EQ2330 Image and Video Processing

Exam, January 7th, 2021, 08:00-13:00

This is a 5 hour, remote exam. The following aids are allowed: A scientific calculator, your course notes, course book, as well as an approved mathematical handbook[1].

Note: show that you understand the principles: do not spend much time on tedious numerical computations that provide little reward. Use reasonable approximations where this is appropriate. If the equations are not simple to derive or very important to the class, then they will be provided.

Start by browsing through the exam. The point value of each problem and subproblem is shown. These points are obtained if the solution is correct and clearly motivated. The maximum exam score is 100 points.

Responsible teacher:	Markus Flierl	08-790-7425
During the exam, questions are also answered by:	Linghui Zhou	

Good luck!

[1] Beta (Råde, Westergren); Taschenbuch der Mathematik (Bronstein, Semendjajew); Taschenbuch Mathematischer Formeln (Bartsch); Matematicke vzorce (Bartsch); Collection of Formulas in Signal Processing (Dept, S3, KTH); Mathematische Formeln—Erweiterte Ausgabe E (Sieber).

1. *True or False Statements (10p)*

   Are the statements below true or false? You get 1 point if your answer is correct, 1 point is deducted if your answer is wrong, and 0 points if no answer is given. The total score for this problem will not be smaller than zero. Answer only true, false, or no answer. Do not provide any motivation (only this problem).

   (a) ($\pm$1p) The Karhunen-Loeve transform only requires first order statistics.

   (b) ($\pm$1p) DCT always achieves optimal energy concentration.

   (c) ($\pm$1p) Histogram equalization always yields uniform *pmf*.

   (d) ($\pm$1p) Two different images cannot have exactly the same histograms.

   (e) ($\pm$1p) Lloyd-Max algorithm always converges to a globally optimal quantizer.

   (f) ($\pm$1p) Biorthogonal wavelet transform are critically sampled.

   (g) ($\pm$1p) The distortion curve D(R) of a Gaussian source is always convex.

   (h) ($\pm$1p) Generally speaking, median filtering removes outliers whereas low pass filters smooth out the outlier over space.

   (i) ($\pm$1p) Huffman coding can result in a fixed length code.

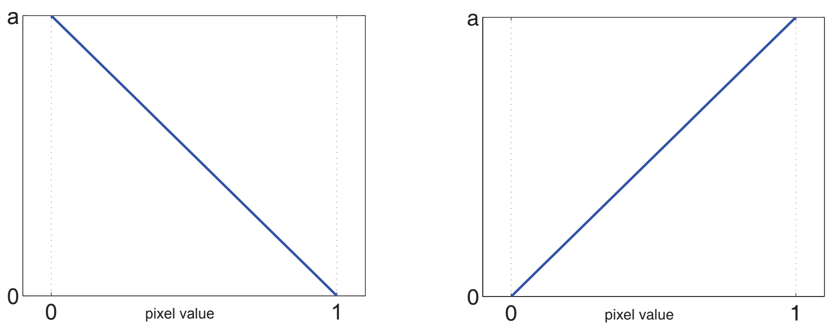
   (j) ($\pm$1p) Rate distortion functions $R(D)$ are monotonically increasing for $D > 0$.

2. *Histograms (15p)*

In this problem we consider histogram equalization of images with pixels taking values from the continuous interval $[0, 1]$. We use the convention that 0 denotes a black pixel and 1 denotes a white pixel. We first consider an image shown in Fig. 1(a). In this problem we slightly abuse the terminology and we use the term "histogram" to refer to the probability density function of the pixel values in the image. The continuous histogram (probability density function) of the image is shown in Fig. 2(a).

(a) Original Image (b) Processed Image 1 (c) Processed Image 2

Figure 1: Images considered in the histogram equalization problem.



(a) Histogram 1 (b) Histogram 2

Figure 2: Histograms considered in the histogram equalization problem.

