Image Processing / NCTU Fall 2019 Exam #2 (1/10/2020)   Name: 陳彥鈞   ID: 0856134

Prob#	1	2	3	4	5	6	Total
Score	14	13	8	15	2	9	61

1. [20%] Answer the following questions.

(a) What is the purpose of adaptive thresholding?

(b) Give an example of a 3x3 spatial filter that has the effect of a HPF, and another 3x3 spatial filter that has the effect of a LPF.

(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 filter is.

(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.

(a) 因為當Image的pixel value為漸增或漸減時,造成光線漸明或漸亮的效果

則若使用global thresholding不太適合,可能導致Image變成 

若以adaptive thresholding對每個區塊選一合適的threshold值,來避免上述之結果,

⇒ 因此,使用adaptive thresholding 而不是只使用一global thresholding value

(b) HPF = Gaussian

LPF = Gaussian

$\frac{C_{current}}{C_{previous}} = C_k$

$1 - \frac{1}{C_k} = ratio$

(c) 這個常數越大,則壓縮率越大,但重建效果越差

反之,常數越小,壓縮率越小,但重建效果越好

(d) 突然之間有一個有効訊號,導致結果圖看起來像是水波紋

(e) h: degradation function

η = noise function

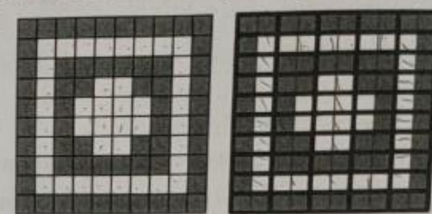
−12

2. [25%] The objective here is to compress the binary image to the right (left plot; white=1, gray=0).

(a) If we treat "the value of a pixel" as one symbol, give the probabilities of the two symbols, 0 and 1. How many bits per pixel are required using Huffman coding?

(b) Consider the case when "the values of a horizontally adjacent pair of pixels" are treated as one symbol. Note that there are now four possible symbol values: "0,0", "0,1", "1,0", and "1,1". The pixels are arranged into pairs as illustrated to the right (right plot). Give the probabilities of the four symbols. How many bits per pixel are required using Huffman coding?

(c) In predictive coding, the difference from a reference value, instead of the value itself, is coded. Now let us apply predictive coding to the symbol values in (b). Specifically, the reference of a pixel pair is the pixel pair right above it. The symbol set representing the differences has four possible values:



d00: both pixels in the pair are the same as the reference
d01: only the second pixel in the pair is different from the reference
d10: only the first pixel in the pair is different from the reference
d11: both pixels in the pair are different from the reference

Give the probabilities of the four symbols. For pixel pairs in the top scanline, use "0,0" as the reference. How many bits per pixel are required using Huffman coding?

(a) symbols probability 100×

0 : 0.6

1 : 0.4

⇒ bits per pixel = 1

50

1 : 6

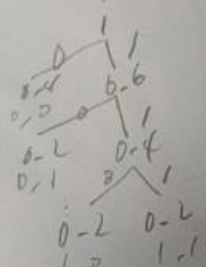
0 : 4

0.4

1

(b) symbols probability

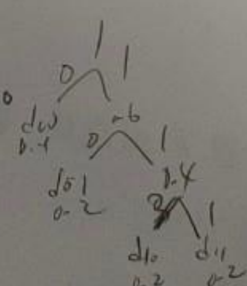
symbols	probability	
0,0	0.4	20
0,1	0.2	10
1,0	0.2	10
1,1	0.2	10



⇒ 0,0 ⇒ 0
0-1 ⇒ 10
1,0 ⇒ 110
1,1 ⇒ 111

bits per pixel : 0.4 + 0.4 + 0.6 + 0.6
= 2/2 − 1

ref : 0,0



(c) symbols probability

symbols	probability
d00	~~0.4~~
d01	~~0.2~~
d10	~~0.2~~
d11	0.2

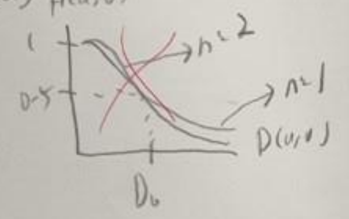
⇒ d00 ⇒ 0
d01 ⇒ 10
d10 ⇒ 110
d11 ⇒ 111

⇒ bits per pixel = 0.4 + 0.4 + 0.6 + 0.6
= 2

3. [15%] Consider this transfer function: $H(u,v)=\dfrac{[D(u,v)]^{2n}}{D_0^{2n}+[D(u,v)]^{2n}}$

(a) Plot the approximate curves for n=1 and n=2.
(b) Is this a LPF or HPF? Briefly explain.
(c) Which one in (a) is more similar to an ideal LPF or HPF (depending on your answer in (a) and (b))?
(d) Which one in (a) is more likely to cause ringing along edges in the filtered image?

