

Multimedia Systems and Applications Midterm (Question Bank)

第一題

1. Briefly explain, in your own words, the difference between “Hypermedia” and “Multimedia”.

Exercise 1.5.3

多媒體主要是一個媒體集合，由那些需要電腦或裝置協助呈現的媒體所組成，如圖片、音樂以及影片等等。可透過 Linear 或 Non-linear 的格式儲存。

但超媒體比較像是一種系統，透過連結的方式使得任何的媒體能夠呈現出來。最著名的例子就是 WWW，透過 HTML 等語言將圖片、音樂及影片呈現於使用者面前。因此，超媒體是以 Non-linear 的格式儲存。

第二題

2. What extra information is multimedia good at conveying?

(a) What can spoken text convey that written text cannot?

(b) When might written text be better than spoken text?

Exercise 2.7.1

(a)

速度、語調、情感以及態度等等的都無法透過文字呈現。

(b)

若以文字呈現，閱讀者可以透過非線性的方式閱讀。

除此之外，透過調整文字的粗細、顏色以及縮排等等，可以達到結構化或提示的目的。

另外，有些語句聽起來一樣，但以文字表現卻不同，這也是無法透過語句所辨識的。範例如下：

- I said “quickly, come here.”
- I said quickly, “come here.”

第三題

3. What information is embedded in (a) Speech (b) Image and (b) Video, respectively?

(a)

包含文字以及其說話的語調、情感、態度等。

(b)

包含二維的亮度、顏色資訊等，進而構成畫面。

(c)

包含了多段的 Speech、Image 以及時間等資訊，構成視訊。

第四題

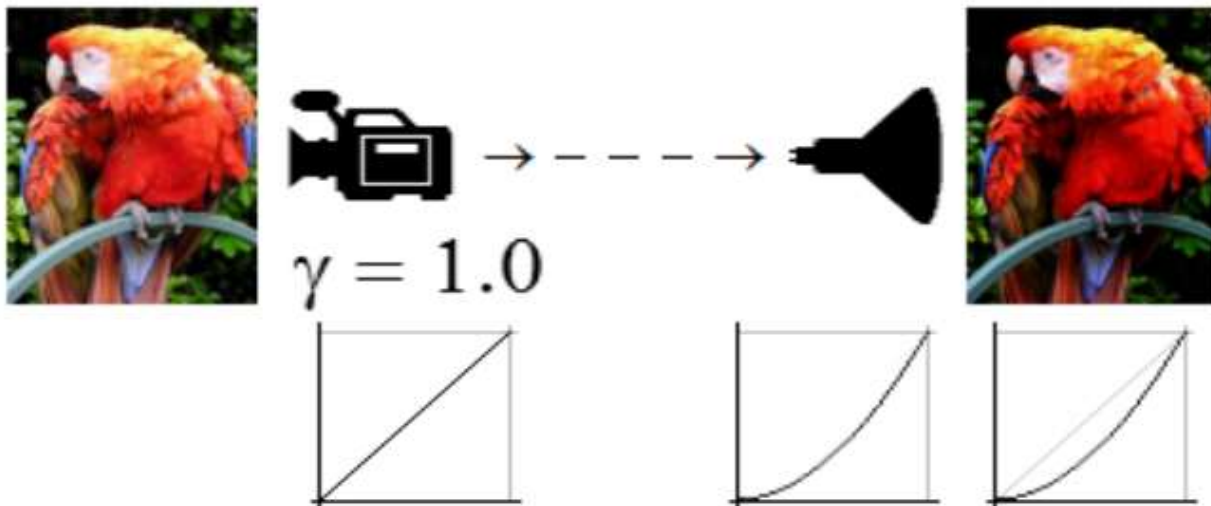
4. Please answer the following questions.

(a) What is gamma correction for the display in the CRT system?

