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

Solutions for exam “Machine Vision” February 14, 2019

Question 1 (2+2+2 points)

Assume, we observe Moiré patterns in an image. Which of the following ideas can be used to avoid the Moiré patterns? Explain your answers briefly by refering to the sampling theorem.

(a) convolute the image with a Gaussian filter

(b) increase the focal length of the camera

(c) reduce the distance between camera and observed object

Solution

Once, Moiré patterns occur in an image we cannot get rid of them by post processing since the sampling theorem was already violated. Therefore, the option (a) does not help. Instead, we must increase the sampling rate of the camera, either by increasing the resolution or by using an optical setup with a larger magnification of objects. Options (b) and (c) increase the optical magnification and can be used.

Question 2 (7 points)

Describe how we could simplify the Hough transform in order to use graylevel gradient information for the estimation of graylevel edges. Which benefits would such an approach yield over the standard Hough transform.

Solution

The gradient provides an information about the local orientation of the edge. Hence, we would not need to consider all angles between 0° and 180° for each pixel in the Hough transform but only the angle which is provided by the graylevel gradient or only angles in a small interval around the angle provided by the graylevel gradient. This would reduce the computational effort and simplify the search for local maxima in parameter space since only relevant lines would enter the accumulator array.

Question 3 (4+3 points)

Given are the points $\vec{x_1} = (-1, 7)^{\mathrm{T}}$ and $\vec{x_2} = (7, 1)^{\mathrm{T}}$. You are supposed to find parameters $\vec{n}$ and c for the line connecting these two points using the line representation from the lecture (Hesse normal form).

(a) How many solutions for $\vec{n}$ exist? Calculate one. What is the corresponding c?

(b) Now, we want to find this line using the total least squares approach. What are the values for α, β and γ in the following Eigenvalue problem?

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \cdot \vec{n} = \lambda \cdot \vec{n}$$

Solution

(a) There are two possible $\vec{n}$: $(\frac{3}{5}, \frac{4}{5})^{\mathrm{T}}$ and $(-\frac{3}{5}, -\frac{4}{5})^{\mathrm{T}}$ all different vectors are either not orthogonal to the line or don't have a length of 1. The corresponding c is respectively -5 and 5.

(b)

$$\begin{aligned} \alpha &= \sum_i x_{i,1}^2 - \frac{1}{N}(\sum_i x_{i,1})^2 = 50 - \frac{1}{2} \cdot 36 = 32 \\ \beta &= \sum_i x_{i,1}x_{i,2} - \frac{1}{N}\sum_i x_{i,1} \sum_i x_{i,2} = 0 - \frac{1}{2} \cdot 48 = -24 \\ \gamma &= \sum_i x_{i,2}^2 - \frac{1}{N}(\sum_i x_{i,2})^2 = 50 - \frac{1}{2} \cdot 64 = 18 \end{aligned} \tag{1}$$

Multiples of α, β, and γ are also okay.

Question 4 (9 points)

Implement a MATLAB function which determines the correction factors c_R, c_B in order to perform a white balance on an input RGB image. The brightest pixel in the image is assumed to be white and should be used as reference pixel to calculate the correction factors. First of all, determine the reference pixel. Afterwards, calculate the correction factors and return them to the user. Keep in mind that the pixel values are stored as an 8-bit unsigned integer for each channel. Below, we provide the function header of the MATLAB function

Remark: The brightness (luminance) L of an RGB value is calculated as weighted sum of the RGB values:

$$L = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot G$$

```
function [c_R, c_B] = white_balance_factors(image)
% calculate the white balance correction factors from an RGB image
% use brightest pixel as reference pixel
% image is an 8-bit encoded unsigned integer RGB image
% add your code below this line
```

Solution

```
function [c_R, c_B] = white_balance_factors(image)
   [v_max, u_max] = size(image(:, :, 1));
   I_white = -1;

   for u = 1:u_max
     for v = 1:v_max
       I = 0.299 * image(v,u,1) +
           0.587 * image(v,u,2) +
           0.114 * image(v,u,3);
       if (I > I_white)
         u_white = u;
         v_white = v;
         I_white = I;
       end
     end
   end

   c_R = double(image(v_white, u_white, 2)) /
         double(image(v_white, u_white, 1));
   c_B = double(image(v_white, u_white, 2)) /
         double(image(v_white, u_white, 3));
```

```
Or even more simplified without using loops:
```

```
function [c_R, c_B] = white_balance_factors(image)
   intensity_image = rgb2gray(image);
```

```
[maximum_intensity, index] = max(intensity_image(:));
[v_white, u_white] = ind2sub(size(intensity_image),index);

c_R = double(image(v_white, u_white, 2)) /
      double(image(v_white, u_white, 1));
c_B = double(image(v_white, u_white, 2)) /
      double(image(v_white, u_white, 3));
```

Question 5 (3+3 points)

Are the following statements true or false? Justify your answers briefly. Provide a counterexample if a statement is false.

(a) If we turn an image by 90° and apply connected components labeling (CCL) on the rotated image we obtain the same partitioning of the image into segments as if we apply CCL on the original image.

(b) If we remove the topmost row of an image and apply CCL on the remainder we obtain the same partitioning of the remaining image as if we would apply CCL on the complete image and ignore the topmost row afterwards.

Solution

(a) This statement is true. CCL is partitioning the image into segments in such a way that two pixels belong to the same segment if and only if there is a path from one pixel to the other pixel such that all the neighboring pixels on the path are similar in color or graylevel, resp. This property is independent of the orientation of the image.

