Polytechnic Institute of NYU
EL5123/BE6223 --- DIGITAL IMAGE PROCESSING

Fall 2012
Yao Wang

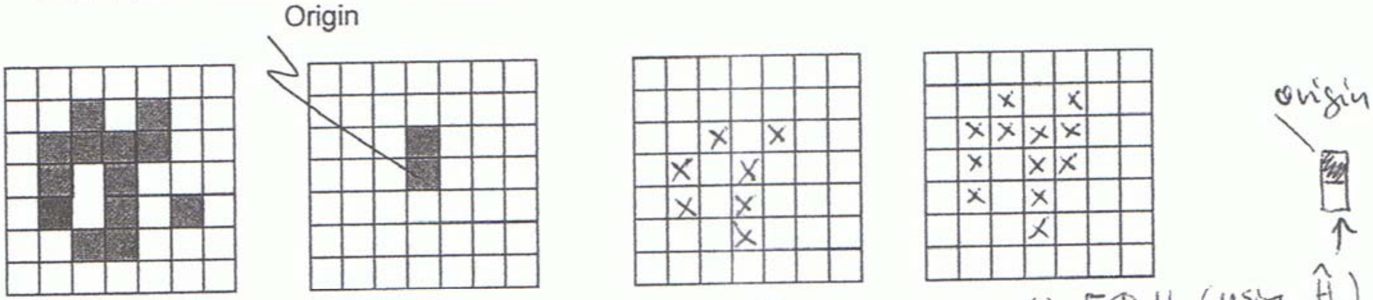
Final Exam (12/19/12, 3:00-5:30PM)

Closed book, 1 sheet of notes (double sided) allowed.

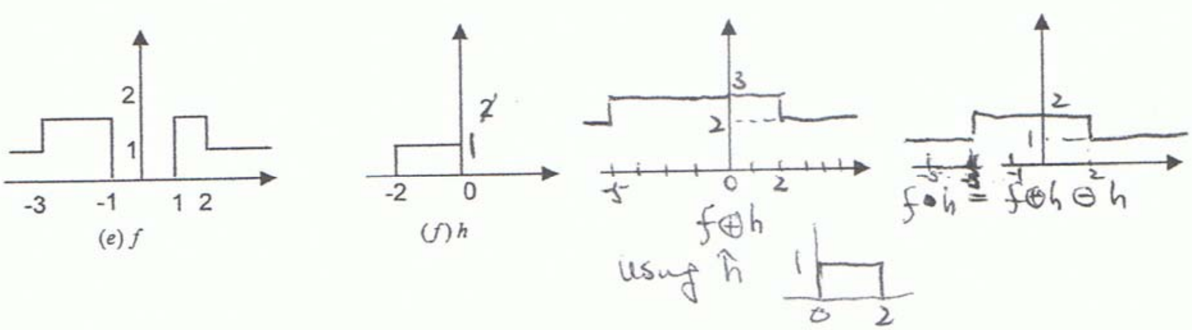
No peeking into neighbors or unauthorized notes. Cheating will result in getting an F on the course. Write your answers on this problem sheet for problems where space is provided. Otherwise write you answer in the blue book. Submit both the blue book and this problem set.

Your name: Solution

1. (10pt) (i) (4pt) Find the opening of the binary image, F, in Fig.(a) by the structuring element H in Fig.(b). You can use the grids in Fig.(c) and Fig.(d) to draw the intermediate and the final results.
(ii) (4pt) Find the gray scale closing of the function sketched in Fig.(e) with the structural element in Fig.(f). Sketch the intermediate and final result in Fig.(g) and Fig.(h).
(iii) (2pt) comment on the effect of opening and closing.



(a) F (b) H (c) F⊖H (c) F⊕H (using $\hat{H}$)



(e) f (f) h $f \oplus h$ using $\hat{h}$ $f \circ h = f \oplus h \ominus h$

— binary opening removes points that do not have vertically adjacent neighbors

— gray scale closing closes up gaps of width ≤ 2.