(a) (2p) Derive a transformation that equalizes the histogram given in Fig. 2(a). Sketch the equalized histogram. *Hint: You need to find a valid expression describing the probability density function shown in Fig. 2(a).*

(b) (1p) The image shown in Fig. 1(b) has been processed by some transformation applied pixel-wise and the processed image has a histogram shown in Fig. 2(b). Your goal in (b) and (c) is to find that transformation. First, find a transformation that equalizes the histogram shown in Fig. 2(b) and proceed to (c).

(c) (2p) Next, using the results from (a) and (b), derive a transformation that applied to the original image shown in Fig. 1(a) yields the image shown in in Fig. 1(b).

(d) (3p) Prove that the transformation computed in (c) and applied to the image in Fig. 1(a) in fact yield the histogram shown in Fig. 2(b). *Hint: Use a relation between the original probability density function of a random variable and the probability density function of a transformed random variable.*

(e) (1p) Now consider the image shown in Fig. 1(c). The image was obtained by applying the transform $T(r) = 1 - r$. Sketch the histogram of the processed image.

(f) (2p) Explain the key difference between the transformation in (c) and (e).

(g) (2p) Assume that all the images considered in this problem, Fig. 1(a) - Fig. 1(c), are quantized using the same uniform quantizer Q resulting in three sets of quantization indices. Assume that all the quantization cells of Q are within the interval $[0, 1]$. Consider the entropy of the indices within respective sets. Which set of the indices is associated with the largest value of the entropy? Explain.

(h) (2p) We perform a pixel-wise addition of the image shown Fig. 1(a) and Fig. 1(c). What is the histogram associated with the result of the addition? Compare the obtained result to the case where images shown in Fig. 1(a) and Fig. 1(b) are added pixel-wise. What is the histogram in this case? Provide explanations.

3. *Linear Image Processing (15p)*

In this problem we focus on linear image processing in spatial and frequency domains.

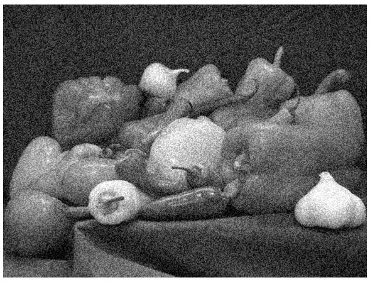
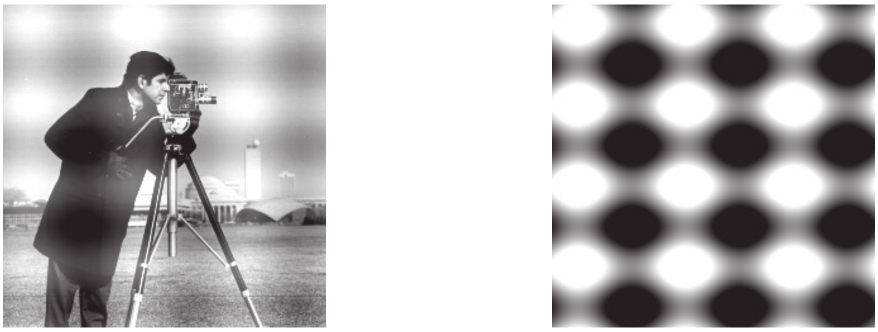


Figure 3: Image of *peppers.*

(a) (2p) The image (Fig. 3) is degraded by uniform noise with zero-mean and unknown variance. You know from the course that such an image can be enhanced by using a simple averaging filter that is applied in the spatial domain. For some reason you prefer to implement the equivalent filter in the frequency domain. Provide a transfer function of such a two-dimensional filter (e.g., sketch the transfer function). Explain, why it can enhance the image.



(a) Degraded image (b) A realization of the degrading signal

Figure 4: Degraded image of *cameraman* and a **possible** realization of the degrading signal.

