Institut für Mess- und Regelungstechnik
mit Maschinenlaboratorium
Karlsruher Institut für Technologie
(KIT)
Prof. Dr.-Ing. C. Stiller

Solutions for exam „Machine Vision“ October 14$^{\text{th}}$, 2016

Contents

Question 1 (6 points)

Calculate the Fourier transform of the following function g:

$$g(x) = \begin{cases} 2 & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Solution

$$\begin{aligned} \hat{g}(k) &= \int_{-\infty}^{\infty} g(x) e^{-2\pi ikx} dx = \int_{-1}^{+1} 2e^{-2\pi ikx} dx = \left[\frac{e^{-2\pi ikx}}{-\pi ik} \right]_{-1}^{+1} \\ &= \frac{e^{-2\pi ik} - e^{2\pi ik}}{-\pi ik} = \frac{2\sin(2\pi k)}{\pi k} \end{aligned}$$

Question 2 (6 points)

Create a MATLAB function that calculates a gray level histogram from an image. It should take as input an $n \times m$ matrix of integers between 1 and 100. The function should return as output a vector `hist` that specifies the count for each gray value, i.e. `hist(i)` should be the count for gray value `i`. The signature of the function should be:

`function [hist] = gray_level_histogram(image)`

Do not use the MATLAB built-in functions `hist` and `histogram`.

Remark: You might find the following MATLAB functions useful for this task

- `A = zeros(n,m)` creates an $n \times m$ matrix filled with zeros

Solution

```
function [hist] = gray_level_histogram(image)
   [n, m] = size(image);
   hist = zeros (1, 100);
   for v=1:n
     for u=1:m
       hist(image(v,u)) = hist(image(v,u))+1;
     end
   end
end
```

Question 3 (6 points)

Create a filter mask h such that $\frac{\partial^2 g}{\partial u \partial v} \approx g * h$ for any gray value image g.

Solution

We can create h as $h = f * f^T$ where f is a filter with which we can calculate the horizontal partial derivative, e.g. $f = \frac{1}{2} \cdot (1,0,-1)$. For this choice of f we obtain

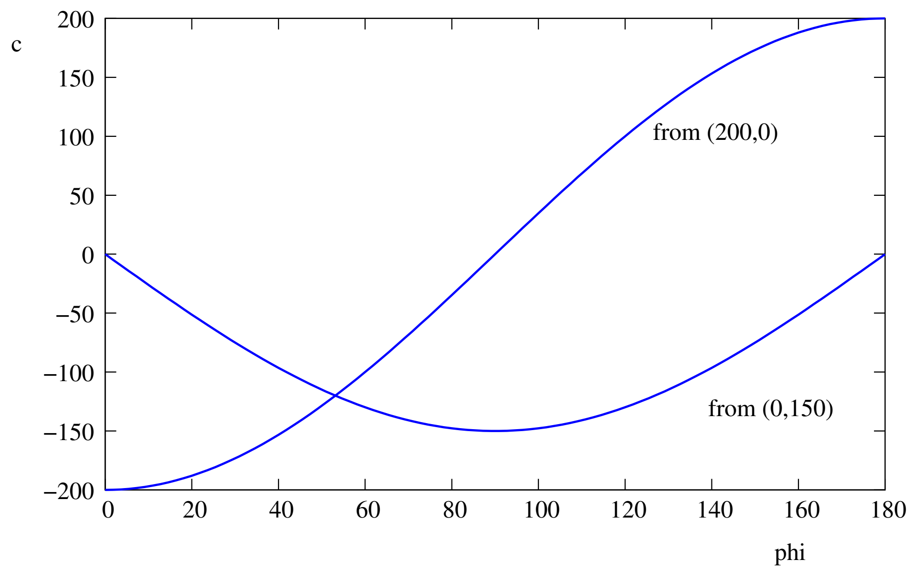
$$h = \frac{1}{2}(1,0,-1) * \frac{1}{2}(1,0,-1)^T = \frac{1}{4}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Question 4 (6 points)

Draw the curves in parameter space that are created by the Hough transform for the points $(x_1,y_1) = (200,0)$ and $(x_2,y_2) = (0,150)$. Do not forget to annotate the axes.

