COMP.SGN.110 Introduction to Image and Video Processing

Exam – 31/03/2021

Instructor: Professor Moncef Gabbouj

Last Name: ____________________________

First Name: ____________________________

Student Id. Number: ____________________

Instructions: Due to COVID-19, we organize a remote exam to allow students to progress in their studies. We therefore count on you to abide by the following rules.

1. Write your name and student ID number on the exam paper. DO THIS NOW.
2. Do not talk to anyone during the exam. Cheating in the exam violates good study practices (https://intra.tuni.fi/en/handbook?page=2255) and University actions described in "Possible consequences of ethical misconduct" may apply.
3. Answer the questions in the space provided, and
4. Submit your exam via Moodle by 20:00 (Finland time). Late submissions will NOT be allowed by the system and will not be graded. If you submit a photo scan of your exam, make sure it is readable, if you cannot read, I cannot read it and it will therefore not be graded. Typed text is preferred, but a good quality scan of hand-written answers is ok.

Last Name: Nguyen.
First Name: Quoc Minh.
Student Id. Number: H290130.

Problem 1 (5 marks):

Consider the following frequency-domain (centered) filter:

$$H(u,v)=\begin{cases}0 & \text{if } (u,v)=(M/2,N/2)\\ 1 & \text{otherwise}\end{cases}$$

(a) What is the output of this filter when the input is an 8-bit grey-scale image $f(x,y)$ of size *M* by *N*?
(b) What is this filter commonly referred to (name of the filter)?
(c) Let $g(x,y)$ be the filtered image, provide the closed-form expression of $g(x,y)$ as a function $f(x,y)$.
(d) Show that *H(u,v)* is a linear filter.
(e) Determine if *H(u,v)* is a lowpass, a high-pass or a bandpass filter.

Solution:

(a) The output will be the image that has averge of zero.
(b) The name of this filter is removal of image average.
(c) We can write $\mathrm{H(u,v)}$ as

$$\mathrm{H(u,v)} = 1 - \delta(\mathrm{u} - \frac{\mathrm{M}}{2}, \mathrm{v} - \frac{\mathrm{N}}{2})$$

Let the Fourier transform of the image is F(u, v), so the Fourier transform of the filter image is

$$G(u,v) = F(u,v) - F(u,v)\,\delta(u - \frac{M}{2}, v - \frac{N}{2})$$

So take the inverse Fourier transform of the above equation to obtain the filter image, here * is the convolution operation.

$$g(x,y) = f(x,y) - \frac{1}{MN} f(x,y) * (-1)^{x+y}$$

(d) Suppose we have to image $f_1(x,y)$ and $f_2(x,y)$ with their linear combination $af_1(x,y) + bf_2(x,y)$. The output in frequency domain is

$$G(u,v) = \left(aF_1(x,y) + bF_2(x,y)\right) - \left(aF_1(x,y) + bF_2(x,y)\right).\,\delta\left(u - \frac{M}{2}, v - \frac{N}{2}\right)$$

$$= a(F_1(u,v) - F_1(u,v)\,\delta(u - \frac{M}{2}, v - \frac{N}{2})) + b(F_2(u,v) - F_2(u,v)\,\delta(u - \frac{M}{2}, v - \frac{N}{2}))$$

We see that this result in the linear combination of filtered result from each image. So this filter is linear.

(e) This is a high-pass filter.

Problem 2 (5 marks):

Consider the following adaptive, local noise reduction filter:

$$g(x,y) = f(x,y) - \frac{\sigma_\eta^2}{\sigma_L^2}[f(x,y) - m_L]$$

where *g(x,y)* is the filter output, *f(x,y)* is the degraded image to be restored, σ_η^2 is the variance of the noise, σ_L^2 is the local variance of the pixels in a local neighborhood S_{xy}, and m_L is the local mean of the pixels in S_{xy}.

Consider a 3x3 local neighborhood S_{xy}.

(a) Give the expression of m_L

(b) Give the expression of σ_L^2

Discuss the behavior of the filter in the following three cases:

(c) when there is no noise,

(d) when the local area contains important details, e.g., edges, and

(e) when the local area is smooth.

Solution:

Denote pixel value at location (m, n) is $p_{m,n}$

(a) Local mean of S_{xy}, center at (i, j)

$$m_L = \frac{1}{9}\sum_{m=-1}^{1}\sum_{n=-1}^{1} p_{i+m,j+n}$$

(b) Local variance of S_{xy}, center at (i, j)

$$\sigma^2 = \frac{1}{9}\sum_{m=-1}^{1}\sum_{n=-1}^{1} (p_{i+m,j+n} - m_L)^2$$

(c) When there is no noise $\sigma_\eta^2 = 0$, then $g(x, y) = f(x, y)$, the filter does nothing and returns the original image.

