DATE: 9. December 2019

University of Stavanger

EXAM IN: ELE510 IMAGE PROCESSING with COMPUTER VISION

DURATION: 4 hours

ALLOWED REMEDIES: Defined, simple calculator permitted.

THE SET OF EXERCISES CONSISTS OF 5 EXERCISES ON 4 PAGES

NOTES: All subproblems are weighted equally.
ANSWERS

Exercise 1

a) Explain how a digital image, $I(x, y)$, is represented in a computer. The answer should include the following: sampling, quantization, color image. Feel free to use drawings and/or mathematical expressions in your answer.

ANSWER: A gray scale digital image is represented as a matrix of numbers, where the number is a representation of the (gray scale) intensity level. Sampling corresponds to the fact that the continous analog world is "sampled" at a spatial grid (in the image sensor) giving us this matrix of numbers. The number of pixels is seen by the dimension of the matrix, i.e. a $N \times M$ image has $N * M$ pixels. The resolution can be seen as the number of pixels for a physical area. The analog continuous values at each point is quantized to a fixed number of different values. Often 8 bits per pixel (bpp) is used, giving us the intensity values from 0 (black) to 255 (white). For a color image, an image sensor typically sensors three different colors, red, blue and green (RBG), and the image in the computer is represented as 3 matrices similar to the one described above for grey level images, all of size $N \times M$, one for the red information, one for the ble and on for the green. Usually this will be represented together as a 3D matrix. A typical color image would have 8 bpp for the red channel, 8 for the blue and 8 for the green, i.e. 24 bpp in total.

b) Sketch a pinhole camera. Mark the focal length, the image plane, virtual image plane, and optical axis and focal point. Use your sketch and and explain how perspective projection forms an image.

ANSWER: See attached notes for sketch

Exercise 2

a) Let $I(x, y)$ denote a gray level image of size $N \times M$, and $I'(x, y)$ the corresponding image after a transformation. What is a geometric transformation? Set up the equation for a flop (mirror) operation. Can you briefly explain the difference between forward mapping and inverse mapping in this context?

ANSWER: Geometric transformation: The location of a pixel change from input to output, but the value does not change. Simple geometric: Each output pixel is dependent on a single input pixel, more complex: require interpolation. Top line forward mapping (use coordinates from input image), second line invers mapping (use coordinates from output image)

$$\begin{aligned} I\acute{}(M-1-x, y) &= I(x, y) && (1) \\ I\acute{}(x\acute{}, y\acute{}) &= I(M-1-x\acute{}, y\acute{}) && (2) \end{aligned}$$

b) Let I_b be an image.

$$I_b = \begin{bmatrix} 0 & 1 & 2 & 5 \\ 1 & 0 & 1 & 7 \\ 1 & 2 & 6 & 7 \\ 0 & 1 & 6 & 6 \end{bmatrix}, \tag{3}$$

find the *normalized histogram* and the *cumulative distribution function* (CDF) of the image.

ANSWER: The histogram $\mathbf{h}$ *as a vector for the indices* $0:7$, *has the following values (counts):*

$$\mathbf{h} = [3\ 5\ 2\ 0\ 0\ 1\ 3\ 2] \tag{4}$$

To normalize we need to divide all elements of $\mathbf{h}$ *with the total count, so that it sums to 1:*

$$\bar{\mathbf{h}} = \frac{1}{16}[3\ 5\ 2\ 0\ 0\ 1\ 3\ 2] = [0.19\ 0.31\ 0.125\ 0\ 0\ 0.06\ 0.19\ 0.125] \tag{5}$$

The CDF is found by the running sum of the normalized histogram:

$$\bar{\mathbf{c}} = \frac{1}{16}[3\ 8\ 10\ 10\ 10\ 11\ 14\ 16] \tag{6}$$

c) Explain the purpose of *histogram equalization*. Feel free to use figures as you explain.

Perform histogram equalization on I_b by help of the CDF from b), and show the output image.