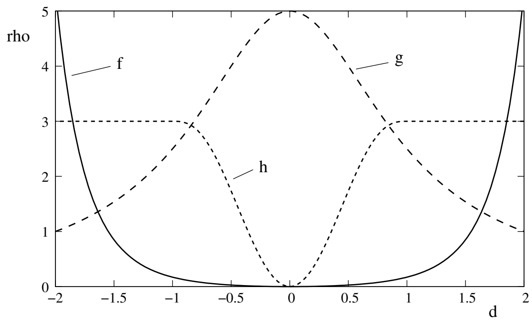
Solution

(x_1,y_1) creates the curve $c(\phi) = -200 \cdot \cos\phi$ while (x_2,y_2) creates the curve $c(\phi) = -150 \cdot \sin\phi$.



Question 5 (6 points)

Which of the three functions f, g, and h provided below can be used as error functions $\rho(d)$ of an M-estimator? The diagram on the right shows a plot of the three functions. Explain your answers briefly.



$$f(d) = \frac{1}{10}(e^{d^2} - 1)$$

$$g(d) = \frac{5}{d^2 + 1}$$

$$h(d) = \begin{cases} 3 \cdot (1 - (1 - d^2)^3) & \text{if } |d| \leq 1 \\ 3 & \text{if } |d| > 1 \end{cases}$$

Solution

f increases much faster than d^2. Therefore, it increases the influence of outliers on the estimation instead of reducing it. It is not a suitable choice.

g is a function that is decreasing for increasing d. Hence it cannot be used as ρ-function.

h is monotonically increasing for positive d and takes its minimum at 0. It is continuous, differentiable, and its derivative is bounded. Therefore, it can be used as ρ-function.

Question 6 (8 points)

The figure below shows a gray level picture with 5 rows and 6 columns. Apply the connected components labeling algorithm (CCL) to segment the image. Two gray values should be treated as similar only if they are equal, i.e. gray pixels and white pixels in the figure below are always dissimilar. Visualize the processing of CCL by inserting into each cell in the figure one of the following letters to indicate the processing step:

N if CCL initializes a new segment at this pixel

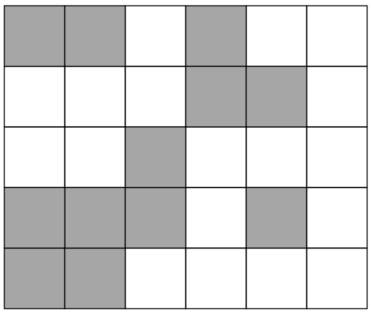
M if CCL merges two segments when passing this pixel

L if CCL assigns the pixel to the segment of its left neighbor, but not to the segment of its upper neighbor

U if CCL assigns the pixel to the segment of its upper neighbor, but not to the segment of its left neighbor

B if CCL assigns the pixel to the segment of its left neighbor and the left and upper neighbor already belong to the same segment

How many segments are finally found?



Solution

N	L	N	N	N	L
N	L	M	U	L	U
U	B	N	N	L	M
N	L	M	U	N	U
U	B	N	M	L	B

Finally, 6 segments are found.

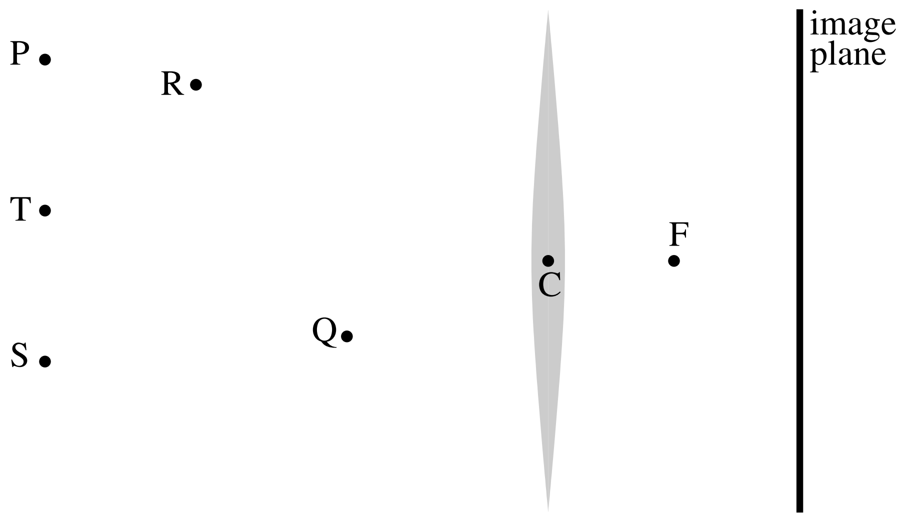
Question 7 (8 points)

The drawing below shows a camera with thin lens and an image plane. C denotes the center of the lens, F the focal point of the lens.