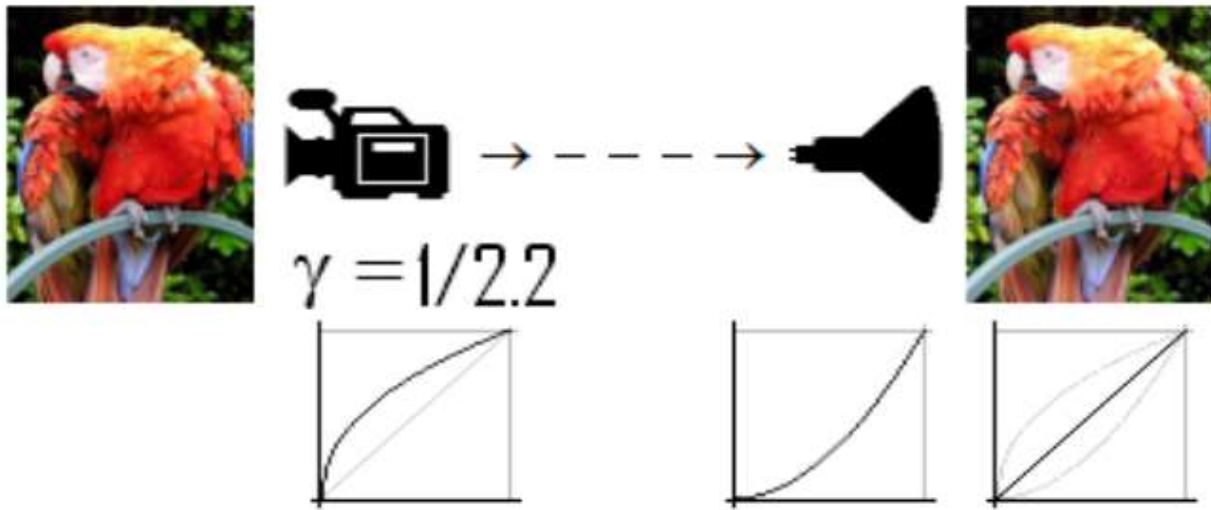
(b) If the color is out of gamut on a device, please provide one method to deal with this problem.

(a)

在 CRT 的系統之下，輸入單位電壓並不會產生等量的亮度，導致顏色亮度失真。



因此，影像在轉換儲存之時，會在亮度的次方部份乘上 $\frac{1}{\gamma}$ 來進行修正，這個過程即為 Gamma Correction。

**(b)**

最簡易的方法就是取支援色域內最接近該色彩的色彩。

建構一條射線，以 White Point 射向該色彩在色域圖中的位置，而射線與支援色域邊界所交的点即為所求。

第五題

5. (a) What is Weber's Law?
 (b) What is Nyquist Theorem?
 (c) Please describe "aliasing."

(a)

在同一類的刺激之下，其差異閾限與標準刺激強度成一定的比例關係，才能引起差別感覺。

意即 $\frac{\Delta (\text{差異閾限})}{I (\text{標準刺激強度})} = k$ (韋伯比例常數)。

(b)

若訊號為 band-limited，並且假設最低頻率為 f_1 ，最高為 f_2 ，則取樣頻率至少要 $2 f_2$ 才不會發生 aliasing。

(c)

因為取樣頻率未遵照 Nyquist Theorem 所述，至少要是兩倍的最高頻率，而導致取樣訊號被還原成連續訊號時產生彼此交疊而失真、無法辨認的現象。

且 $f_{alias} = f_{sample} - f_{true}$ for $f_{true} < f_{sample} < 2 f_{true}$ 。

第六題

6. (a) What is meant by the Multimedia and Hypermedia? Distinguish between these two concepts.
 (b) Why is file or data compression necessary for Multimedia activities, specially when transmitted on the Internet?
 (c) Please describe "aliasing."

(a)

參考第一題。

多媒體是包含了文字、圖像、音訊以及視訊等媒體的集合。

超媒體是透過連結關係，將多種媒體資訊封裝起來，以利於呈現或操作。

(b)

多媒體的資料量相較於其他類型都十分的大，因此為了節省儲存空間，會進行壓縮的動作。而因為國際網路速度的限制，更需要壓縮的技術，以減少傳輸時間。

(c)

同第五題的 (c)。

第七題

7. My old SoundBlaster card is an 8-bit card

- (a) What is the definition of Signal-to-Quantization-Noise (SQNR)
 (b) What is the best SQNR it can achieve?