2. a) $f(x,y)=\sin(4\pi x+2\pi y)$

$F(u,v)=\frac{1}{2j}[\delta(u-2,v-1)-\delta(u+2,v+1)]$

magnitude of $\delta(u-2,v-1)$ is $|\frac{1}{2j}|=\frac{1}{2}$, $\delta(u+2,v+1)$ is $|\frac{-1}{2j}|=\frac{1}{2}$

b) $\Delta x=\Delta y=\Delta=1/3$. $f_{sx}=f_{sy}=3$

Seen in Figure (b)

c) Frequency domain, after multiply by $H_1(u,v)$, the reconstructed signal is

$\hat{F}(u,v)=-\frac{1}{2j}\delta(u-1,v+1)+\frac{1}{2j}\delta(u+1,v-1)$. So in spatial domain $\hat{f}(x,y)=f_{r1}(x,y)=-\sin(2\pi x-2\pi y)$

d) $h(x)=\begin{cases}1-\frac{|x|}{1/3} & |x|<\frac{1}{3}\\ 0 & \text{others}\end{cases}$ $H(u)=\left(\frac{\sin\pi\frac{1}{3}u}{\pi u}\right)^2$

$h_2(x,y)=h(x)h(y)$, so $H_2(u,v)=H(u)H(v)=\left(\frac{\sin\frac{\pi}{3}u}{\pi u}\cdot\frac{\sin\frac{\pi}{3}v}{\pi v}\right)^2$

If $H_2(u,v)$ is bounded by $|u|<\frac{7}{2}$ and $|v|<\frac{7}{2}$, So $\hat{F}(u,v)$ can be seen as

$$\hat{F}(u,v)=\frac{H(2,1)}{2j}[\delta(u-2,v-1)-\delta(u+2,v+1)]+\frac{H(2,2)}{2j}[\delta(u-2,v+2)-\delta(u+2,v-2)]$$
$$-\frac{H(1,2)}{2j}[\delta(u-1,v-2)-\delta(u+1,v+2)]-\frac{H(1,1)}{2j}[\delta(u-1,v+1)-\delta(u+1,v-1)]$$

$$\hat{f}(x,y)=f_{r2}(x,y)=H(2,1)\sin[2\pi(2x+y)]+H(2,2)\sin[2\pi(2x-2y)]$$
$$-H(1,2)\sin[2\pi(x+2y)]-H(1,1)\sin[2\pi(x-y)]$$

As $H(u,v)$ is even in x, y axis, so $H(2,1)=H(-2,-1)$, and so on.