ANSWER: Purpose: To make the probability distribution of the available grey levels (here 0-7, alternatively 0-255) as even as possible, i.e. to maximize the use of the bits we have available. This will typically enhance the image visually. In a continuous world all levels should be equally probable, but since it is discrete that will not be possible. The information should not be destroyed. I.e. if a pixel at position (x_A, y_A) is darker than pixel at position (x_B, y_B) prior to equalization, the pixel at (x_A, y_A) should still be darker (or equal to) the pixel at position (x_B, y_B) after equalization. to perform histogram equalization, we find the new pixel level l' as a function of the pixel level on the input l using the CDF:

$$l' = ROUND(MAX_{l'} \cdot \bar{\mathbf{c}}(l)) \tag{7}$$

Here $MAX_{l'} = 7$, corresponding to 3 bit per pixel (often it is 255). Upper row represent original pixel intensity levels, and lower row represent the new pixel values after histogram equalization:

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 4 & 4 & 4 & 5 & 6 & 7 \end{bmatrix} \tag{8}$$

$$I_{histeq} = \begin{bmatrix} 1 & 4 & 4 & 5 \\ 4 & 1 & 4 & 7 \\ 4 & 4 & 6 & 7 \\ 1 & 4 & 6 & 6 \end{bmatrix}, \tag{9}$$

d) *Frame differencing* is a simple method for detecting moving object in video frames. The simplest approach is to compute a difference image:

$$I'(x, y) = |I_t(x, y) - I_{t-1}(x, y)| > \tau \tag{10}$$

Make a simple sketch of 3 consecutive frames in a video with a moving ball on an even background. Illustrate the output of Eq. 10. What is the well-known problem of this approach? Propose another well-known approach to solve that problem, and sketch the output.

ANSWER: See attached notes for sketch. Well known problem: double image problem, you see the foreground object both at the position from the first and the second frame (I_t, and I_{t-1}). Illustrated in the sketch. Solution: double difference image (also called three-frame difference) or the triple difference image given by (first double, than triple):

$$I' = |I_t - I_{t-1}| > \tau \text{ AND } |I_{t+1} - I_t| > \tau \tag{11}$$

$$I' = (|I_{t-1} - I_t| + |I_{t+1} - I_t| - |I_{t-1} - I_{t+1}|) > \tau \tag{12}$$

Exercise 3

a) Let $T(\cdot)$ be a system (filter), and $I(x, y)$ an image; The output image can be written as $I'(x, y) = T(I(x, y))$. Define what we mean by: *linear* system, *shift invariant* system, FIR and IIR?

ANSWER: Let a, b be constant. For a linear system the scaling and the additivity property holds. Together these properties are referred to as superposition, and can be written as:

$$T(a \cdot I_1(x, y) + b \cdot I_2(x, y)) = a \cdot T(I_1(x, y)) + b \cdot T(I_2(x, y)) \tag{1}$$

A shift invariant system has the following property: If

$$I^{'}(x, y) = T(I(x, y)), \tag{2}$$

then the system is shift-invariant if and only if:

$$I^{'}(x - x_0, y - y_0) = T(I(x - x_0, y - y_0)) \tag{3}$$

b) 2D Convolution is shown below with notation from the book:

$$I'(x, y) = I(x, y) * G(x, y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(x + \tilde{w} - i, y + \tilde{h} - j) G(i, j) \tag{4}$$

where $\tilde{w}, \tilde{h}$ indicates the half width and height of a filter kernel, G.

Using the expressions from a), *when* can we use *convolution* to find $I'(x, y)$? Write up the kernel, G(i,j), of a 3×3 mean filter, and explain with a few sentences what it does.

ANSWER: Convolution can be used to describe the output from a system that is LINEAR and SHIFT-INVARIANT. A mean filter finds the mean of the pixels under the kernel-support. For a 3x3 filter kernel that means that the mean of all pixel in a 3x3 neighbourhood is calculated, and that will be the output of the pixel at the position corresponding to the pixel under the center of the kernel. When convolution is used, we can think of it as the kernel (a fliped version) is slided over the entire image, and the output pixel is found (at the center position of the pixels under the kernel) by convolving the image pixels under the kernel with the filterkernel.

$$G(i, j) = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{5}$$

c) Some 2D filter-kernels are separable. What is the advantage of separable filter kernels? Let the following describe a 2D convolution of an image $I(x, y)$ with a separable Gaussian filter:

$$I\,'(x, y) = I(x, y) * \left(\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \right) = .. \tag{6}$$

Write the decomposed version of the equation.