Exercise 6.4.2

(a)

將訊號為一連續的電壓值，在量化時會發生所謂的進位誤差，而 SQNR 便是用以評估量化品質的指標。其公式如下：

$$SQNR = 20 \log_{10} \frac{V_{signal}}{V_{quan_noise}}$$

(b)

以 PSQNR 為答案。

$$SQNR = 20 \log_{10} \frac{2^{8-1}}{\frac{1}{2}} = 20 \log_{10} 2^8 = 160 \log_{10} 2 \approx 48.16 \approx 48$$

第八題

8. For a quantization accuracy of N bits per sample, what is the worst case of the SQNR?

將最大的訊號值映射至 $2^{N-1} - 1$ ，最小的映射至 -2^{N-1} 。

而最糟的情況就是「量化誤差剛好是區間值的一半」，則 SQNR 可以表示成如下公式：

$$SQNR = 20 \log_{10} \frac{2^{N-1}}{\frac{1}{2}} = 6.02 N$$

若輸入訊號為正弦波，因為量化誤差均勻分佈在 0 至一半間，所以 SQNR 公式則變為：

$$SQNR = 6.02 N + 1.76$$

第九題

9. Describe the process of the digitization of an analog sound. Define what a sample is, the frequency of occurrence and how these factors can affect the quality of sound.

類比音訊數位化的步驟如下：

1. Sampling
2. Quantization

所謂的 Sampling 就是間隔固定時間對類比訊號進行取樣，而此頻率就是取樣頻率。至於 Quantization 則是在取樣後，將類比訊號數值化以便儲存的過程。

而上述所說的都將會影響數位化後的音訊品質。

以 Sampling 來講，若取樣頻率未依照 Nyquist Theorem，則會造成 aliasing 的現象。

以 Quantization 來講，若數值化後的值與實際訊號差異過大，也會影響還原後的音訊品質。

第十題

10. (a) Dithering is often used when converting greyscale images to monochrome. What is the basic idea of dithering algorithm?

(b) For the given 2x2 dither matrix, briefly describe the ordered dithering algorithm.

$$\begin{pmatrix} 0 & 2 \\ 3 & 1 \end{pmatrix}$$

(c) Use the same dither matrix, what is the result for the following input? Assume that the input is greyscale intensities normalized to 0 to 1.

$$\begin{matrix} 0.66 & 0.18 & 0.03 & 0.19 \\ 0.54 & 0.13 & 0.56 & 0.37 \\ 0.70 & 0.99 & 0.88 & 0.46 \\ 0.67 & 0.17 & 0.67 & 0.98 \end{matrix}$$

(a)

以題幹所提之灰階圖為例，Dithering 會透過特定演算法決定該像素是否被印出。

該原理是利用人類視覺會混合相鄰像素色彩的特性，使得印出點密集之處感覺較黑，稀疏之處較白，而達到灰階的呈現效果。

常用於無法支援 multi-level 的印表機，透過 dithering 可以使其只用 2-level 色彩來呈現資訊。

(b)

以灰階圖為例，像素值介於 0 到 255。

1. 將 256 levels 分割為 5 個區間，每一區間的大小為 $\frac{256}{5} = 51.2$
使得 Dither Matrix 中對應的像素值如下：

$$\begin{bmatrix} 51.2 & 153.6 \\ 204.8 & 102.4 \end{bmatrix}$$

2. 在處理位置 (x, y) 的像素點時，對 (x, y) 除以 2 取餘數得 (i, j)
3. 若 $I(x, y) > D(i, j)$ 則印出，否則不印。

(c)

將 Dither Matrix 一樣標準化，使得範圍為 0 到 1。

$$\begin{bmatrix} 0.2 & 0.6 \\ 0.8 & 0.4 \end{bmatrix}$$

再套用 ordered dithering

$$\begin{bmatrix} 0.66 > 0.2 & 0.18 < 0.6 & 0.03 < 0.2 & 0.19 < 0.6 \\ 0.54 < 0.8 & 0.13 < 0.4 & 0.56 < 0.8 & 0.37 < 0.4 \\ 0.70 > 0.2 & 0.99 > 0.6 & 0.88 > 0.2 & 0.46 < 0.6 \\ 0.67 < 0.8 & 0.17 < 0.4 & 0.67 < 0.8 & 0.98 > 0.4 \end{bmatrix}$$

可得輸出矩陣如下：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

第十一題

11. Instead of using 8 bits per pixel, you'd like to use 48 bits per pixel in RGB. How could you store the original grayscale images so that in the new format they would appear the same as they used to, visually?