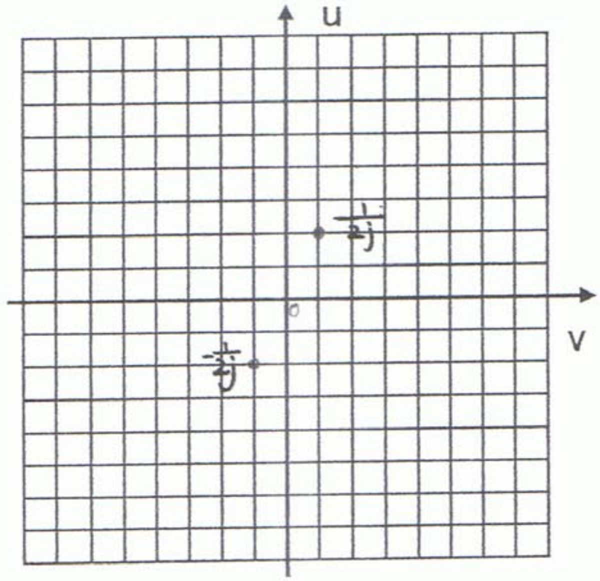


Fig. (a): Original Signal

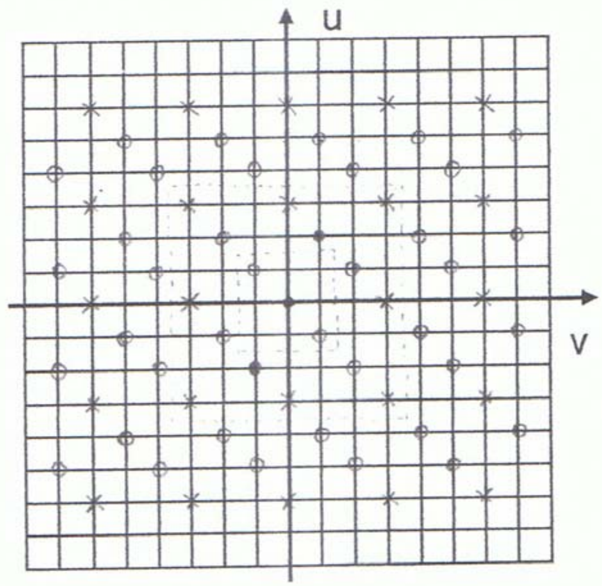


Fig.(b): Sampled Signal

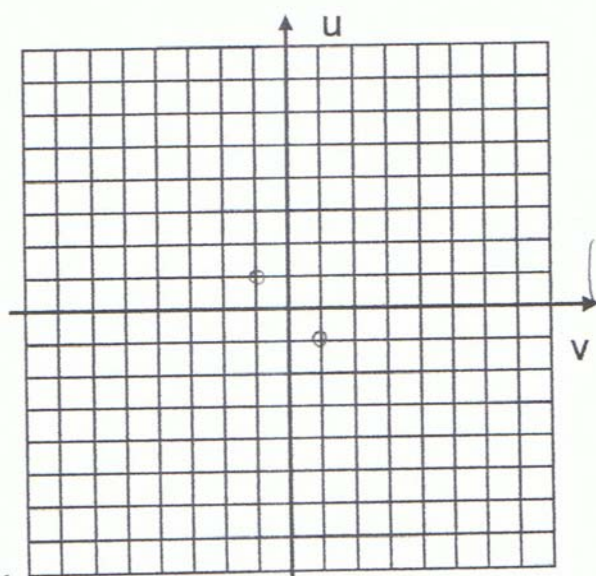


Fig. (c): Reconstructed Signal

3. a) EOL

Line 8
EOL
2, 1, 3, 1, EOL
2, 1, 3, 1, EOL
2, 5, EOL
2, 1, 3, 1, EOL
2, 1, 3, 1, EOL
2, 1, 3, 1, EOL

EOL

b)

Symbols	probability	Huffman codes
EOL	9/31	11
1	10/31	10
2	6/31	01
3	5/31	001
5	1/31	000

'1' '0' $\frac{19}{31}$ '1' '0' '1' '0' '1' '0' $\frac{6}{31}$ $\frac{12}{31}$ 1

c) Huffman coding average Length $l = \sum_{i=1}^{5} P_i \times L_i = \frac{9\times 2}{31} + \frac{10\times 2}{31} + \frac{6\times 2}{31} + \frac{5\times 3}{31} + \frac{1\times 3}{31} = \frac{68}{31} \approx 2.1935$ bits/Symbol

Entropy is defined as $H = -\sum_{i=1}^{5} P_i \log_2 P_i \approx 2.0875$

So $H < L < H+1$, which satisfy the shannon coding theorem.

d) $\text{bits/pixel} = \frac{\text{total \# of bits}}{\text{total \# of pixels}} = \frac{2.19 * 31}{81} = 0.838$

It is more compressed.

Compression ratio $= 1/0.838$

4.

	c ↓	
	D	C ← b
a →	B	A

$$\begin{bmatrix} R(1,1) & R(2,1) & R(3,1) \\ R(1,2) & R(2,2) & R(3,2) \\ R(1,3) & R(2,3) & R(3,3) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} R(0,1) \\ R(0,2) \\ R(0,3) \end{bmatrix}$$

$\hat{A} = aB + b\cdot C + c\cdot D$

"0" "1" "2" "3" So

$$\begin{bmatrix} \sigma^2 & \rho_d\sigma^2 & \rho_v\sigma^2 \\ \rho_d\sigma^2 & \sigma^2 & \rho_h\sigma^2 \\ \rho_v\sigma^2 & \rho_h\sigma^2 & \sigma^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \rho_h\sigma^2 \\ \rho_v\sigma^2 \\ \rho_d\sigma^2 \end{bmatrix} \quad a)$$

b) if $\rho_h = \rho_v = \rho_d = \rho$. then a) has to be $\begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \rho \\ \rho \\ \rho \end{bmatrix}$ $\quad \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{\rho}{2\rho+1} \\ \frac{\rho}{2\rho+1} \\ \frac{\rho}{2\rho+1} \end{bmatrix} = \frac{\rho}{2\rho+1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

c) $\sigma_p^2 = E\{|f_0 - \hat{f}_0|\} = R(0,0) - \sum_{i=1}^{3} a_k R(k,0) = \sigma^2 - \frac{\rho}{2\rho+1}[\rho+\rho+\rho]\sigma^2 = \frac{2\rho+1-3\rho^2}{2\rho+1}\sigma^2$

5. (a) Scalable coding means that the coded bit stream can be truncated at diffent point, with more bits leading to better decoded video quality.

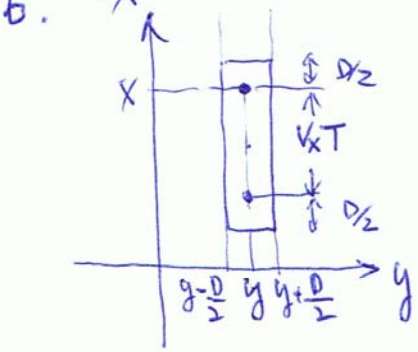
(b) Spatial scalability means more retrieved bits correspond to higher spatial resolution.

(c) Amplitude scalability means more bits lead to more accurate rendition of pixel values.