(b) (4p) The image shown in Fig. 4(a) (256 ×256 pixels) is degraded by additive two-dimensional noise

$$n(x, y) = a\sin(2\pi x f_x + \phi_x) + a\sin(2\pi y f_y + \phi_y) \tag{1}$$

with unknown phases ϕ_x and ϕ_y and constant frequencies f_x and f_y, as shown in Fig. 4(b). Sketch the transfer function of an ideal notch filter that eliminates the noise $n(x, y)$. Characterize the class of images for which the notch filter is MSE-optimal.

(c) (2p) Consider a filter as described by the following filter mask:

$$\mathbf{h} = \frac{1}{10}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Suppose that we apply this filter to an arbitrary image. We take the result and apply the filter again. Suppose that we make n such iterations. What is the final result if $n \to \infty$?

(d) (3p) Show that two-dimensional convolution with the following filter mask can be decomposed into two one-dimensional convolutions.

$$\mathbf{g} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

State the condition for the two-dimensional mask that allows for such a decomposition.

(e) (2p) An image represented by $f(x, y)$ was multiplied by $(+1)^{(x+y)}$. How would you undo the effect of this multiplication by processing in the frequency domain?

(f) (2p) An image was degraded first by Gaussian noise and then by salt and pepper noise. Suppose you can apply filters to the degraded image that perform averaging and median filtering. Propose an order for these filtering operations that will yield the best enhancement for the degraded image. Provide an explanation!

4. *Multiresolution Processing (15p)*

Images can be represented at more than one resolution. The image pyramid is a simple structure for a multiresolution representation. Fig. 5 is the system block diagram for creating an image pyramid. The original image is on the base level J of the pyramid and j is an intermediate level in the pyramid, where $0 \leq j \leq J$. A fully populated pyramid is composed of $J+1$ resolution levels. The base of the pyramid contains a high resolution representation of the image being processed, while the apex contains a low resolution approximation. As you move up the pyramid, the resolution decreases.

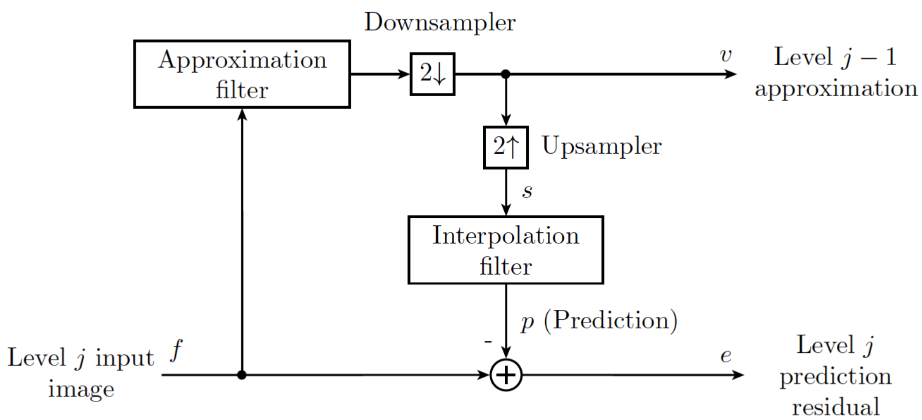


Figure 5: System block diagram.

Now, we consider a one-dimensional signal as input for simplicity. The Z-transform of a discrete signal $x[n]$ is given as $X(z) = \sum_{n=-\infty}^{\infty} x[n]z^n$. Using the Z-transform, the signal with downsampling and upsampling by factor 2 can be expressed as follows:

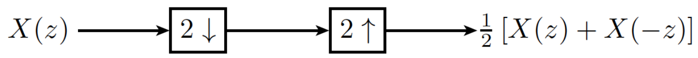


Figure 6: Downsampling and upsampling by factor 2.

Assume the frequency spectrum of the one-dimensional input signal is given in Fig. 7.

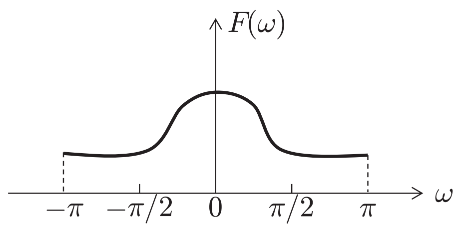


Figure 7: Frequency spectrum.