Exercise 3.3.5

使用 48-bits 來代表 RGB，每個 channel 使用 16-bits，故範圍為 2^{16} 。

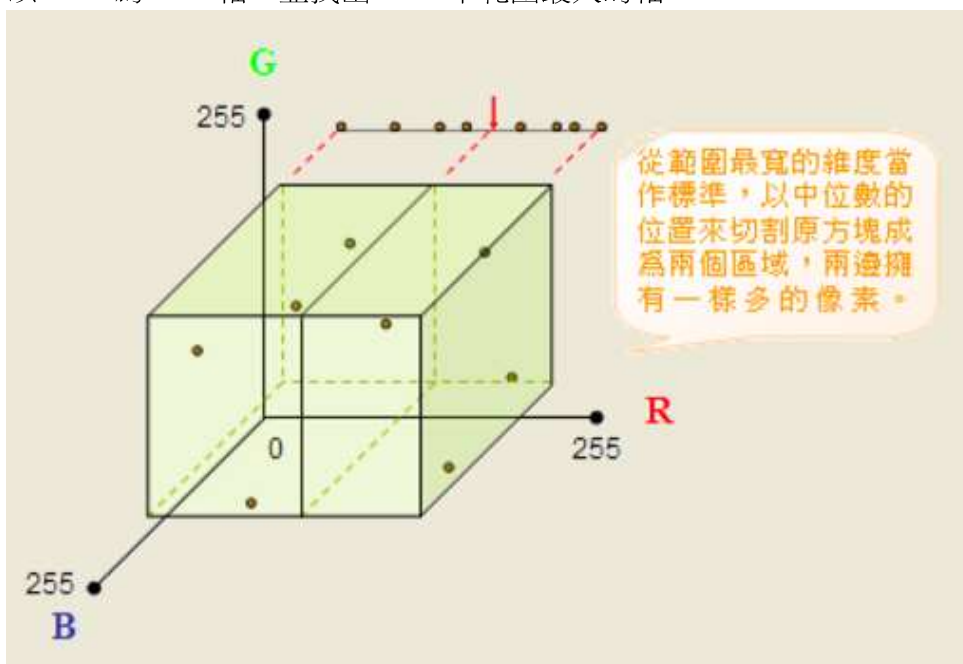
灰階圖原先的像素值範圍為 2^8 ，若要使得圖像看起來相同，則必須將原本的像素值乘以 2^8 ，使其落在 2^{16} 範圍，最後再將該值同時存入 RGB 三個 channel 即可。

第十二題

12. Please describe the steps to devise a Color Lookup Table to make 8-bit lookup color out of 24-bit color. You can choose your own way or median-cut algorithm.