(d) JPEG 2K achieves spatial scalability by employing wavelet transform. One can retrieve increasingly higher spatial resolutions by decoding more fine scale wavelet subbands.

(e) JPEG 2K achieves amplitude scalability by applying sequential bit plane coding. One can get increasingly higher amplitude resolution by decoding more less significant bit planes.

6.

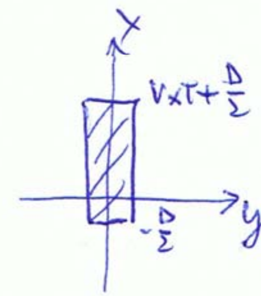


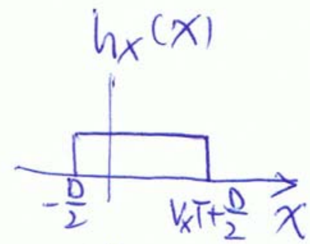
(a) Note that a sensor looking at (x,y) over a time of T will essentially integrate the original image intensity over the region shown on the left. Therefore

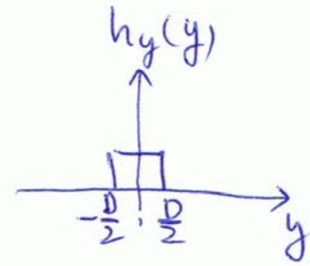
$$g(x,y) = \int_{\beta:\,-\frac{D}{2}}^{\frac{D}{2}} \int_{-\frac{D}{2}}^{\frac{D}{2}+V_xT} f(x-\alpha, y-\beta)\, d\alpha\, d\beta$$

(b) $g(x,y) = f(x,y) * h(x,y)$, with

$$h(x,y) = \begin{cases} \frac{1}{\Delta} & (x,y) \in \text{Region to the right} \\ & \Delta = D\cdot(D+V_xT) \\ 0 & \text{otherwise} \end{cases}$$







(c) $h(x,y) = h_x(x)\, h_y(y)$

$$H_x(u) = \int_{-\frac{D}{2}}^{V_xT+\frac{D}{2}} e^{-j2\pi xu}\, dx = \frac{1}{-j2\pi u}\left(e^{-j2\pi(V_xT+\frac{D}{2})u} - e^{j2\pi\frac{D}{2}u}\right)$$

$$= e^{-j\pi V_xTu}\,\frac{\left(e^{-j2\pi(\frac{V_xT}{2}+\frac{D}{2})u} - e^{j2\pi(\frac{V_xT}{2}+\frac{D}{2})u}\right)}{-j2\pi u}$$

$$= e^{-j\pi V_xTu}\,\frac{\sin(\pi(V_xT+D)u)}{\pi\cdot(V_xT+D)u}\cdot(V_xT+D$$

$\rightarrow$ Sinc((V_xT+D

$$H_y(v) = \int_{-\frac{D}{2}}^{\frac{D}{2}} e^{-j2\pi vy}\, dy = -\frac{1}{j2\pi v}\left[e^{-j2\pi v\frac{D}{2}} - e^{+j2\pi v\frac{D}{2}}\right]$$

$$= \frac{\sin \pi D v}{\pi D\cdot v}\cdot D = D\cdot \mathrm{Sinc}(Dv)$$

$$H(u,v) = H_x(u)\cdot H_y(v)$$

d) $$Q(u,v) = \frac{1}{H(u,v)}$$

(e) Since $H(u,v)$ has zeros at some frequencies, $Q(u,v)$ will be ∞ at those frequencies. This will amplify noise greatly. To fix this problem we can use several possible solutions (you only need to list one)

i) $$Q(u,v) = \begin{cases} \frac{1}{H(u,v)} & -B_1 < u < B_1, \quad -B_2 \le v < B_2 \\ 0 & \text{otherwise} \end{cases}$$

B_1 is before the first zero of $H_x(u)$
B_2 〃 〃 〃 〃 〃 $H_y(v)$

ii) $$Q(u,v) = \begin{cases} \frac{1}{H(u,v)} & \text{if } |H(u,v)| > T \\ 0 & \text{otherwise} \end{cases}$$

YW

7. (15 pt) For the given image below, (a) (4pt) Compute its approximation pyramid (also known as Gaussian pyramid) using an averaging filter. (b) (4pt) Compute the Laplacian pyramid using the pixel-replication interpolation filter. (c) (3pt) Quantize the pixels in the Laplacian pyramid using a uniform quantizer with a quantization stepsize of ~~4~~ 2. You should assume that the top level pixel has a distribution with a minimum value being zero, whereas pixels at other levels has a symmetric distribution with respect to the origin. (d) (4pt) Reconstruct the original image from the quantized Laplacian pyramid. Please use the provided grids to show your results. • For all processing, round all non integer value to keep only one decimal point

