Image Processing / NCTU Fall 2020 Exam #2 (12/25/2020) Name: 張銘仁 ID: 309553045

Prob#	1	2	3	4	5	6	7	8	Total
Score	20	13	10	9	9	5	12	6	84

1. [24%] Answer the following questions:

(a) What is the purpose of subband coding? How is it used in discrete wavelet transform of an image?

(b) Consider the periodic sequence "1 3 2 1 3 2 1 3 2 1 3 2…" down-sampled by a factor of 2. Use this as an example to explain the cause of "aliasing". How is this related to the sampling theorem?

(c) For JPEG coding, the quantization matrix ($Z(u,v)$) is a pre-specified normalization matrix multiplied by a positive constant. How is this constant related to the compression ratio and the quality of the reconstructed image?

(d) Explain what the "impulse response" of a frequency-domain filter is, and use it to explain the "ringing" caused by ideal LPF in images.

(e) The "degradation function" of an image is given as $g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$. Here f and g represent the images before and after degradation, respectively. Explain what h and η represent.

(f) Explain the source of wrap-around error in discrete Fourier transform, and the standard procedure to prevent it.

2. [16%] These are questions related to the group presentations. Each is worth 2 points, up to a total of 16 points. Clearly indicate the questions you're answering.

(a) Give two applications of image inpainting.

(b) The active contour energy function involve internal and external energy terms. Which is intended to make the contour "fit to image edges"?

(c) What is the meaning of "saliency" in images? human 關注

(d) Name a type of convolutional neural networks used in image compression.

(e) Pairs of image patches are used to train models for learning based super resolution. What is the relation between the two images in a pair?

(f) Give two methods/techniques of obtaining depth information of images.

(g) Describe the least-significant-bits method of image watermarking.

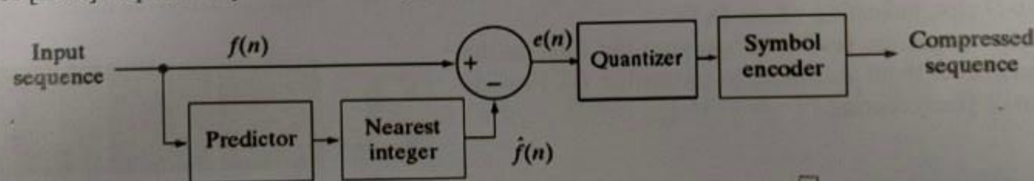
(h) One denoising method using CNN aims to learn a residual function. What does this function try to model?

(i) What information is used when grouping pixels into superpixels with the SLIC algorithm?

(j) What does the "blending" step in image stitching do?

(k) What differences are in the multiple images used to produce a HDR image?

3. [10%] Explain why this block diagram for a lossy predictive coder is incorrect, and how you can fix it.



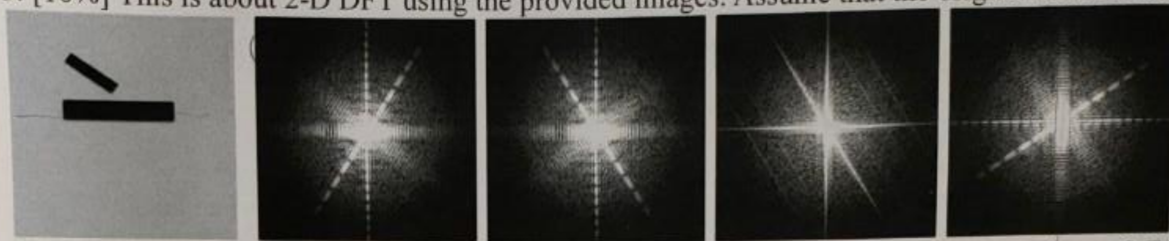
4. [10%] This problem is about arithmetic coding. Given the intervals for the four symbols a_1, a_2, a_3, a_4,

(a) Give the 3-element string obtained by decoding 0.25.

(b) Find the shortest binary number needed to encode the 2-element string a_2a_3.

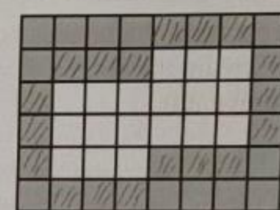
Source Symbol	Probability	Initial Subinterval
a_1	0.2	[0.0, 0.2)
a_2	0.2	[0.2, 0.4)
a_3	0.4	[0.4, 0.8)
a_4	0.2	[0.8, 1.0)

5. [10%] This is about 2-D DFT using the provided images. Assume that the original size is 100x100.



