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

Solutions for exam "Machine Vision" February 28, 2017

Question 1 (7 points)

Let $f : \mathbb{R} \to \mathbb{R}$ be a function and $g(t) = f(\alpha \cdot t)$ with $\alpha \in \mathbb{R}$, $\alpha > 0$. Prove the following property of the Fourier transform:

$$\hat{g}(k) = \frac{1}{\alpha} \cdot \hat{f}(\frac{k}{\alpha})$$

Note, that $\hat{f}$ denotes the Fourier transform of the function f and $\hat{g}$ the Fourier transform of g.

Solution

let $u = \alpha \cdot t$

$$\hat{g}(k) = \int_{-\infty}^{\infty} f(\alpha \cdot t) \cdot e^{-2\pi i k t} dt = \int_{-\alpha\infty}^{\alpha\infty} f(u) \cdot e^{-2\pi i k u/\alpha} d\frac{u}{\alpha}$$

since $\alpha > 0$

$$\int_{-\alpha\infty}^{\alpha\infty} f(u) \cdot e^{-2\pi i k u/\alpha} d\frac{u}{\alpha} = \frac{1}{\alpha} \int_{-\infty}^{\infty} f(u) \cdot e^{-2\pi i k u/\alpha} du = \frac{1}{\alpha} \cdot \hat{f}(\frac{k}{\alpha})$$

Question 2 (6 points)

The figure on the right depicts a small gray level image. The numbers represent the gray value of the respective pixel. Apply the median filter to the image. The median filter should consider a block of 3×3 pixels to calculate the median. Leave the boundary pixels blank.

1	1	4	6	0	7
1	2	3	5	4	5
0	1	7	6	4	3
1	0	0	2	2	1

Solution

	1	4	4	5	
	1	2	4	4	

Question 3 (8 points)

(a) Which function $c(\phi)$ in parameter space is created by the Hough transform for the point $(x,y) = (0,0)$?

(b) For which image point (x,y) does the Hough transform generate the curve $c(\phi) = -100 \cdot \sqrt{2} \cdot \sin(\phi + \frac{\pi}{4})$ in parameter space?

Remark: $\sin(\frac{\pi}{4}) = \sin(\frac{3\pi}{4}) = \frac{1}{2}\sqrt{2}$

Remark: you might apply appropriate values for ϕ to solve the task

Remark: The subtasks can be solved independently.

Solution

(a) By inserting the point coordinates $(x,y) = (0,0)$ into the line equation we obtain

$$0 \cdot \cos\phi + 0 \cdot \sin\phi + c = 0$$

It follows

$$c(\phi) = 0$$

(b) We create the two lines for $\phi = 0$ and $\phi = \frac{\pi}{2}$. The requested point is the point of intersection of these two lines.

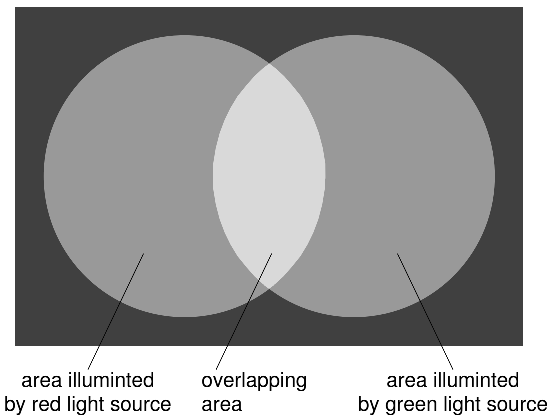
$$\phi = 0 \quad \Rightarrow \quad c(0) = -100 \quad \Rightarrow \quad x \cdot \cos(0) + y \cdot \sin(0) - 100 = 0 \quad \Rightarrow \quad x = 100$$

$$\phi = \frac{\pi}{2} \quad \Rightarrow \quad c(\frac{\pi}{2}) = -100 \quad \Rightarrow \quad x \cdot \cos(\frac{\pi}{2}) + y \cdot \sin(\frac{\pi}{2}) - 100 = 0 \quad \Rightarrow \quad y = 100$$

Hence, the image point is $(x,y) = (100,100)$.

Question 4 (6 points)

Assume that a white surface is illuminated by two light sources, a red light source and a green light source as depicted in the figure below. There are no other light sources involved. Explain why humans perceive the color of *yellow* in the overlapping area which is illuminated by both light sources.



Solution

According to the Young-Helmholtz theory the color perception of humans is based on the neural responses of the cones in the retina of the human eyes. There are three kind of cones, S-, M-, and L-cones which respond on light of different wavelength. Red light creates strong responses of the L-cones and weaker responses of the M-cones while green light creates strong responses of the M-cones and weaker responses of the L-cones. Hence, in the overlapping area both, the L-cones and the M-cones show strong responses. The wavelength of yellow light is between red and green light. Therefore, yellow light creates equally strong responses of the L- and M-cones. The S-cones are neither responding on red, yellow, or green light. Hence, the response pattern of the three kind of cones is the same if we illuminate a surface with yellow light or with red and green light at the same time.