(d) When the local area contains important details such as edges, σ_L^2 will have large value. So the second term on the equation in the problem would be small, the $g(x, y)$ will be close to $f(x, y)$. Thus, the important details are preserved.

(e) When the local area is smooth, σ_L^2 will have small value since we have the assumption $\sigma_L^2 >= \sigma_\eta^2$, so the second term in the equation in the problem will be close to $f(x, y)$ - m_L, so the output $g(x, y)$ will be close to m_L. The filter behaves like an arithmetic mean filter.

Problem 3 (5 marks):

A skilled medical technician is charged with the job of inspecting a certain class of images generated by an electron microscope. In order to simplify the inspection task, the technician decides to use digital image enhancement. He examined a set of representative images and identified a number of problems. You are asked to assist the technician by **proposing** and **justifying** a solution to each problem below:

(a) Several bright, dark and isolated dots in the images are of no interest and must be removed, but how?
(b) the images suffer from lack of sharpness, so what should he do?
(c) some of the images lack contrast, so what to do?
(d) some shifts in the average gray-level value were noticed, when this value should be V to perform correctly certain intensity measurements, how to fix this problem?
(e) The technician wants to display in white all grey levels in the range [I_L, I_2] while keeping the remaining grey levels as they are. What should he do?

Solution:

(a) We can use median filter. Since the bright and dark spot in the isolated region are pixel that has outlier value from other pixel value in local region, the median filter can remove those pixel values.
(b) We can use Laplacian filter to enhance the sharpness of the image.
(c) We can enhance the contrast of image by using power-law transformation $T(r) = cr^{\gamma}$ for $c > 0$, and $0 < \gamma < 1$.
(d) We can add the difference between the desired average gray-level and image average gray-level to all the pixel in the image, also remember to cast those pixel value outside of the limit back to the allowed region.
(e) Suppose pixel value is in $[0, L]$. We can apply transformation: intensity-level slicing:

$$f(x) = \begin{cases} x, & \text{if } 0 <= x < I_1 \\ L, & \text{if } I_1 <= x <= I_2 \\ x, & \text{if } I_2 < x <= L \end{cases}$$

Problem 4 (5 marks):

An archaeologist who studies human history through the excavation of sites and the analysis of artefacts and other physical remains stumbled into a set of old grey scale images of ancient Egyptian pyramids. The image content seems valuable, but the image quality was poor. A summer intern who just completed the first course on image and video processing was hired and asked to **propose** and **justify** image processing tools to restore the images. These are some of the issues which were identified:

(a) horizontal streaks (lines) are visible across some of the images and need to be removed.
(b) noise is apparent in all images, how to find out what types of noises are there?
(c) among the different types of noises, some random black dots and some random noise are clearly visible in the images, how to remove them?
(d) a bright region appears in a relatively clean corner of one image, it is believed to originate from a streetlight, the archaeologist thinks it can be useful to restore his images, what should be done?
(e) the results are so far not satisfactory, so the archaeologist is getting desperate and asks the trainee to try the best he learned on these images. What drastic measures should be done in addition to these ad-hoc methods?

You are the trainee, help the expert archaeologist by designing individual techniques and a sequence of processing algorithms to achieve the desired goals.

Solution:

(a) We can use the Fourier transform to the image, and move the DC component to the center of the transform. Those horizontal streaks would appear as bright and maybe isolated spot in the vertical axis of the magnitude of the transform. The we can suppress those spots by mask those spot with zeros value and keep other coefficient unchanged. If those spot are not too isolated, we may need to make a rectangle mask to suppress a large region the vertical axis that we think it relevant to the horizontal streak, this may result in some detail loss. After all that, we convert the transform to the spatial domain to obtain the clean image.

(b) First we choose a region in the image that we think that it is the smooth region without the noise, maybe the surface of object with no texture. Then we draw the histogram of that region. By the effect of the noise, the histogram of a flat region will be shaped by shaped of the distribution of the noise. For example, if the histogram has bell shape, the noise maybe Gaussian type, if the histogram somehow remains at the same level, this might be uniform noise.

(c) We can use median filter or contraharmonic filter with $Q > 0$ to remove the random black dot, and then apply mean filter to remove random noise.