(a) Among the right four images, which is the likely Fourier transform of the leftmost image? Explain.
(b) The transform images have the brightest part near the center. How do we achieve this centering?
(c) What happens if the two rectangles in the original image are rotated clockwise by 10 degree?
(d) What happens if the two rectangles are moved downward by 30 pixels?
(e) What happens if the original image is resampled to 200x200 while keeping the appearance unchanged?

6. [10%]



(a) In the binary image, the white pixels are the foreground pixels. Using a structuring element consisting of the center pixel and its 4 neighbors, draw the foreground pixels after the morphological operation given below. Here set A contains the foreground pixels and set B is the structuring element. The operation $\oplus$ is dilation.

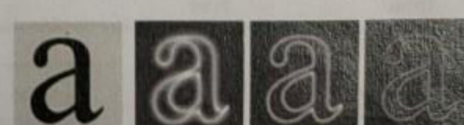
$$(A \oplus B) - A$$

(b) Give the two structuring elements needed for using hit-or-miss transform to detect the shape represented by B in (a).

7. [12%] For the following sequence of pixel values: **1 2 3 4 5 6 7 8 7 6 5 4 3 2 1 0**
(a) Compute the entropy of the sequence itself.
(b) Create the "difference sequence" by replacing every pixel value (except for the first one) with its difference from the previous pixel value.
(c) Compute the entropy of the difference sequence.
(d) Use the above results to explain the benefit of predictive coding in image compression.
The equation for entropy is given by $H = -\sum_f p(f)\log_2[p(f)]$. Here $p(f)$ is its probability of a symbol f.

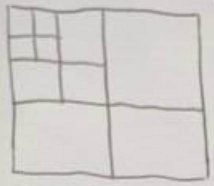
8. [8%] The transfer function of a filter is given by $H(u,v) = \dfrac{[D(u,v)]^2}{D_0^{\,2} + [D(u,v)]^2}$.

(a) Draw its H-vs-D plot, indicating where D_0 is.
(b) The three right images are results of applying the filter to the leftmost image. Order them according to increasing D_0. Provide an explanation of your ordering.



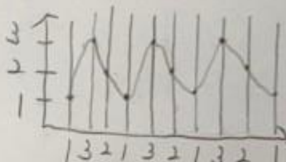
學號: 309553045

(a) subband coding 為了降低需儲存之資訊，原來在DWT需存下 1/4之原圖，以及其他相鄰 pixel 之差異量，用 subband coding 對縮小之圖片做同樣事情

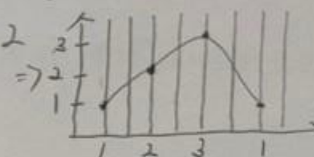


-1

(b) 原先之 sequence



若 downsample by 2 =>



由於 downsample 導致 sequence 改變，形成不同之頻率 稱為 aliasing。

-1

(c) 在 JPEG 當中，經過 DCT 之結果會由 $Z(u,v)$ quantization，$Z(u,v)$ 值愈大，量化後值愈小，所需 bit 愈少，但品質愈差，因為量化會將區間內值壓成一樣，區間愈大壓縮值愈多，而乘上之 positive constant 則可控制此量化區間 ($Z(u,v)$) positive constant 大 compression ratio 上升 quality 下降

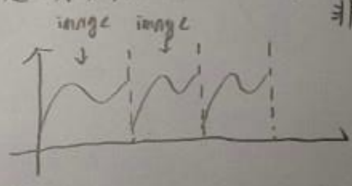
(d) impulse response 為以 delta function 當作 filter 在 frequency-domain，用以得到一頻率之值，ideal LPF 可想成多個在低頻之 delta function，由於 delta function 轉回 spatial domain 為 constant function ?

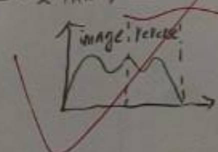
-2

(e) $h(x,y)$ 為 degradation function (通常為一特定之 filter)

$n(x,y)$ 為 noise function

(f) 由於圖片在做 DFT時，會將圖片當作一段波峰到波谷，因此在做 DFT之前會想成無限 repeat之 image 拼成之訊號，但圖片左右，上下不會正好能相連，導致訊號有非連續點，因此~~將圖片反轉拼貼~~，使其形成連續訊號





2.

(a) 還原因傳輸而損失之 pixel，去除圖片中之特定物體 +2

(b)

(c) human 關註之圖片區域 +2

(d) SRCNN，SRFlow

(e) 一張為 HR (High resolution) image，一張為 LR (Low resolution image)，兩張為相同圖片僅解析度不同 +2

(f) 1. 利用 Lidar，kinect，去得到 depth information.