• use stepsize of 2

6	5	4	3
5	4	3	2
4	3	2	1
3	2	1	1

Original image

A

6	5	4	3
5	4	3	2
4	3	2	1
3	2	1	1

$\rightarrow$ B

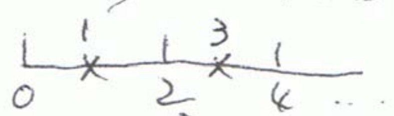
5	3
3	$\frac{5}{4}=1.3$

$\rightarrow$ $\boxed{3.1}$

Averaging every 2x2 pixels

Approximation (Gaussian) pyramid

Quantizer for top pel



F=A-E

1	0	1	0
0	-1	0	-1
1	0	0.7	-0.3
0	-1	-0.3	-0.3

1.9	-0.1
-0.1	-1.8

D=B-C

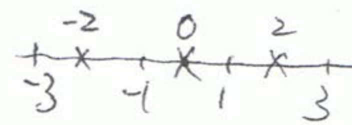
$\boxed{3.1}$ $\xrightarrow{\uparrow 2}$

3.1	3.1
3.1	3.1

C

Laplacian pyramid

Quantizer for other levels



$\uparrow 2$

5	5	3	3
5	5	3	3
3	3	1.3	1.3
3	3	1.3	1.3

E

2	0	2	0
0	0	0	0
2	0	0	0
0	0	0	0

2	0
0	-2

$\boxed{3}$

G= Q(D)

Quantized Laplacian pyramid

Take the convention

$Q(f) = 2 \quad -1 < f \leq 1$

$Q(f) = 0 \quad -2 \leq f \leq -1$

H=Q(F)

7	5	5	3
5	5	3	3
5	3	1	1
3	3	1	1

5	3
3	1

$\boxed{3}$ $\rightarrow$

3	3
3	3

I

J=I+G

Reconstructed approximation pyramid from Quantized Laplacian pyramid

if a student assumes

$Q(f)=0 \quad -1 \leq f \leq 1$

or

$Q(f)=\pm 2$ if $f=\pm 1$

Also OK

L=K+H

is the result for part (d)

5	5	3	3
5	5	3	3
3	3	1	1
3	3	1	1

K

9\. 1) Find $N \geq 4$ corresponding points between the two images, denoted by $(u_n, v_n) \to (x_n, y_n)$, $n = 1, 2, \dots, N$

2) In the intermediate image, those points will have positions (s_n, t_n), with $s_n = \frac{1}{2}(u_n + x_n)$ $t_n = \frac{1}{2}(v_n + y_n)$

3) Generate one intermediate image G_1 from F_1 so that (u_n, v_n) maps to (s_n, t_n), using inverse mapping
$$\begin{cases} u = u_1(s,t) = a_0 + a_1 s + a_2 t + a_3 st \\ v = v_1(s,t) = b_0 + b_1 s + b_2 t + b_3 st \end{cases}$$

where (a_i, b_i) are obtained by solving

$$\begin{bmatrix} 1 & s_1 & t_1 & s_1 t_1 \\ 1 & s_2 & t_2 & s_2 t_2 \\ & \cdots & & \\ 1 & s_N & t_N & s_N t_N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_N \end{bmatrix} \quad \text{or} \quad A \cdot \bar{a} = \bar{u}$$

$$\bar{a} = (A^T A)^{-1} A^T \bar{u}$$

similarly $\bar{b} = (A^T A)^{-1} \cdot A^T \cdot \bar{v}$

For every integer (s,t), $G_1(s,t) = F_1(u_1(s,t), v_1(s,t))$
(use interpolation in F_1 if $u_1(s,t), v_1(s,t)$ are non-integers)

4) Similarly generate another intermediate image G_2 from F_2, so that (x_n, y_n) maps to (s_n, t_n) using
$$\begin{cases} x = x_2(s,t) = c_0 + c_1 s + c_2 t + c_3 st \\ y = y_2(s,t) = d_0 + d_1 s + d_2 t + d_3 st \end{cases}$$

with $\bar{c} = (A^T A)^{-1} A^T \cdot \bar{x}$, $\bar{d} = (A^T A)^{-1} A^T \cdot \bar{y}$

For every integer (s,t), $G_2(s,t) = F_2(x_2(s,t), y_2(s,t))$

5) Combine the two intermediate image

$$G(s,t) = \frac{1}{2}\left[G_1(s,t) + G_2(s,t)\right].$$