(d)

Problem 5 (5 marks):

An airport security firm hires a new image processing engineer and tasks the engineer with the job of designing a detector of laptops hidden in travelers' bags going through X-ray imaging at security check points. The engineer recalls a similar problem in the image processing course at the University, see the academic problem below, and wanted to follow similar steps to solve the assignment. Argue the claims below and answer the questions:

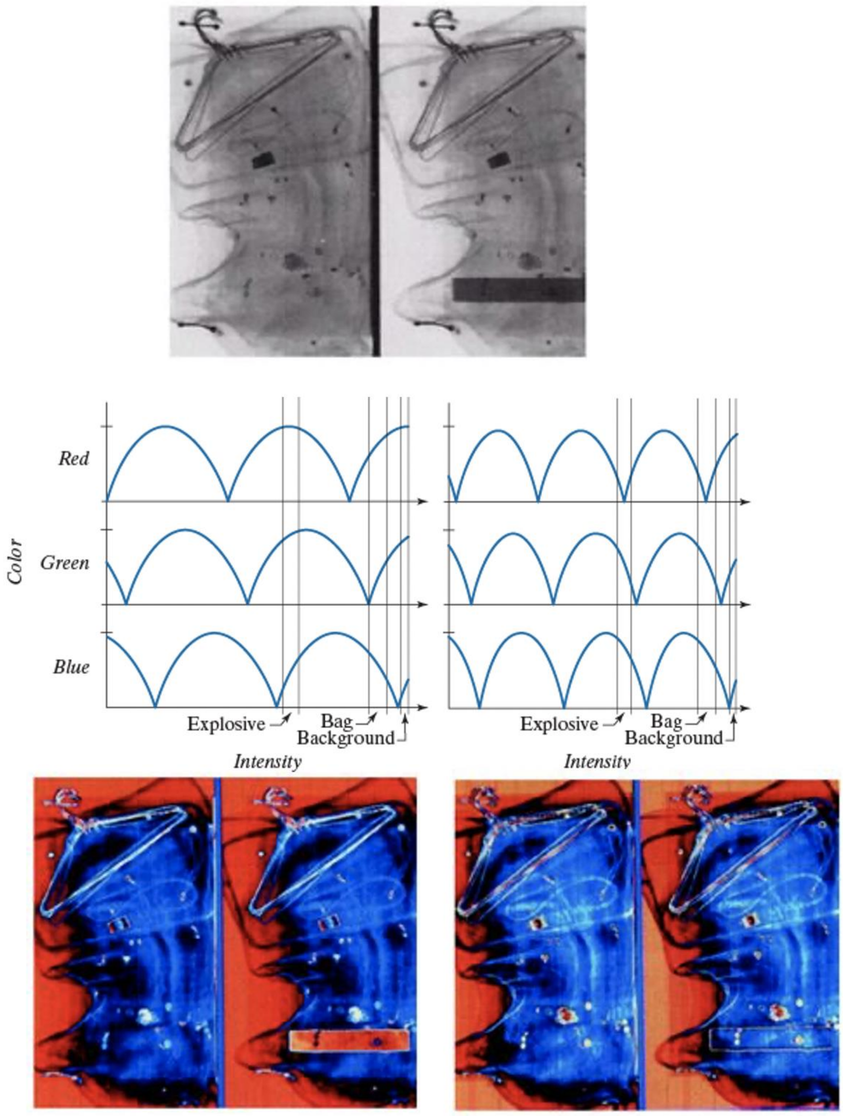


Figure 1: Academic Problem: (top) X-ray images; (middle) two color transformations: Transformation 1 on the left-hand side, Transformation 2 on the right-hand side; (bottom) pseudo-colored images corresponding to Transformation 1 (Solution 1 on the left-hand side) and Transformation 2 (Solution 2 on the right-hand side).

(a) Assume that the laptop has similar characteristics as the "Explosive" in the Academic Problem above, argue whether any of the two solutions presented in the Academic Problem is viable for the problem at hand.
(b) Propose an alternative solution and show how it addresses the problem at hand.
(c) The engineer wants to further impress his new boss and thought about applying a modified Gaussian high-pass filter (see expression below) as a pre-processing step before applying his color transformations.