Question 5 (6 points)

Create a potential function ϕ for a random field based image segmentation that models that pixels close to the left boundary of the image more likely belong to the segment labeled with 1 while the other pixels more likely belong to the segment labeled with 0. ϕ should depend on the position of a pixel (u,v) and its label $l(u,v)$ where u denotes the image column and v the image row.

Solution

Since small values of a potential function indicate prefered labelings, a suitable potential function would be

$$\phi((u,v), l(u,v)) = \begin{cases} 1 - \frac{u}{u_{\max}} & \text{if } l(u,v) = 0 \\ \frac{u}{u_{\max}} & \text{if } l(u,v) = 1 \end{cases}$$

or in one line:

$$\phi((u,v), l(u,v)) = |1 - l(u,v) - \frac{u}{u_{\max}}|$$

Question 6 (6 points)

Assume a pinhole camera with intrinsic parameters $\alpha' = \beta' = 500$, $(u_0, v_0) = (300,200)$, and skewing angle $\theta = 90°$. The images of the camera are 600 pixels wide and 400 pixels high. Assume that we rescale the images to half width and half height, i.e. the scaled images are 300 pixels wide and 200 pixels high. What are the values of α', β', u_0, and v_0 for the scaled images? Explain your solution.

Solution

The scaling factors α' and β' are created by multiplying the focal length of the camera with the resolution of the imager, i.e. the number of pixels per area. The focal length does not change by rescaling, however, we can interpret the rescaling procedure as taking just half of the resolution. That means, the new values for α' and β' are $\frac{1}{2} \cdot 500 = 250$. The position of the principle point physically does not depend on the resolution of the image. It is still in the center of the imager. However, due to the modified resolution its position in image coordinates becomes $(u_0,v_0) = (150,100)$.

Question 7 (7 points)

Assume that we trained a Support Vector Machine (SVM) on the three one-dimensional training examples that are given in the table below. The SVM uses the kernel function

$$K(x,y) = (x \cdot y + 1)^2$$

The SVM solver provides the Lagrange multipliers α_i for all training examples that are listed in the fourth column of the table below.

(a) Calculate the margin ρ of the trained SVM

(b) Calculate the bias weight b of the trained SVM

pattern No.	class label $d^{(i)}$	pattern $x^{(i)}$	Lagrange multiplier α_i
1	$+1$	2	$\frac{2}{11}$
2	-1	1	$\frac{2}{11}$
3	-1	0	0

Solution

As described on the lecture slide 17 of chapter 10 the Lagrange multipliers and the support vectors are sufficient to calculate ρ and b. However, since we use a non-linear kernel function, we have to replace the dot product in the equations on slide 17 by the kernel function.

(a) We obtain the margin

$$\rho = \frac{1}{\sqrt{\sum_{i=1}^{3} \sum_{j=1}^{3} (\alpha_i \alpha_j d^{(i)} d^{(j)} K(x^{(i)}, y^{(j)}))}} = \frac{1}{\frac{2}{11}\sqrt{25 - 9 - 9 + 4}} = \frac{1}{2}\sqrt{11}$$

(b) We obtain b from one of the two support vectors, e.g. from $x^{(1)}$ as

$$b = d^{(1)} - \sum_{i=1}^{3} \alpha_i d^{(i)} K(x^{(1)}, x^{(i)}) = 1 - (\frac{2}{11} \cdot 25 - \frac{2}{11} \cdot 9) = -\frac{21}{11}$$

Question 8 (8 points)

Which of the following strategies are appropriate to improve the classification performance of a classifier? Justify your answers briefly.

(a) collect additional examples and add them to the set of training examples

(b) add all validation examples to the set of training examples

(c) generate variations of the existing training examples by rotating the examples slightly

(d) randomly change the label of a subset of the training examples

Solution

(a) adding additional training examples is very useful since the additional examples will provide additional information on how examples typically look like so that the classifier can better distinguish between examples of different classes

(b) however, adding validation examples to the set of training examples will lead to overfitting so that the real performance of the classifier drops

(c) slight variations of existing training examples might also be a help for the classifier like in case (a). Since only slight rotations are mentioned in the task we can be sure that the rotated examples still belong to the same class as the unrotated examples.

(d) relabeling examples does not help but it confuses the classifier so that the accuracy drops.

Question 9 (6 points)

Implement a MATLAB function that takes a gray level image and adds jitter (stochastically independent random noise). The jitter should be generated from a Gaussian distribution with zero mean and given standard deviation. The signature of the function should be:

`function [ J ] = add_jitter ( I, sigma )`

where `I` is the input image, `sigma` is the standard deviation of the Gaussian distribution, and `J` is the resulting image.

Remark: You might find the following MATLAB functions useful for this task:
`M = normrnd(mu,sigma,m,n)` generates an $m \times n$ matrix of random numbers independently generated from a Gaussian distribution with expectation value `mu` and standard deviation `sigma`. The third and fourth argument `m` and `n` are optional. If they are not provided to the function, it returns a single random number.

Solution

```
function [ J ] = add_jitter ( I, sigma )
   [ rows cols ] = size(I);
   J = I + normrnd (0, sigma, rows, cols);
end
```

Gesamtpunkte: 60