(a) (3p) The original image has a size of $N \times N$. The downsampling factor is 2 in each direction. Calculate the total number of elements in the fully populated pyramid. Is this multiresolution representation critically sampled?

(b) (3p) In the following, consider a one-dimensional signal as input. Let the approximation filter and the interpolation filter in Fig. 5 be ideal lowpass filters that preserve the spectrum within $[-\frac{\pi}{2}, \frac{\pi}{2}]$ over a 2π period. Sketch the frequency spectrum of the prediction signal p. Motivate.

(c) (1p) Determine and sketch the frequency spectrum of the residual signal e.

(d) (3p) Now, the residual signal e is fed into the system as shown in Fig. 8. To obtain an output $\hat{e}$ that is identical to e, what conditions are needed for $H(z)$ and $G(z)$? Are they highpass or lowpass filters? Please sketch the two filters.

$$e \longrightarrow \boxed{H(z)} \longrightarrow \boxed{2\downarrow} \longrightarrow \boxed{2\uparrow} \longrightarrow \boxed{G(z)} \longrightarrow \hat{e}$$

Figure 8: Downsampling and upsampling with filters $H(z)$ and $G(z)$.

(e) (1p) In general, the approximation filter and the interpolation filter in Fig. 5 are not ideal filters. What resulting effect on the signal will the samplers introduce?

(f) (4p) Draw a multiresolution system that cancels the resulting effect for non-ideal filters. Analyze your system with the help of Fig. 6 and give the conditions for the filters to cancel such an effect and to allow perfect reconstruction for the signal.

5. *Motion Compensation (15p)*

Motion compensation is commonly used in video coding. In this problem, we consider motion compensation used on a bandlimited one-dimensional signal s. Assume the signal $s[x]$ with $x \in \mathbb{Z}$ is sampled at integer positions. The autocorrelation sequence of $s[x]$ is

$$\phi_{ss}[x] = \mathbb{E}\{s[x_0 + x]s^*[x_0]\} = \sigma_s^2 \rho^{|x|}, \ x \in \mathbb{Z}, \tag{2}$$

where σ_s^2 is the variance of the signal, and $0 < \rho < 1$ the correlation coefficient. The motion-compensated prediction signal is modeled as

$$c[x] = s[x - \Delta] - n[x], \ x \in \mathbb{Z}, \tag{3}$$

where $\Delta \in \mathbb{Z}$ is the displacement error and $n[x]$ the noise. We assume that the noise $n[x]$ is statistically independent of $s[x]$. The variance of the noise is σ_n^2 and the autocorrelation sequence of the noise is

$$\phi_{nn}[x] = \sigma_n^2 \delta[x], \tag{4}$$

where $\delta[x]$ is the Kronecker delta. The error between the original signal $s[x]$ and the motion-compensated signal $c[x]$ is

$$e[x] = s[x] - c[x], \ x \in \mathbb{Z}. \tag{5}$$

(a) (2p) What is the purpose of using motion compensation in video coding?

(b) (4p) Derive the autocorrelation sequence $\phi_{ee}[x]$ of the error $e[x]$ in terms of ϕ_{ss} and ϕ_{nn} for a given $\Delta \in \mathbb{Z}$.

(c) (2p) Determine the variance of the error $e[x]$ in terms of σ_s^2, σ_n^2, and ρ.

(d) (5p) Determine the expressions of $\phi_{ee}[x]$ for $\Delta = 0$, $\Delta = 1$, and $\Delta \to \infty$. Compare the three cases and discuss the quality of motion compensation in terms of the displacement error.

(e) (2p) How can the above motion-compensated prediction be improved?