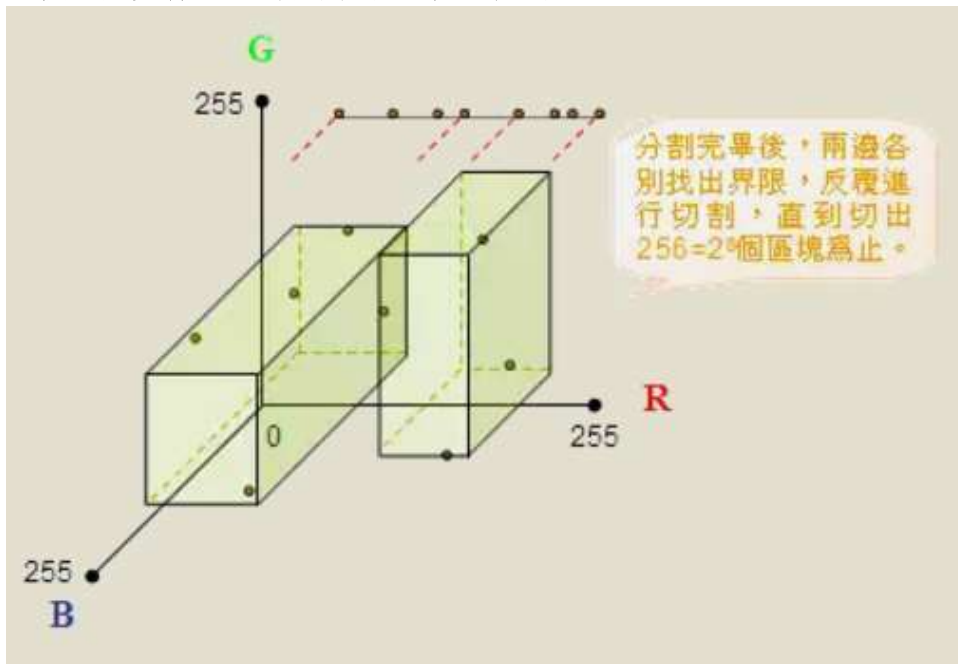
以 Median Cut 演算法為例。

1. 以 RGB 為 XYZ 軸，並找出 RGB 中範圍最大的軸。

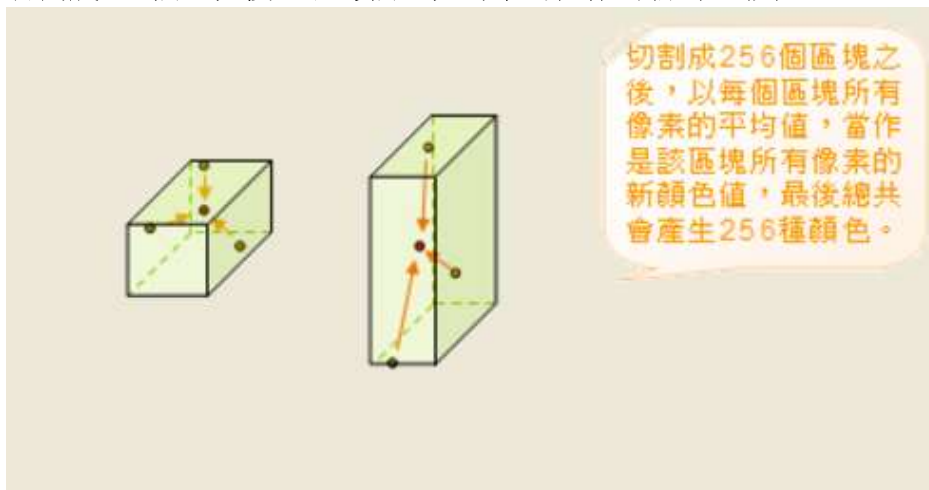


2. 以其軸的中位數做切割，使得兩邊像素數量一樣。

3. 重複上述步驟，直到切割出 2^8 個區塊為止。



4. 切割成 2^8 個區塊後，以每個區塊的平均值作為新的色彩值。



第十三題

13. Please answer the following questions.

- (c) What is gamma correction for the display in the CRT system?
- (d) If the color is out of gamut on a device, please provide one method to deal with this problem.

同第四題。

第十四題

14. Suppose we use a new set of color-matching functions $x^{\text{new}}(\lambda)$, $y^{\text{new}}(\lambda)$, $z^{\text{new}}(\lambda)$ with values

λ (nm)	$x^{\text{new}}(\lambda)$	$y^{\text{new}}(\lambda)$	$z^{\text{new}}(\lambda)$
450	0.2	0.1	0.5
500	0.1	0.4	0.3
600	0.1	0.4	0.2
700	0.6	0.1	0.0

In this system, what are the chromaticity values (x, y) of equi-energy white light $E(\lambda)$ where $E(\lambda)=1$ for all wavelengths λ ? Explain.