ANSWER: Computationally less expensive, can do 2 times 1D convolution instead (over the rows, and over the columns).

$$I\,'(x, y) = \left(I(x, y) * \frac{1}{4} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right) * \left(\frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \right) \tag{7}$$

d) Explain the purpose of the three filter-kernels below. Perform the convolution $I\,'_b(x, y) = I_b(x, y) * h_2$ (I_b from Eq. 11). Let $I\,'_b(x, y)$ have the same size as the input $I_b(x, y)$, and use zero padding.

$$\mathbf{h}_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} . \tag{8}$$

$$\mathbf{h}_2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} . \tag{9}$$

$$\mathbf{h}_3 = \frac{1}{64} \begin{bmatrix} 1 & 6 & 1 \\ 6 & 36 & 6 \\ 1 & 6 & 1 \end{bmatrix} . \tag{10}$$

ANSWER: filter kernels are doing:
$\mathbf{h}_1$ *is a highpassfilter (it sums to 0). It is actually the Laplacian (sum of second derivatives), or the Laplacian of Gaussian (LoG) with* $\sigma = 0$ *for the Gaussian (ie. no smoothing). The Laplacian corresponds to the sum of the horizontal and vertical second derivative kernels. Here it is represented with a negative sign compared to the book, but that is not important.*
$\mathbf{h}_2$ *is It is also highpass (sums to zero). But finds edges, it is the Sobel operator in y-direction. It smoothes over the rows, and find the derivative over the columns.*
$\mathbf{h}_3$ *is a lowpass kernel (sums to 1) It will smoothen the image. It is a Gaussian with* $\sigma^2 = 0.25$ *.*

The output after filtering $I_b(x,y)$ with $\mathbf{h}_2$:

$$I_b^{'} = \begin{bmatrix} 2 & 2 & 9 & 15 \\ 3 & 7 & 11 & 8 \\ -1 & 6 & 10 & 3 \\ -4 & -11 & -21 & -20 \end{bmatrix}, \tag{11}$$

e) Equation Eq. 12 shows the magnitude of an ideal 1D low pass filter in the frequency domain, with cutoff freq. of w_0.

$$|H(w)| = \begin{cases} 1 & \text{if} |w| < w_0 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Sketch roughly how such an ideal filter in frequency domain appear in the *spatial domain*. Use the formula for the invers discrete time Fourier Transform of a 1D signal to find a mathematic expression for the same:

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(w) e^{jwn} dw \tag{13}$$

Hint: Use Eulers formula (formula pages, F. 7) *ANSWER: See notes for sketch.*

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(w) e^{jwn} dw \tag{14}$$

$$h(n) = \frac{1}{2\pi} \int_{-w_0}^{w_0} 1 \cdot e^{jwn} dw \tag{15}$$

$$h(n) = \frac{1}{2\pi} \int_{-w_0}^{w_0} (\cos(wn) + j\sin(wn)) dw \tag{16}$$

$$h(n) = \frac{1}{2\pi} \left[\int_{-w_0}^{w_0} \cos(wn) dw + j \int_{-w_0}^{w_0} \sin(wn) dw \right] \tag{17}$$

$$\tag{18}$$

since $\sin()$ *is odd-symmetric, the imaginary part disapears (=0). We get:*

$$h(n) = \frac{1}{2\pi} \int_{-w_0}^{w_0} \cos(wn) dw \tag{19}$$

$$h(n) = \frac{1}{2\pi} \left[\frac{sin(wn)}{n} \right]_{-w_0}^{w_0} \tag{20}$$

$$h(n) = \frac{\sin(w_0 n)}{\pi n} = \frac{1}{w_0 \pi} \textit{sinc}(w_0 n) \tag{21}$$

This corresponds with the sketch.

Exercise 4

a) A small (7x7) image is depicted in Figure 1 with pixel intensity values. The circle is roughly pointing out a foreground object, this is not a part of the image itself. We wish to do segmentation to get the foreground object. Explain thresholding by using this image as example; Let us use p-tile method for finding the threshold, and from the size of the circle we can say that the object should be approximately 12 pixels. Sketch the output image. Can you mention other ways of finding the threshold?

0	1	0	1	1	2	0
1	1	0	3	4	1	0
0	2	4	5	4	0	1
0	0	4	5	4	1	1
1	0	3	3	2	0	2
0	1	2	0	1	0	5
1	0	1	0	0	3	4

Figure 1: Figure to problem 4 a) and b)

ANSWER First, the histogram for pixel values $0:7$*:*

$$\mathbf{h} = [17\ 14\ 5\ 4\ 6\ 3\ 0\ 0] \tag{1}$$

p-tile method: find the thershold so that the right proportion of pixels are in the foreground. Here we want approximately 12 pixels in the foreground. Look at the histogram, and see if we set $T = 3$*, we get* $4+6+3 = 13$ *pixels in the foreground. However, in addition to the desired foureground object we also get the lower right corner.*

b) Let us look at the image from Figure 1 again and try segmentation by region growing instead. Let the seed point be column=4, row=3, (intensity level = 5). Let our feature, $f(x, y)$, be the pixel intensity value, and the similarity measure be defined $d = |f(x, y) - \bar{f}|$, where $\bar{f}$ denotes the mean of the feature over the region. Include a point in the region if $d < 2$, and use a 4NB (neighbourhood). Go through entire boundary before $\bar{f}$ is updated (as done in class). Do the segmentation, Show the development of $\bar{f}$ in each round, and sketch the final output image. Compare with a) and comment.
ANSWER See attached notes

c) The Harris Stephens corner detector is a classical algorithm for finding corners, or feature points, in an image. Can you explain minimum two applications where finding such feature points are is necessary or useful? Mention some *properties* that are good for a feature point detector.

ANSWER: Some possible answers (there are more things that this can be used for) : 1) Useful for finding known objects in an image, by firstly have a typical feature point representation of the object. When looking for the object in other images, the feature point pattern can be search for, for example by feeding a classifier. 2) Useful for image mosaicing where we are stitching together different images with some overlapping parts (panorama images made by a phone, for example) 3) useful for finding corresponding points in stereo vision. Good properties: rotation invariant, translation invariant, scale invariant.

Exercise 5

A camera has a focal length of $f = 30mm$ and the image sensor have 3000 x 6000 pixels covering an area of 10x20 mm. The camera has zero skew. The camera is facing a wall, perpendicular to the optical axis, at a distance of 2 meters measured along the optical axis from the focal point. Let $\{X_c, Y_c, Z_c\}$ correspond to the camera coordinate system.

a) What is the Field of View (FOV) in the X_c and Y_c direction (i.e. given by angles in the $Z_c - X_c$ and $Z_c - Y_c$ planes)? How large area on the wall (height and width in m) is covered by the image?

ANSWER $w_x = 20mm$, $w_y = 10$*mm*, $f = 30mm$, $l = 2m$.

$$FOV_x = 2 \cdot \mathit{Arctan}\left(\frac{w_x}{2f}\right) = 2 \cdot \mathit{Arctan}\frac{1}{3} \tag{1}$$

$$FOV_y = 2 \cdot \mathit{Arctan}\left(\frac{w_y}{2f}\right) = 2 \cdot \mathit{Arctan}\frac{1}{6} \tag{2}$$

Area on wall:

$$W_x = l\frac{w_x}{f} = l\frac{2}{3} = \frac{4}{3}m \tag{3}$$

$$W_y = l\frac{w_y}{f} = l\frac{1}{3} = \frac{2}{3}m \tag{4}$$

Area on the wall: $\frac{4}{3} \cdot \frac{2}{3} = \frac{8}{9}m^2$

b) Find the internal camera calibration matrix.

ANSWER: No skew, simplifies:

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

$f_x = \frac{f}{\Delta x}$, $\Delta x = \frac{20*10^{-3}}{6000} = 3.33 * 10^{-6}$ *m/pixel*, $f_x = 9000$
$f_y = \frac{f}{\Delta y}$, $\Delta y = \frac{10*10^{-3}}{3000} = 3.33 * 10^{-6}$ *m/pixel*, $f_y = 9000$
$x_0 = 6000/2 = 3000$, $y_0 = 3000/2 = 1500$.

c) Assume you have two cameras of the same type, and you will use them for stereo imaging. Let the cameras be *rectified*, i.e. the cameras are translated to each other parallell to the scan lines (rows). Explain what we mean by corresponding points in the images, and how these can be used to find the depth, i.e. distance, to objects in the image scene. What are epipolar lines in the context of stereo imaging. (brief explanations)

ANSWER For rectified images, corresponding points have to lie on the same scan-line (image - row). Corresponding points are the same physical point, seen be the two different cameras. The DISPARITY, i.e. the difference between these two points, can be used to find the depth : $d = x_L - x_R = \frac{fb}{z_w}$*, where* x_L, x_R *are the image coordinates,* f *is the focal length,* b *is the baseline, ie difference of the focal points between the two cameras (measured along the scan line) .* z_w *is the depth of the object, or distance in the z-direction from the focal points. Even if the cameras are NOT rectified, we can still find a line where the corresponding points must lie in the other image. This is an advantage so we do not need to search the entire image for corresponding points. The lines are called epipolar lines, andto find them we need to know the gemoetry of the placement of the cameras relative to each other.*

d) Let the cameras be rectified, and let the maximum disparity be 20. Which of the following pixel coordinates (x,y) in the right image could not match the pixel (52,3) in the left image, and why : i) (26,3) ii) (48,13), iii) (64,3) iv) (48,3) v) (59,6)

ANSWER:
i) exceeds maximum disparity of 20 (i.e. 52-26 = 26 > 20),
ii) wrong row (13 not 3)
iii) not possible because of the cheirality constraint (i.e 64>52. needs to be smaller),
iv) (48,3) is a possible match (The ONLY possible match of the ones listed in the question)
v) wrong row (6 not 3)

Formulas

Gaussian function

$$Gauss_{\sigma^2}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{\left(-\frac{x^2+y^2}{2\sigma^2}\right)} \tag{6}$$

Eulers formula

$$e^{jwn} = cos(wn) + jsin(wn) \tag{7}$$

SVD decomposition

$$f = U\Lambda^{\frac{1}{2}}V^T \tag{8}$$

Discrete Fourier transform (DFT) and the invers DFT:

$$\hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{9}$$

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{10}$$

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_y \sum_x f(x,y)h(\alpha - x, \beta - y) \tag{11}$$

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y) \cdot r(x,y) \tag{12}$$

Between class variance:

$$\sigma_B^2(t) = \frac{[\mu(t) - \bar{\mu}\theta(t)]^2}{\theta(t)(1-\theta(t))} \tag{13}$$

LoG function:

$$LoG = [\frac{x^2+y^2-2\sigma^2}{\sigma^4}]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{14}$$

Minimum errror threshold:

$$\theta p_o(t) = (1-\theta)p_b(t) \tag{15}$$

Harris Stephens corner detection:

$$\mathcal{H} = \sum_{window} \{(\nabla I)(\nabla I)^T\} \tag{16}$$

$$= \sum_{window} \begin{bmatrix} I_xI_x & I_xI_y \\ I_yI_x & I_yI_y \end{bmatrix}. \tag{17}$$

$$R = \det(\mathcal{H}) - k(\frac{\mathsf{trace}(\mathcal{H})}{2})^2. \tag{18}$$

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{19}$$

$$\lambda\mathbf{p} = \mathcal{K}\Pi_0\mathbf{TR}\,{}^{W}\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{20}$$

Here $\mathbf{p} = [x\ y\ 1]^T$ is the image coordinates in number of pixels and ${}^{W}\mathbf{P} = [X\ Y\ Z\ 1]^T$ the world coordinates in meter.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{21}$$

where $\alpha = kf = \frac{f}{\Delta x}$ and $\beta = lf = \frac{f}{\Delta y}$.

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{22}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{23}$$

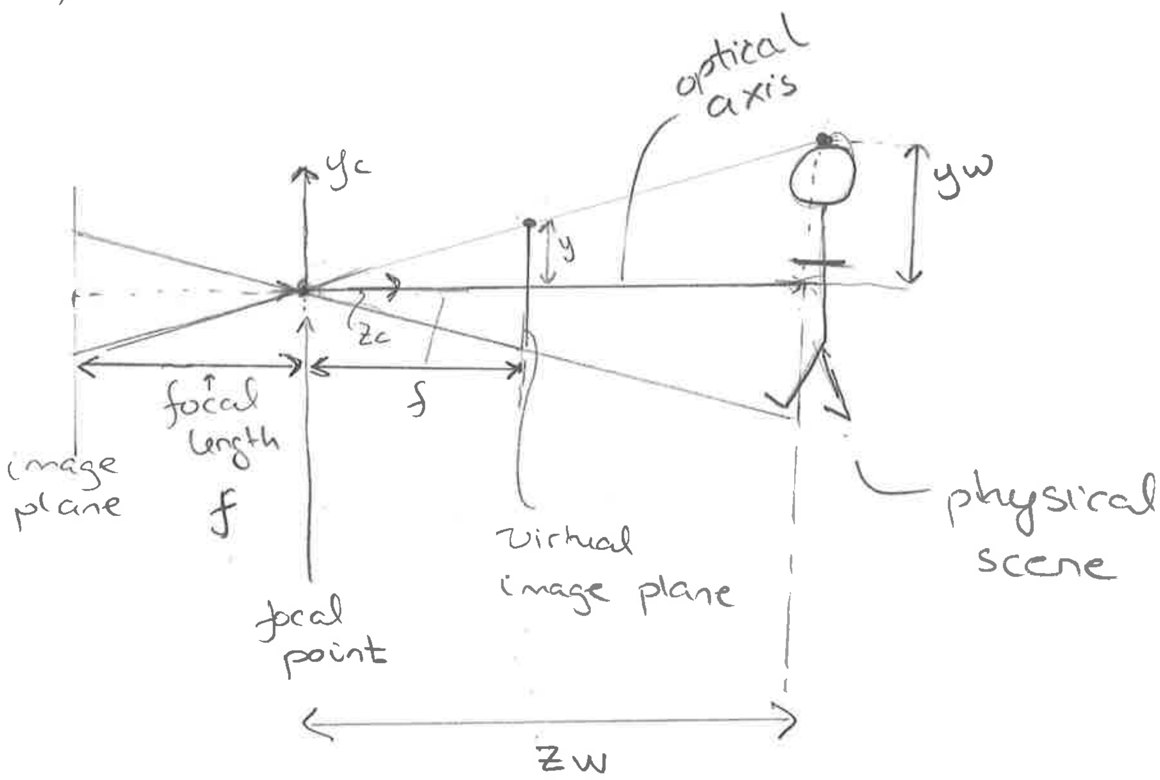
$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{24}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{25}$$

ELE 510 Image processing with computer vision

Exam 09.12.2019 notes, attachement to solution

1 b) pinhole camera



perspective projection:

light reflect from the physical surface and travel through the focal point to the image plane. We look at the virtual image plane for simplicity (to not have to think upside-down)

see that $\frac{y}{f} = \frac{y_w}{z_w}$ since the shape of the triangles are the same

$$\Rightarrow y = f \cdot \frac{y_w}{z_w}$$

and

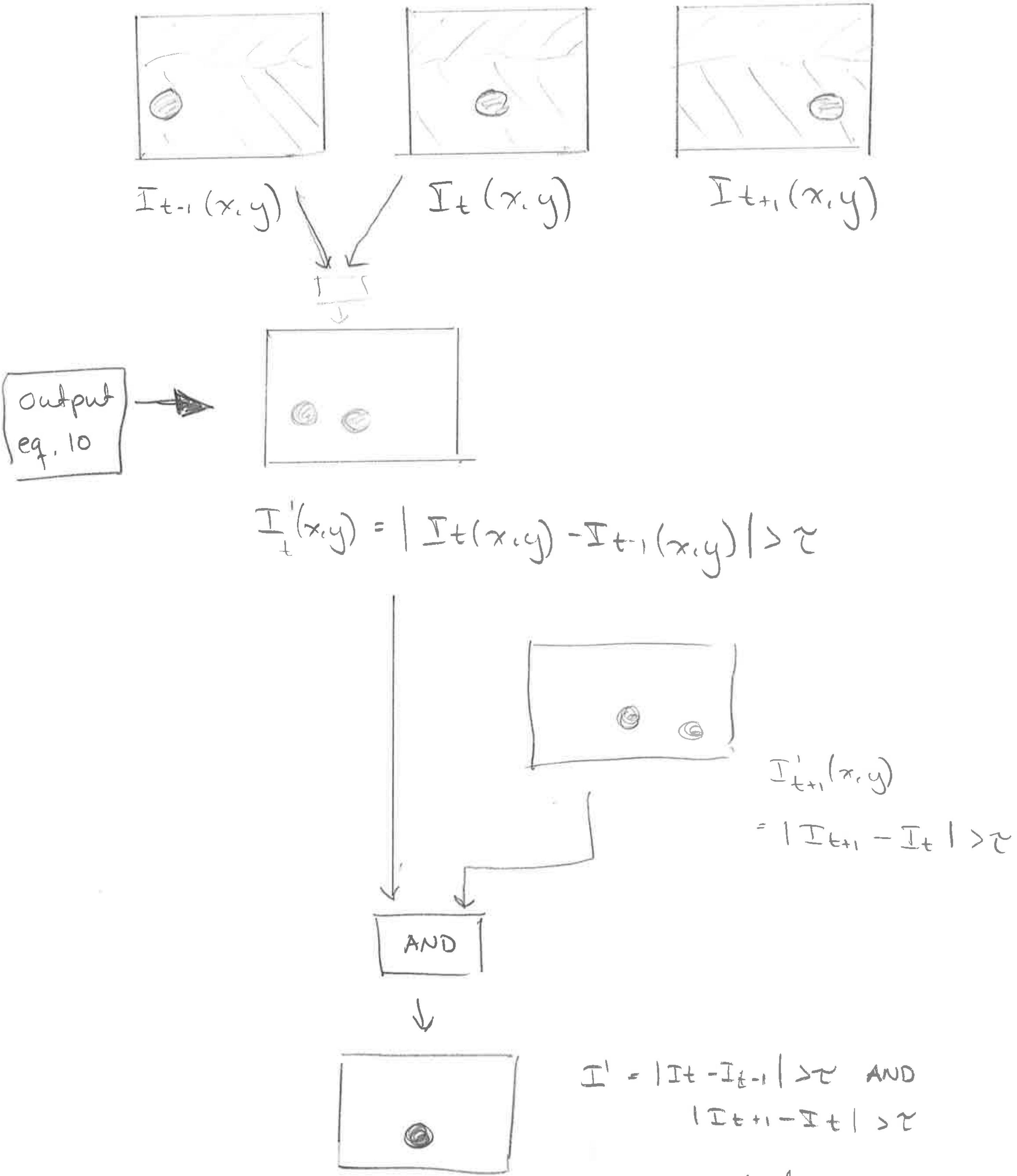
$$x = f \cdot \frac{x_w}{z_w}$$

x_w, y_w, z_w is in world coordinates

x, y is on the virtual image plane, camera coord. system.

ELE 510, exam 09.12.2019, notes, attachment to solution

2 d)



double difference image picks out foreground object at time t.

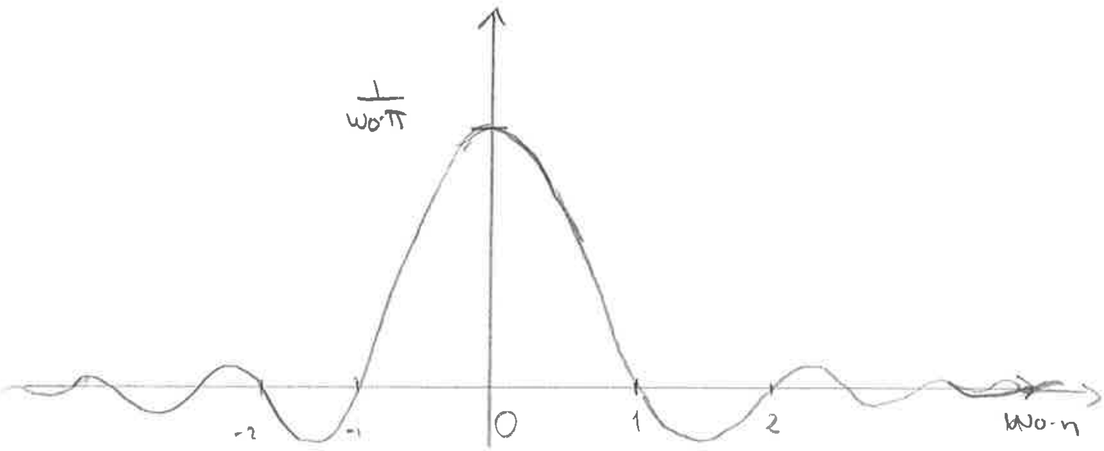
ELE 510 exam 09.12.2019, Notes attachement to solution

3 e)

ideal 1D low pass filter in freq. domain

$$|H(\omega)| = \begin{cases} 1 & |\omega| < \omega_0 \\ 0 & \text{otherswise} \end{cases}$$

sketch how this filter appear in the spatial (or time) domain:



$$h(n) = \frac{\sin(w_0 \cdot n)}{\pi \cdot n} = \frac{1}{w_0 \cdot \pi} \operatorname{sinc}(w_0 \cdot n)$$

Ex 4 b) ELE 510 exam dec. 2019

$$\begin{array}{ccccccc}
0 & 1 & 0 & 1 & 1 & 2 & 0 \\
1 & 1 & 0 & 3 & 4 & 1 & 0 \\
0 & 2 & 4 & 5 & 4 & 0 & 1 \\
0 & 0 & 4 & 5 & 4 & 1 & 1 \\
1 & 0 & 3 & 3 & 2 & 0 & 2 \\
0 & 1 & 2 & 0 & 1 & 0 & 5 \\
1 & 0 & 1 & 0 & 0 & 3 & 4
\end{array}$$

$d = |f(x,y) - \bar{f}|$

use 4NB , include if $d < 2$

1) $\boxed{5}$ look at $\begin{array}{ccc} & 3 & \\ 4 & \boxed{5} & 4 \\ & 5 & \end{array}$ $d < 2$? $\Rightarrow$ $\begin{array}{ccc} 4 & 5 & 4 \\ & 5 & \end{array}$

$\bar{f} = 5$

2) new $\bar{f} = \dfrac{5 + (4+4+5)}{1+3} = \dfrac{18}{4} = 4.5$ look at: $\begin{array}{ccccc} & 0 & 3 & 4 & \\ 2 & 4 & 5 & 4 & 0 \\ & 4 & 5 & 4 & \\ & & 3 & & \end{array}$ $\Rightarrow$

3) $\begin{array}{ccc} & 3 & 4 \\ 4 & 5 & 4 \\ 4 & 5 & 4 \\ & 3 & \end{array}$ new $\bar{f} = \dfrac{18 + 3+4+4+4+3}{4+5} = \dfrac{36}{9}$ $\Rightarrow$

4) $\begin{array}{ccc} & 3 & 4 \\ 4 & 5 & 4 \\ 4 & 5 & 4 \\ 3 & 3 & \end{array}$ new $\bar{f} = \dfrac{36+3}{9+1} = \dfrac{39}{10} = 3.9$ $\Rightarrow$

5) $\begin{array}{cccc} & & 3 & 4 \\ 2 & 4 & 5 & 4 \\ & 4 & 5 & 4 \\ & 3 & 3 & 2 \\ & 2 & & \end{array}$ new $\bar{f} = \dfrac{39 + 2+2+2}{10+3} = \dfrac{45}{13} = 3.46$

6) no new

final segmented area

We include the 2's (diff. between result after it 4) and 5)) which we did not do in a), but we do not get the lower left corner. Seems more correct compared to threshholding.