(a) Assume, the pinhole camera model applies to the drawn camera. In which order (from top to bottom) are the points P, Q, R, S, and T mapped onto the image plane? Explain your answer briefly in textual form or by providing a drawing. You might use the drawing on the questionnaire (i.e. the drawing below) for that purpose.

(b) Assume, the drawn camera is a telecentric camera. In which order (from top to bottom) are the points P, Q, R, S, and T mapped onto the image plane? Explain your answer briefly in textual form or by providing a drawing. You might use the drawing on the questionnaire (i.e. the drawing below) for that purpose.

(c) Explain why the lens of a telecentric camera must be at least as large as the object that should be observed with the camera.

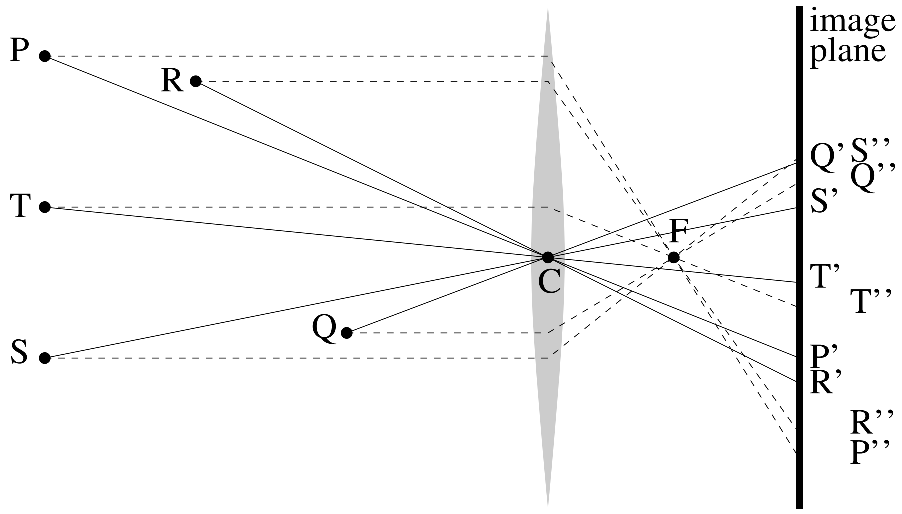


Solution

(a) The mapping of a scene point in a pinhole camera can be described by straight lines that pass through the scene point and point C, the center of the lens. The solid lines in the drawing below illustrate these lines for the points P-T and the points P'-T' denote the image points. Hence, the order of points on the image plane is Q, S, T, P, R.

(b) The mapping in a telecentric lens can be described by lines parallel to the optical axis that are refracted at the lens, and pass through the focal point of the lens. The dashed lines in the drawing below illustrate these lines for the points P-T and the points P"-T" denote the image points. Hence, the order of points on the image plane is S, Q, T, R, P.

(c) A telecentric lens filters out all lines of sight that are not parallel to the optical axis. Therefore, points are not mapped into the image if they are farer away from the optical axis than the radius of the lens.



Question 8 (6 points)

Assume, we trained four different classifiers c_1, c_2, c_3, and c_4 on the same training set and evaluated all of them on the same test set. The table below provides the classification accuracy on the training and test set.

	c_1	c_2	c_3	c_4
training set accuracy	82%	89%	86%	98%
test set accuracy	82%	71%	87%	65%

Which classifier performs best? Justify your answer and discuss which of the classifiers suffer from overfitting and/or from underfitting.

Solution

Relevant for the quality of a classifier is only its performance on the test set. Hence, classifier c_3 is the best one. For c_2 and c_4 the performance on the test set is much lower than the performance on the training set. This means, that those classifiers are overfitted. Classifier c_1 shows the same performance on the training and the test set. Therefore, it is not overfitted. However, we can conclude that it suffers from underfitting since it is clearly outperformed by c_3.

Question 9 (8 points)

Which among the following properties are necessary for a kernel function K that is used in a support vector machine and which properties are not necessary? Justify your answers briefly.

(a) K must be symmetric, i.e. $K(x,y) = K(y,x)$ for all values of x and y

(b) $K(x,y)$ must always be nonnegative

(c) K must be differentiable

(d) $K(x,x) \geq K(x,y)$ must hold for any value of x and y

Solution

(a) This property is necessary since the kernel is based on the dot product in feature space which is symmetric

(b) The dot product itself is a counterexample, e.g. $\langle 3, -1 \rangle < 0$

(c) The histogram intersection kernel is a counterexample.

(d) This property does not hold. E.g. $\langle 1, 1 \rangle \not\geq \langle 1,2 \rangle$

Gesamtpunkte: 60