Exercise 4.4.6

Chromaticity values 是由下方公式計算而得：

$$(X, Y, Z) = \left(\int E(\lambda) \bar{x}(\lambda) d\lambda, \int E(\lambda) \bar{y}(\lambda) d\lambda, \int E(\lambda) \bar{z}(\lambda) d\lambda \right)$$

而題幹已經假設 $E(\lambda) = 1$ ，因此根據所給予的表格，可知：

$$(X, Y, Z) = \left(\int \bar{x}(\lambda) d\lambda, \int \bar{y}(\lambda) d\lambda, \int \bar{z}(\lambda) d\lambda \right) = (1, 1, 1)$$

$$\text{因此 } (x, y) = \left(\frac{X}{X+Y+Z}, \frac{Y}{X+Y+Z} \right) = \left(\frac{1}{3}, \frac{1}{3} \right)$$

第十五題

15. Digital video uses chroma subsampling. What is the purpose of this? Why is it feasible?

Exercise 5.5.7

由於人類視覺對於色彩的敏感度不及黑白 (亮度)，因此刪除部份色彩資訊是可以被接受的。日常中最常見的應用就是 TV 系統，在 TV 系統上的亮度資訊頻寬通常都比色彩頻寬還要來的大。

而此處所提的 Chroma Subsampling 也是使用了相同的原理，保存亮度資訊、減少色彩資訊，藉此便可減少所要儲存或傳輸的資料量。

第十六題

16. If a set of ear protectors reduces the noise level by 30 dB, how much do they reduce the intensity (the power)?

Exercise 6.4.4

每減少 10 dB 代表能量減少 10 倍。

因此總共減少 1000 倍。

第十七題

17. Suppose the sampling frequency is 1.5 times the true frequency. What is the alias frequency?

Exercise 6.4.6

由於 $f_{alias} = f_{sample} - f_{true}$ for $f_{true} < f_{sample} < 2 f_{true}$

故 $f_{alias} = 0.5 f_{true}$

第十八題

18. Suppose we use a predictor as follows:

$$\hat{f}_n = \text{trunc}\left[\frac{1}{2}(\tilde{f}_{n-1} + \tilde{f}_{n-2})\right]$$

$$e_n = f_n - \hat{f}_n$$

Also, suppose we adopt the quantizer

$$\tilde{e}_n = Q[e_n] = 16 * \text{trunc}[(255 + e_n) / 16] - 256 + 8$$

$$\tilde{f}_n = \hat{f}_n + \tilde{e}_n$$

If the input signal has values as follows:

20 38 56 74 92 110 128

What is the output from a DPCM coder (without entropy coding)?

Exercise 6.4.22

假設 $\tilde{f}_0 = \hat{f}_1 = f_1 = 20$, $\tilde{e}_1 = 0$, 則 $\tilde{f}_1 = 20$ 。

運算過程如下：

Notation	20	38	56	74	92	110	128
\hat{f}	20	20	32	50	65	81	97
e	0	18	24	24	27	29	31
\tilde{e}	0	24	24	24	24	24	24
\tilde{f}	20	44	56	74	89	105	121

第十九題

19. Suppose we use a predictor and the uniform delta modulation (DM) as follows:

$$\hat{f}_n = \tilde{f}_{n-1}$$

$$e_n = f_n - \hat{f}_n$$

Also, suppose we adopt the quantizer as

$$\tilde{e}_n = \begin{cases} 10 & \text{if } e_n > 0 \\ -10 & \text{otherwise} \end{cases}$$

$$\tilde{f}_n = \hat{f}_n + \tilde{e}_n$$

If the input signal has values as follows:

25 34 39 54 62 70 88

What is the output from a DM coder?

假設 $\tilde{f}_0 = \hat{f}_1 = f_1 = 25$ ， $\tilde{e}_1 = 0$ ，則 $\tilde{f}_1 = 25$ 。

運算過程如下：

Notation	25	34	39	54	62	70	88
\hat{f}	25	25	35	45	55	65	75
e	0	9	4	9	7	5	13
\tilde{e}	0	10	10	10	10	10	10
\tilde{f}	25	35	45	55	65	75	85