פתרונות
בחינה בקורס מבוא לעיבוד תמונות

203.2730 סמסטר ב' מועד א' תשס"ד
שם המרצה: ד"ר חגית הל-אור.
משך הבחינה: שעתיים.

Question 1

4 sections - each section 8 points.

א Since the image histogram does not preserve spatial locations associated with gray value, the only possibility of reconstructing an image is when the image is of uniform gray value. That is the Histogram shows 0 pixels at all gray values except for a unique gray value where the count equals the number of pixels in the image.
Thus the number of images that can be reconstructed from its histogram equals
The number of possible gray values.

ב The histogram shows that 50 pixels are of value 10 and 50 are of value 20.
Thus the number of images equals the number of possibilities of choosing 50 pixels of value 10 from amongst the 100 image pixels = '100 choose 50' =

$$\binom{100}{50} = \frac{100!}{50!50!}$$

ג In addition to the above, the x-derivatives are constrained to be of value 10 or -10.
This means that there no neighboring pixels along a row that have the same gray value.
Thus each row of the image must have alternating values of 10 and 20.
There are 2 such possibilities per row (determined by the 2 possibilities of setting the 1^{st} pixel of the row). Since the rows are independent of each other, the number of possibilities is:

$$\underbrace{2*2*2*\ldots..*2}_{\text{10 times}} = 2^{10}$$

ד In addition to the above, the y-derivatives are constrained to values of 10 and -10,
Thus values within each column of the image alternate between 10 and 20.
Alternating in rows and in columns, constrains the image to be a 'chess' image of values 10 and 20. Since the any one pixel (say at coordinate (1,1)) determines the 'chess' pattern, there are only 2 possible images.

Question 2

2 sections – Section א 18 points, Section ב 15 points.

א The text is assumed to be of brightest (left image – I1) and darkest (right image – I2) gray value. The text is assumed to be of constant value.

Step 1: Determine threshold values for each image.

T1 = max(I1(:))

T2 = min(I2(:))

or: create histogram of images and find max/min bin with count greater than 0.

(Note: we can not assume the text values are 255 and 0).

Step 2: perform thresholding of each image and combine the results:

finalImage = (I1 >= T1) + (I2<=T2)

Final image has background 0 and text 1.

ב To estimate I(x,y) from the 100 acquired images Fi, average the Fi images and subtract 2. The reason this works is as follows:

$$\frac{1}{100}\sum_{i=1}^{100} F_i(x,y) - 2 = \frac{1}{100}\left(\sum_{i=1}^{100} I(x,y) + \sum_{i=1}^{100} N_i(x,y)\right) - 2 =$$

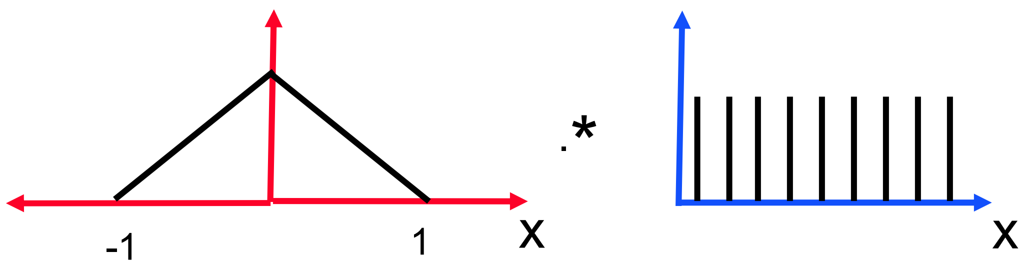
$$I(x,y) + \underbrace{\frac{1}{100}\sum_{i=1}^{100} N_i(x,y)}_{\text{mean noise = 2}} - 2 \cong I(x,y) + 2 - 2 = I(x,y)$$

The standard deviation is meaningless except for the fact that it implies that the mean noise does not deviate a lot from 2.

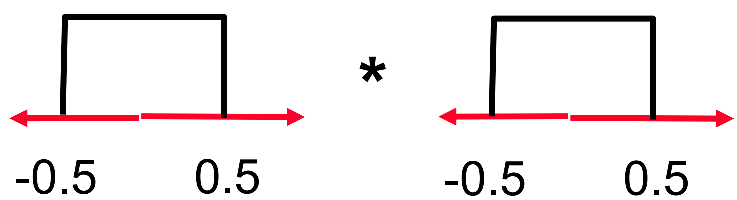
Question 3

2 sections – Section א 16 points, Section ב 17 points.