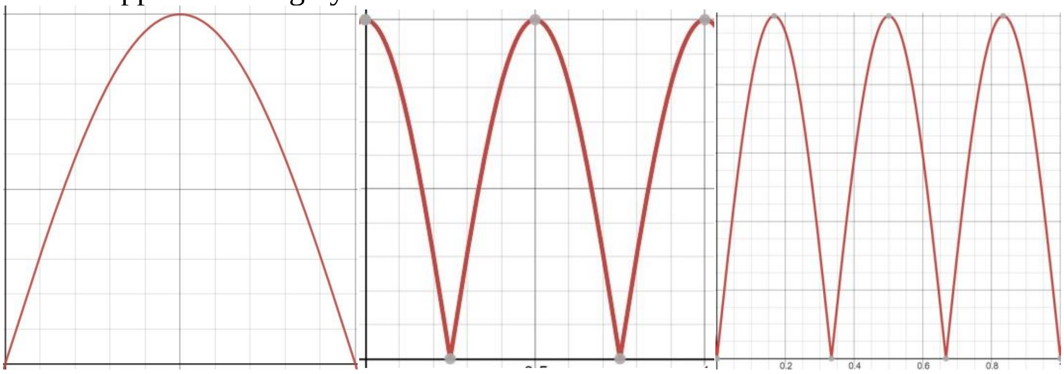
$$H(u,v) = (\gamma_H - \gamma_L)\left[1 - e^{-c\left[D^2(u,v)/D_0{}^2\right]}\right] + \gamma_L$$

Help the engineer motivate this choice of pre-processing (what is the relevance in the problem at hand) and explain to the boss how the method works in practice (how the filter works, how to select it parameters)?

Solution:

(a) The second transformation (right) assigns the "explosive" and the bag to the same color, so we can see through to "explosive". So this transformation is not useful to detect the object. The first transformation (left) assigns the "explosive" and the bag to different color and since the "explosive" is inside the bag, we can distinguish it. So the first transformation can be used to detect the laptop but we must be careful since this transformation assigned the target object to the same color with the background.

(b) Different from above, we can make the frequency of transformation in each coloring channel to be lower, so each gray level in will have the unique color after the coloring. Example of transformation is below, the horizontal axis is from the lower limit to upper limit of gray level.



Red Green Blue

We can also shift in phase for each color channel transformation to make the distinguish between the laptop and the background and the bag to be better.

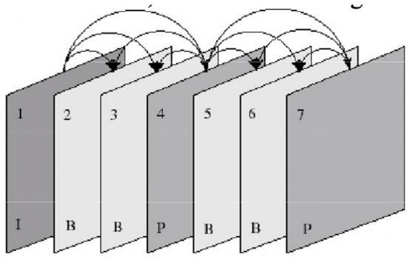
(c) This filter helps to lower the effect of the background illumination, for example the effect of the cover of the bag will be attenuated help we see clearer what is inside the bag. This filter simultaneously boosting the contrast of the image, for example it makes the object inside the bag look more sharper and easier to distinguish between the objects.
For low frequencies component, the exponential term will be close to 1, result in $\mathrm{H(u,v)}$ closes to γ_{L} . For high frequency component, the exponential term will be close to zero, result in H(u, v) close to γ_{H}. So to obtain the effect of attenuate low

frequency and boosting high frequency, we need to choose $\gamma_H > 1$ and $\gamma_L < 1$. Parameter c controls the slope of the transition of $H(u, v)$, we can experiment with it to find suitable value.

Problem 6 (5 marks):

When encoding a group of pictures (GOP), there is a risk of error propagation in video encoding.

(a) What is error propagation
(b) How can it be controlled or stopped?
(c) What is the cost of stopping error propagation?
(d) How can we reduce this cost?
(e) Given the GOP below, identify the frames from which errors may propagate.



Solution:

(a) In encoded step, the predicted frame may contain some error, if we use this predicted frame as reference frame to encode other frame, the error will propagate to this new frame.
(b) Error propagation can be solved by encoding reference regions or frames, called I-frame at constant intervals of time. Some frame called P-frame is used, this frame is only encoded with reference to previous frame. Other frames between the reference frame and P-frame will be encoded with reference to the reference frame and that P-frame.
(c) Since some frames or regions is encoded as a still image, the compression performance will be decrease.
(d)
(e) Suppose frame 1 contains no error.
Frame 4 is encoded with reference to frame 1, so error may start at frame 4.
Frame 2 and frame is encoded with refernce to frame 1 and frame 4, so the error propagates from frame 4 to frame 2 and from frame 4 to frame 3.
Frame 7 is encoded with reference to frame 4, so error propagates from frame 4 to frame 7.
Frame 5 and 6 is encoded with reference to frame 4 and frame 7, so the error propagates to frame 5 and frame 6.