2. 利用 CNN 預測單張影像深度 ✓ +2

(g)

(h) inverse noising function. ✗ +1

(i)

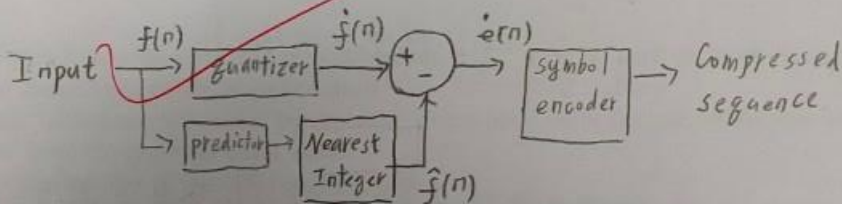
(j) 在 image stitching 時，直接將 image 拼貼會有不連續感，用 image blending 將 image 拼起來以消除此情形。 blending $= \alpha A + (1-\alpha) B$，A, B 為二圖 ✓ +2

(k) 為不同曝光時間之照片，以此來重建其光場，形成 HDR image ✓ +2

3.

由於 $e(n)$ 為 2 相鄰 bit 之差異，以此做 quantize，quantizer 所造成之 loss 會影響到還原時的預測，使其愈來愈不準。

正確為



4.

$0.0 \overset{a_1}{—} 0.2 \overset{a_2}{—} 0.4 \overset{a_3}{—} 0.8 \overset{a_4}{—} 1.0 \rightarrow a_2$ (0.25)

$0.2 — 0.24 — 0.28 — 0.36 — 0.4 \rightarrow a_2$ (0.25)

$0.24 — 0.248 — 0.256 — 0.272 — 0.28 \rightarrow a_2$ (0.25)

(a) $a_2 a_2 a_2$ ✓

$\underline{a_2 a_3} \rightarrow 0.28 \sim 0.36 \rightarrow$ 最短 binary number 為 $0.3125 = \underline{\frac{1}{4} + \frac{1}{16}}$

$0.3125_{10} = 0.\cancel{00}101$

0.0101 −1

張銘仁 309553045

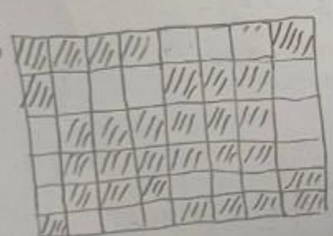
(a) 第一張，若以此為例  由於此圖片在x軸訊號為 
比y軸之訊號來之寬，因此需更多高頻資訊，所以x方向之間距較小，由此推出 ①

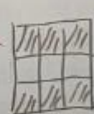
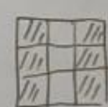
(b) 乘上 $f(x,y) = -1^{(x+y)}$

(c) 在 frequency-domain 也會跟著旋轉 10度

(d) 在 frequency-domain 不會改變

(e) 在 frequency-domain 會 rescale 2倍. -1

6.
(a) 

(b)  and 

7.

0	1	2	3	4	5	6	7	8	
1	2	2	2	2	2	2	2	1	total = 16

(a) $-\sum_{i=0}^{8} P(i) \log P(i) = -\left(\frac{1}{16} \cdot \log \frac{1}{16} \times 2 + \frac{2}{16} \cdot \log \frac{2}{16} \times 7\right)$

$= \frac{4}{16} \times 2 + \frac{6}{16} \times 7 = \frac{50}{16} = \frac{25}{8}$ #

(-4, -3)

(b)
1 2 3 4 5 6 7 8 7 6 5 4 3 2 1 0
→ 0 1 1 1 1 1 1 1 -1 -1 -1 -1 -1 -1 -1 -1

(c)

-1	0	1
8	1	7

$-\sum_{i=-1}^{1} P(i) \log P(i) = -\left(\frac{8}{16} \cdot \log \frac{8}{16} + \frac{1}{16} \log \frac{1}{16} + \frac{7}{16} \log \frac{7}{16}\right)$

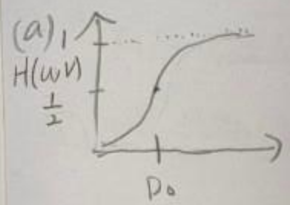
$= -\left(-\frac{1}{2} - \frac{1}{4} + \frac{7}{16} \log \frac{7}{16}\right)$

-1

(d) 由於相鄰 bit 通常有許多相同之訊息，用 predictive code 減少訊息量以利於壓縮

8. $H(u,v) = \dfrac{1}{1+\left[\dfrac{D_0}{D(u,v)}\right]^2}$

(a)



$\rightarrow$ high-pass

(b) D_0愈大，能通過之波就愈高頻，