(b) This statement is not true. Assume a 2×3 graylevel image with white pixels in the first row and a white, a black, and a white pixel in the second row. If we apply CCL on the complete picture, we obtain two segments. The first segment contains all white pixels while the second segment contains the black pixel. If we would apply CCL only on the second row, we would obtain three segments, one for each pixel.

Question 6 (7 points)

Assume a pinhole camera that is described by its matrix of intrinsic parameters $A = \begin{pmatrix} 1000 & 0 & 500 \\ 0 & 1000 & 500 \\ 0 & 0 & 1 \end{pmatrix}$ and a circle on the plane $z = 10$ with center at position $(x,y,z) = (3,-1,10)$ and radius 2. Derive to which geometric curve this circle is mapped in the image of the camera. Describe the shape, position, and size of the curve. (x,y,z) refer to coordinates in the camera coordinate system.

Solution

All points on the circle can be described by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 10 \end{pmatrix} + 2 \cdot \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} \quad \text{width } \phi \in [0,2\pi)$$

By applying these points to the perspective projection we obtain

$$\begin{aligned} z \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} &= A \cdot \left(\begin{pmatrix} 3 \\ -1 \\ 10 \end{pmatrix} + 2 \cdot \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} 8000 \\ 4000 \\ 10 \end{pmatrix} + 2000 \cdot \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} \end{aligned}$$

Hence, we obtain

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 800 \\ 400 \end{pmatrix} + 200 \cdot \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$$

which is the representation of a circle in the image plane with center $(800,400)$ and radius 200.

Question 7 (3+3+3 points)

Derive for each of the following kernel functions $K(x,y)$ its respective nonlinear transformation $\Phi(x)$ such that $K(x,y) = \langle \Phi(x), \Phi(y) \rangle$. You may assume $0 < x,\ y < 1$.

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

(b) $K(x,y) = \cos(x - y)$

(c) $K(x,y) = \frac{1}{1-x^2y^2}$

Remark:

$$cos(\theta \pm \psi) = cos(\theta)cos(\psi) \mp sin(\theta)sin(\psi)$$

Remark:

$$\sum_{k=0}^{\infty} a_0 q^k = \frac{a_0}{1-q} \quad \text{for any } 0 < q < 1,\ a_0 \in \mathbb{R}$$

Solution

(a)

$$\begin{aligned} K(x,y) &= (x \cdot y + 1)^2 = x^2 \cdot y^2 + 2 \cdot x \cdot y + 1 \\ &= < \begin{pmatrix} x^2 \\ \sqrt{2} \cdot x \\ 1 \end{pmatrix}, \begin{pmatrix} y^2 \\ \sqrt{2} \cdot y \\ 1 \end{pmatrix} > \\ \Rightarrow \Phi(x) &= \begin{pmatrix} x^2 \\ \sqrt{2} \cdot x \\ 1 \end{pmatrix} \end{aligned}$$

(b)

$$\begin{aligned} K(x,y) &= \cos(x - y) = \cos(x) \cdot \cos(y) + \sin(x) \cdot \sin(y) \\ &= < \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}, \begin{pmatrix} \cos(y) \\ \sin(y) \end{pmatrix} > \\ \Rightarrow \Phi(x) &= \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} \end{aligned}$$

(c)

$$
\begin{aligned}
K(x,y) &= \frac{1}{1-x^2y^2} = \sum_{k=0}^{\infty}(xy)^{2k} \\
&= < \begin{pmatrix} 1 \\ x^2 \\ x^4 \\ x^6 \\ x^8 \\ \ldots \end{pmatrix}, \begin{pmatrix} 1 \\ y^2 \\ y^4 \\ y^6 \\ y^8 \\ \ldots \end{pmatrix} > \\
\Rightarrow \Phi(x) &= \begin{pmatrix} 1 \\ x^2 \\ x^4 \\ x^6 \\ x^8 \\ \ldots \end{pmatrix}
\end{aligned}
$$

Question 8 (6 points)

Assume a cascade classifier with three classification stages c_1, c_2, and c_3. We are facing four images I_1, I_2, I_3, and I_4 which are applied to the cascade classifier. The classification results are shown in the following table. Enter the output of the cascade classifier (either $+1$ or -1) for each image in the table. Furthermore, enter to the table which stages are evaluated for each image and which stages are not evaluated.

image	classification output			output of cascade	stages evaluated	stages not evaluated
	c_1	c_2	c_3			
I_1	$+1$	-1	-1			
I_2	$+1$	$+1$	$+1$			
I_3	-1	$+1$	$+1$			
I_4	$+1$	-1	$+1$			

Solution

image	classification output			output of cascade	stages evaluated	stages not evaluated
	c_1	c_2	c_3			
I_1	$+1$	-1	-1	-1	c_1, c_2	c_3
I_2	$+1$	$+1$	$+1$	$+1$	c_1, c_2, c_3	none
I_3	-1	$+1$	$+1$	-1	c_1	c_2, c_3
I_4	$+1$	-1	$+1$	-1	c_1, c_2	c_3

Question 9 (3 points)

Name three regularization techniques for the training of convolutional networks.

Remark: If more than three answers are given, only the first three answers count.

Solution

Possible answers are, e.g. early stopping, L2-regularization/weight decay, dropout, multi task learning

Gesamtpunkte: 60