א The input image is a pixel-by-pixel multiplication of a ‘pyramid’ and a comb:



The ‘pyramid is a convolution of 2 ‘rect’s :



Thus we have that the input image = (rect * rect) .* comb

Following the Convolution theorm, the Fourier transform of the image equals:

(F(rect) .* F(rect)) * F(comb)

The fourier transform of rect is a sinc function, the transform of the comb is a comb with 2 peaks (per cycle) at distance N/2 of each other (this was proven In class). Thus the transform of the image is a convolution of sinc^2 with the 2-peak comb, giving:

ב. The solution is to perform image rectification.

* Select the 4 corners of the mosaic denote them x_i, y_i and in matrix notation:

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{bmatrix}$$

* Select the destination coordinates of these corners.

$$X' = \begin{bmatrix} x_1' & x_2' & x_3' & x_4' \\ y_1' & y_2' & y_3' & y_4' \end{bmatrix}$$

The destination must be square but exact coordinates were not requested, so we can choose. Without loss of generality we can choose coordinates (0,0) (1,0) (0,1) (1,1).

* Determine the transformation **A** that maps the coordinates of the new image to Their source location in the original image.

$$AX' = X$$

Since the system is over-determined, use pseudo-inverse to determine **A** :

$$A = X \ pinv(X')$$

* Perform inverse-mapping by applying **A** to all coordinates of the new image and Determining the source location in the original image.
Use Bilinear interpolation or NN to determine gray value of each location.

<u>NOTES:</u>
One can estimate **A** as the direct mapping and then invert **A** to perform the inverse mapping.

One can add another feature to serve as the origin of the transformation **A** (one of the corners or the center of the mosaic). In this case an origin in the original image and an origin in the destination image must be defined. The features X and X' must be adjusted to be around the selected origin (by subtracting the corresponding origin's coordinates) prior to estimating **A**.
Adding origins is not necessary though it might improve the result because A can then be more precise.

Question 4

9 sections – each 3 points.

Consider first the images A,F,H : These are results of correlation. The masks are reflected parts (or whole) of the original image, so that convolution with them performs correlation.

A – Mask 7

Correlation of an image with itself produces a peak response when the mask is centered on the image. Assuming the origin of the mask is at the center, the peak should be in the image center as in A.

H – Mask 8

Correlation with a mask that equals the original image + a shifted version of the image. The result should be a peak response at the center of the image (assuming again the origin of the mask is in the center) + another peak at the shifted location.
The location is of course mirrored with respect to the origin and thus the second peak appears above and to the left of the origin in the resulting image.

F – Mask 3

Correlation with a part of the image. Since non-normalized correlation is performed, response values will be high for any area in the original image that is bright (large gray values).
Thus the resulting image has high values in the part of the image corresponding to the mask and also in other bright areas such as the melon in the top left corner of the image).

Next consider images B and D : these have mean 0 (and have been normalized to 0..255).
There are only 2 masks that sum to 0 : Mask 1 and 4. Note mask 5 does NOT sum to 0.

D – Mask 1

The mask is an oriented laplacian. Thus only vertical high frequencies corresponding to horizontal edges, are seen in the resulting image. To distinguish between D and B, notice that there are mach bands along the edges (i.e. a dark band next to a bright band). If this is not clear in the reproduction, one can at least distinguish the gray value difference at the edge – see explanation for B.

B – Mask 4

The mask is an edge detector mask, detecting horizontal edges. These can be seen in the resulting image. Consider the edge between the apple at the top left corner and the orange below it. The apple is dark and the orange is bright. Thus the edge detector causes the dark values to be multiplied by -1 and the bright values by 1 (remember the mask is flipped for convolution). Resulting in high values at the edge as seen in the resulting image as bright pixels (this contrasts with the dark pixels dominant for the same edge in image D.

Next consider the remaining images.

G – Mask 2
Image G is the original image with a horizontal ‘echo’. That is the image is the original image + a shifted version of the original. (this is NOT a blurred image).
This corresponds to convolving the image with 2 delta functions as can be seen in the mask.
Note the scale: the mask presents 2 deltas with a shift of 5 and this is a small shift compared to the size of the original image (256x256).

C – Mask 5
This mask is a high frequency enhancing mask. It equals a laplacian + delta (a laplacian mask would have 8 in the center so that the mask sums to 0).
The sum of the mask values equals 1 so that the image average is maintained.
The resulting image is a high frequency enhanced (sharpened) original.
One can actually see the Mach bands along the edges.

Finally consider Images E and I.
Both these images are blurred versions of the original but image I has added artifacts which We called ‘ringing’ artifacts. One can not really tell which image is blurred stronger.

E – Mask 6
Mask 6 is a gaussian mask (values sum to 1).

I – Mask 9
The mask is a sinc function, thus its Fourier transform is a rect function.
Multiplying the transform of the original image with the rect (pixel-by-pixel) maintains the low frequencies and zeros the high frequencies – that is we have an ideal filter.
As discussed in class, the ideal filter suffers from Ringing.