以只針對特定頻率做壓縮就好

(2) ∵音樂常會有 repeated melody，∴我們可以針對重複的音樂只壓縮其中一段就好

(3) ∵聲音會連續一段時間，∴可以只針對那一段時間的那個音高做一次壓縮就好 #

6.

(a) (i) ∵ DCT 只涉及到實數的計算，在表示上比較容易轉換成實際可以理解的物理意義

(ii) DCT 是連續函數，由 cos 函數所構成，而 DFT 不是，∴ DCT 具有可微分等特性

(iii) DCT 計算複雜度較低

(b) (i) frequency distribution of an image usually varies with the location

(ii) less buffer size

(iii) less complexity ⇒ $O(MN)$ time

6

(c) ∵在影像中，相鄰兩個 block ~~pixel~~ 的亮度 大部分是 difference=0
是差不多一樣的，∴我們可以只記錄 difference 的部分，這樣如果差異是0的話，我們只要記錄其中一個 original value 就好，可以省去資料記錄的空間 #

(d) ∵如果照 zig zag 順序的話，我們就可以從左邊 rows 和 columns 頻率較低的往頻率較高的去採計，而通常影像較高頻的部分是較少的，通常是較低頻，因此 zig zag 後段值幾乎都是0，∴可以直接用指標知道只要記錄到哪邊有值就好，可以省去要記錄的資料 #

7. (i) $\text{ceil}\left(N\frac{\text{entropy}}{\log 2}\right) \le \text{total coding length} \le \text{floor}\left(\frac{N\text{entropy}}{\log 2}+N\right)$

$N=5000$, $\text{entropy} = -\sum_{n=1}^{\infty} p(n)\log p(n)$, $p = \frac{e^{-0.97}\times 0.97^n}{n!}$

⇒ 經過計算，entropy = 0.9196

⇒ total coding length 在 66342 ~ 116341 之間 #

7(ii)

$$\text{ceil}\left(N*\frac{\text{entropy}}{\log k}\right) \le \text{total length} \le \text{floor}\left(N\frac{\text{entropy}}{\log k} + \log_k 2 + 1\right)$$

⇒ total coding length 在 66342 ~ 66344 之間

Extra: 人耳 20~200w Hz 是多少分貝的聲音??

Ans: