NTNU – Trondheim
Norwegian University of
Science and Technology

Department of Computer and Information Science

Examination paper for TDT4265 Computer Vision

Academic contact during examination: Christian Schellewald
Phone: 94123453

Examination date: 26.5.2015
Examination time: 09:00-13:00
Permitted examination support material: code D

No printed or hand-written support material is allowed.
Specified simple calculator permitted.

Other information:

Time allowed: 4 hours = 240 minutes.

The point scores for each question part are shown in the text. Use these point scores to help you budget your time and effort in answering.

READ EACH QUESTION THROUGH CAREFULLY BEFORE BEGINNING YOUR ANSWER.
ANSWER ALL QUESTIONS. SHOW ALL YOUR WORKING.

Total point score: 70

Language: English
Number of pages: 6
Number of pages enclosed: 6

Checked by:

Date Signature

1 Theme: Short Questions

1.1 Short Questions

1.1.1 Each answer to a short question is worth maximum two points.

(a) [2 points] How many and which basic image processing processes are performed by a **Laplacian of Gaussian** (LoG) filter.

(b) [2 points] How can we decide if a two-dimensional filtering kernel/mask is separable?

(c) [2 points] What is the core idea behind the K-Nearest-Neighbor-Classifier (KNN-Classifier)?

(d) [2 points] What is the **disparity** in a rectified stereo image pair?

(e) [2 points] Name four factors that make tasks in Computer Vision difficult.

2 Theme: Image Processing

2.1 Filtering

2.1.1 You plan to apply the normalized one dimensional box-filter (of size 1x2) several times.

(a) [2 points] What kind of filter is the box-filter? Give the box-filter-mask (of size 1x2) along with its normalization factor.

(b) [2 points] Determine the filter that can perform four (=4) box-filterings (of size 1x2) directly? Which filter is approximated?

2.1.2 A filter can be separated into the following two filter masks:
The filter masks are $(1\ 2\ 1)$ and $(1\ \ -2\ \ 1)^T$ [Note, the first is a row-vector and the second is a column vector.]

(a) [2 points] How does the original filter look like? (Give the filter-mask)

2.1.3 [2 points] When is a filter called a nonlinear filter?

2.1.4 [2 points] What is the main idea behind the nonlinear diffusion filter (Perona/Malik: Anisotropic diffusion)? What is better preserved in the images that are filtered by this filter compared to images filtered by a Gaussian filter.

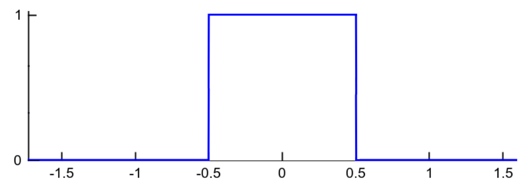
2.2 Convolution with a unit impulse.

2.2.5 Determine the convolution of a 1D rectangular function $f(x)$ with a unit impulse $g(x) = \delta_{x_o}(x) = \delta(x - x_0)$ at location $x_0 = 7.5$. A unit impulse in the continuous case (Dirac delta distribution) is defined as:

$$\delta(x) = \begin{cases} \infty & \text{for } x = 0 \\ 0 & \text{otherwise.} \end{cases} \quad \text{where} \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

The rectangular function is:.

$$f(x) = \begin{cases} 0 & \text{if } |x| > 0.5 \\ 1 & \text{if } |x| \leq 0.5 \end{cases}$$



(a) [3 points] Draw a sketch of the function $h(x) = (f * g)(x)$ which is the result of the convolution of f and g where g is the unit impulse at location $x_0 = 7.5$ as defined above. There is no need to compute the convolution explicitly, but indicate important coordinates and the length of function-segments in your sketch.

(b) [2 points] Give a discrete filter mask of a unit impulse. (Use a 1x5 filter-mask)

2.3 Distance functions

2.3.1 Determine the distance of two-dimensional points.

(a) [3 points] Give the Absolute-, Euclidean-, and Maximum- distance of the following 2 two-dimensional points A and B: $A = (7, 5)$; $B = (10, 9)$

(b) [2 points] Which distance corresponds to the Manhattan-distance and which to the Chessboard-distance?

3 Theme: Mathematical Morphology

3.1 Morphological Operations

3.1.1 Assume that the structuring element b is given and a gray-value image is denoted f. Then gray-scale dilation of the image can be expressed as $f \oplus b$ and erosion as $f \ominus b$.

3.1.2 [3 points] Use both operations erosion and dilation to define a **morphological gradient**. Give a brief explanation of your found expression by explaining what is actually computed at a single image pixel.

3.1.3 [2 points] Which of the following operations are **not idempotent** ? Erosion ($\ominus$), dilation ($\oplus$), closing ($\bullet$), opening ($\circ$). (Answer in a complete sentence e.g.: X,Y and Z are not idempotent.)

4 Theme: Filtering in the Frequency Domain

4.1 Fourier transform

4.1.1 [3 points] Consider the 1D rectangular function shown below.

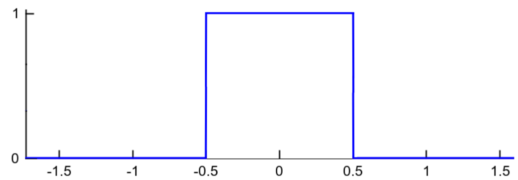


Figure 1: A 1-dimensional rectangular function

The function is defined by the equation below:

$$f(x) = \begin{cases} 0 & \text{if } |x| > 0.5 \\ 1 & \text{if } |x| \leq 0.5 \end{cases}$$

Calculate the continuous 1D Fourier transform of this rectangular function.
Reminder: The 1D Fourier transform is defined by the integral:

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi ux} \, \mathrm{d}x \tag{1}$$

You might find Euler's formula useful:

$$e^{ix} = \cos x + i \sin x \tag{2}$$

4.1.2 [2 points] Is the Fourier-transform invariant against rotation? To answer this, you might consider the Fourier transform of two versions of a rotated pattern like the ones shown in figure 2. Are the amplitude images of their Fourier transform $|\mathcal{F}\{f\}|$ different ? (Answer this question in a complete sentence !)

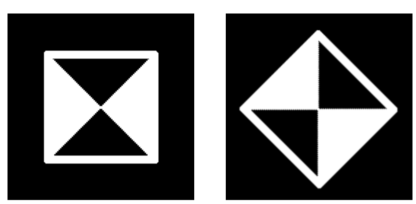


Figure 2: Rotated pattern. Is the Fourier-transform invariant against rotation?

5 Theme: Algorithms

5.1 Clustering: K-Means

5.1.1 Explain the k-Means Clustering approach by answering the following questions.

(a) [3 points] Describe an algorithm that computes the k-Means-partition on a point set.

(b) [3 points] Which are the two main-problems of the k-Means approach? Give suggestions to cope with these problems.

(c) [2 points] Which properties define a single k-Means-partition?

(d) [2 points] When does k-Means work reasonably well, when is it likely to fail ?

6 Theme: Image Structure and Shape Description

6.1 Curvature Scale Space

6.1.1 Use the following questions as a guide to explain how the **Curvature Scale Space image** (CSS-image) of a contour can be created and how it can be used to describe the shape of the contour. Assume that you have extracted the consecutive x- and y- coordinates of the N contour-points by tracing around a shape. (You do not need to recall details on the formulas!)

(a) [2 points] If needed, which pre-processing should you perform on the raw contour data? Why?

(b) [2 points] Which points of the contour are of special interest ?

(c) [2 points] Describe the iterative process that builds the CSS-image of the contour.

(d) [2 points] When can you stop iterating?

(e) [2 points] How can the CSS-image be used to describe a contour?

7 Theme: Hierarchical Structures / Computer Vision Tasks

7.1 KD-trees

7.1.1 Tree structures turn out to be an efficient way to organize data in computer vision. One of them is the KD-tree. Assume that the split-dimension for the root node is the x-coordinate and changes with the depth of the nodes.

(a) [3 points] Insert the following 2D-Points into the KD-tree structure (in alphabetic order) and draw the resulting tree: $A = (3, 4)$ $B = (1, 6)$ $C = (2, 1)$ $D = (5, 2)$ $E = (7, 1)$

(b) [2 points] For which kind of data would a KD-tree usually be used? Which alternative tree-structure would you suggest to use for 2D data?

7.2 Real World Computer Vision Task

7.2.1 [5 points] Describe the steps you would use to determine the locations of the knots (crossing points) in a fishing net within images like the ones shown in Figure 3. How would you cope with the fact that you could be far or close to the net when the image is recorded? Note that this is a challenging real world task, so make guesses when your proposed method or parts ot it might fail. Start at the pre-processing stage. (Level of detail: Try to divide your method-description into 5 to 10 steps and indicate which aims the steps pursue.)

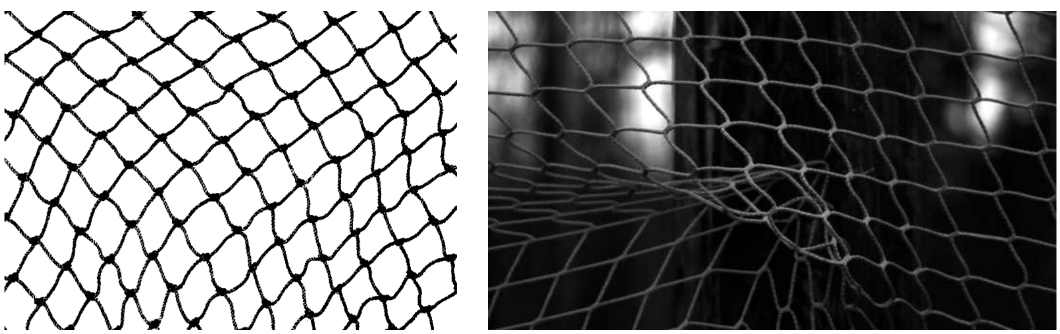


Figure 3: Fishnet example images.

———————- Congratulations this is the last page of the exam ! ———————————