(a) $H(u,v)$



(b) ~~LPF~~

n 越大 越接近 ideal

(c) n=2

(d) n=2

4. [15%] The purpose is to do block-based transform coding. The following is a 2x2 gray-level image block (left) and four "basis blocks".

20	32
28	40

1/2 x

1	1
1	1

1/2 x

1	1
-1	-1

1/2 x

1	-1
1	-1

1/2 x

1	-1
-1	1

(a) Obtain the four coefficients by projecting the image block to the four basis blocks.
(b) Reconstruct the image block by dropping the last coefficient (i.e., using only the first three basis blocks and their respective coefficients.)
(c) Reconstruct the image block using only the DC coefficient.
(d) For the results of (b) and (c), compute the sum of squared errors with the formula below. Which is closer to the original?

$$e=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x,y)-f(x,y)\right]^2 \quad (M \text{ and } N\text{: size of the block})$$

(a)

$$\left[\begin{array}{cccc|c} 1&1&1&1&20\\ 1&1&-1&-1&32\\ 1&-1&1&-1&28\\ 1&-1&-1&1&40 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1&1&1&1&20\\ 0&0&-2&-2&12\\ 0&-2&0&-2&8\\ 0&-2&-2&0&20 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1&1&1&1&20\\ 0&0&1&1&-6\\ 0&1&0&1&-4\\ 0&1&1&0&-10 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccc|c} 1&1&1&1&20\\ 0&0&1&1&-6\\ 0&1&0&1&-4\\ 0&0&1&-1&-6 \end{array}\right]$$

$\Rightarrow \begin{cases} a=30\\ b=-4\\ c=-6\\ d=0 \end{cases} \Rightarrow \begin{cases} a=60\\ b=-8\\ c=-12\\ d=0 \end{cases}$

For first block, coefficient is 60
second block, coefficient is -8
third block, coefficient is -12
forth block, coefficient is 0

(b)

20	32
28	40

(c)

30	30
30	30

(d) $e_{(b)}=0$

$e_{(c)}=100+4+4+100=208$

$\Rightarrow$ (b) is closer

5. [10%] Apply the basic global algorithm (given to the right) to find the threshold for an 8-level gray-level image with the histogram given below. The initial threshold is the mean intensity. Use a floating-point number to represent your threshold here. Stop when the threshold does not change between two iterations.

> REPEAT
> compute the two means of values above the below the threshold, respectively
> Use the midpoint of the two means as the new threshold.
> UNTIL convergence

gray-level value	0	1	2	3	4	5	6	7
percentage	30	30	10	0	10	20	0	0

mean: 0.3 + 0.2 + 0.4 + 1 = 1.9 (T)

-8

$m_{below T}$ = 0.5

$m_{above T}$ = 0.2 + 0.4 + 1 = 4 ⇒ new threshold = 0.95

$m_{below T}$ = 0

$m_{above T}$ = 0.3 + 0.2 + 0.4 + 1 = 1.9 ⇒ threshold = 0.95 (not change)

⇒ Threshold = ~~0.95~~

6. [15%] The following are questions from the group presentations. One question is given for each topic. You should only answer five questions in the space given below. No more than two sentences required for each question.

(1) What are the "edge weights" in min-cut based segmentation?
(2) Describe an application of image inpainting.
(3) Watermarking for copyright protection requires the property of <u>sturdiness</u>. What does it mean?
(4) What are the purposes of the two main steps in image stitching: <u>alignment</u> and <u>blending</u>?
(5) What metric is used to compare denoising methods according to the presentation?
(6) Give at least two parameters used to specify a Gaber filter.
(7) What is a saliency map of a layer of convolutional neural network?
(8) Give the two types of information used to compute distances between pixels and cluster centers used in SLIC superpixel algorithm.
(9) Define the task of super-resolution, and explain why direct interpolation is not sufficient.
(10) What is the meaning of dynamic range?
(11) Describe the image capture process in computed tomography.
(12) For "full reference" image quality assessment, what is used as the reference?

	no.	your answer
-3	9	把低解析度的圖變成高解析度，且它只做插值是不足夠的，因為通常低解析度的會有些 noise，還需要 denoise
-2	10	pixel value 的範圍
	2	去除圖片上的人、事物，只留背景
	3	浮水印不會因為一些破壞而無法正確辨識
-1	4	對齊是指將2個不同視角的 image 經由重疊的對應；blending：指利用對齊後的 image 做延伸，以達到 stitching 效果