6. *Huffman Coding (15p)*

Consider the simple 4×8, 8-bit image:

$$\begin{matrix} 23 & 23 & 23 & 97 & 171 & 247 & 247 & 247 \\ 23 & 23 & 23 & 97 & 171 & 247 & 247 & 247 \\ 23 & 23 & 23 & 97 & 171 & 247 & 247 & 247 \\ 23 & 23 & 23 & 97 & 171 & 247 & 247 & 247 \end{matrix} \tag{6}$$

(a) (2p) Compute the entropy of the pixel values.

(b) (2p) Compress individual pixel values of the image by using Huffman coding.

(c) (2p) Does the Huffman code satisfy Kraft's inequality?

(d) (2p) Compute the average code word length and the redundancy of the Huffman code.

(e) (2p) Consider Huffman encoding of pairs of adjacent pixels rather than individual pixels. That is, we form pairs of adjacent pixels, where we represent the paired pixels with a new vector symbol. What is the entropy of the image when encoded as pairs of adjacent pixels?

(f) (3p) Consider coding the differences between horizontally adjacent pixels. For the new difference image, we keep the first column of the original image to maintain the image size of 4x8. What is the entropy of the new difference image? Can such a method be used for image compression? Motivate your answer.

(g) (2p) Explain the entropy differences in (a), (e) and (f).

7. *Image Segmentation (15p)*

Let us assume that an image contains only two principal gray level regions. Suppose that through a statistical analysis of the image, we find that the probability density function is given by

$$p(x) = P_{fg}f_{fg}(x) + P_{bg}f_{bg}(x),$$

where x denotes the gray level values, P_{fg} the probability that a pixel belongs to the foreground, $f_{fg}(x)$ the probability density function for pixels belonging to the foreground, P_{bg} the probability that a pixel belongs to the background, and $f_{bg}(x)$ the probability density function for pixels belonging to the background. The conditional probability density functions are shown in Fig.9. Since any given pixel belongs either to the foreground or to the background, we therefore have $P_{fg}+P_{bg} = 1$. In this problem, we consider *supervised thresholding* which classifies all image pixels with gray level greater than the threshold T as foreground.

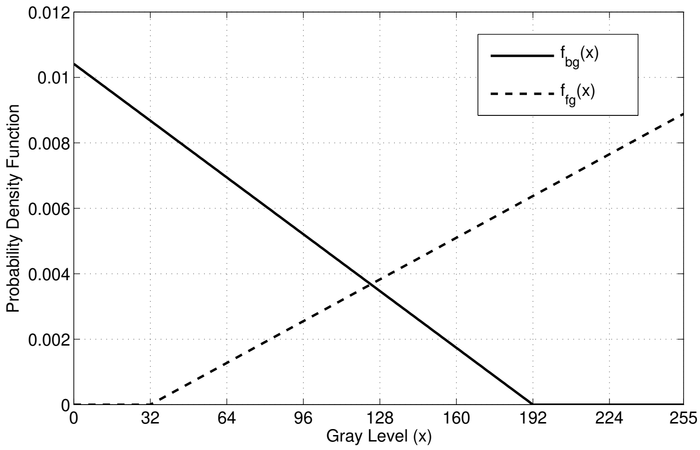


Figure 9: Gray level probability density functions of background and foreground regions in the image.

(a) (3p) Give the mathematical expressions for the background pixel probability density function $f_{bg}(x)$ and the foreground pixel probability density function $f_{fg}(x)$ by using Fig. 9.

(b) (4p) Optimal supervised thresholding with threshold T aims at minimizing the probability of misclassification, which is a sum of two misclassification probabilities. For example, consider $T = 96$ and mark the regions that represent the two misclassification probabilities. Explain these two terms shortly and give the corresponding expressions.

(c) (4p) We want to select the threshold T to minimize the probability of misclassification. What is the optimal segmentation threshold in the two limiting cases $P_{bg} = 0$ and $P_{fg} = 0$, respectively?

(d) (4p) For general probabilities P_{bg}, determine the expression to find the optimal threshold T that minimizes the probability of misclassification by using your expressions of $f_{bg}(x)$ and $f_{fg}(x)$. Give also the result for the special case $P_{bg} = P_{fg} = 1/2$.