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

Solutions for exam „Machine Vision“ August 9, 2017

Question 1 (6 points)

Assume we want to sample with a sampling interval of δ from the function f defined as

$$f(x) = 4\cos(2\pi x) + \frac{1}{2}\sin(\pi x)$$

Which condition on δ must be met to avoid Moiré patterns?

Solution

Since the function f is a linear combination of sine/cosine functions, its Fourier transform is zero almost everywhere except of the reciprocals of the period lengths of the sine/cosine functions. In this case, the periods are 1 and 2, i.e. the non-zero frequencies in the Fourier transform are ± 1 and $\pm\frac{1}{2}$. Due to the Nyquist-Shannon theorem, the sampling frequency has to be larger than twice the largest non-zero frequency, i.e. the sampling frequency must be larger than 2. Thus, δ must be less than $\frac{1}{2}$.

Question 2 (4 points)

Implement a MATLAB function that takes an 8-bit encoded graylevel image `IN` and returns a binary image `OUT` of the same size that satisfies the equation below:

$$\texttt{OUT}(\texttt{v},\texttt{u}) = \begin{cases} 1 & \text{if } \texttt{IN}(\texttt{v},\texttt{u}) \text{ is overexposed} \\ 0 & \text{otherwise} \end{cases}$$

Solution

```
function [ OUT ] = check_overexposed ( IN )
    OUT = (IN==255);
end
```

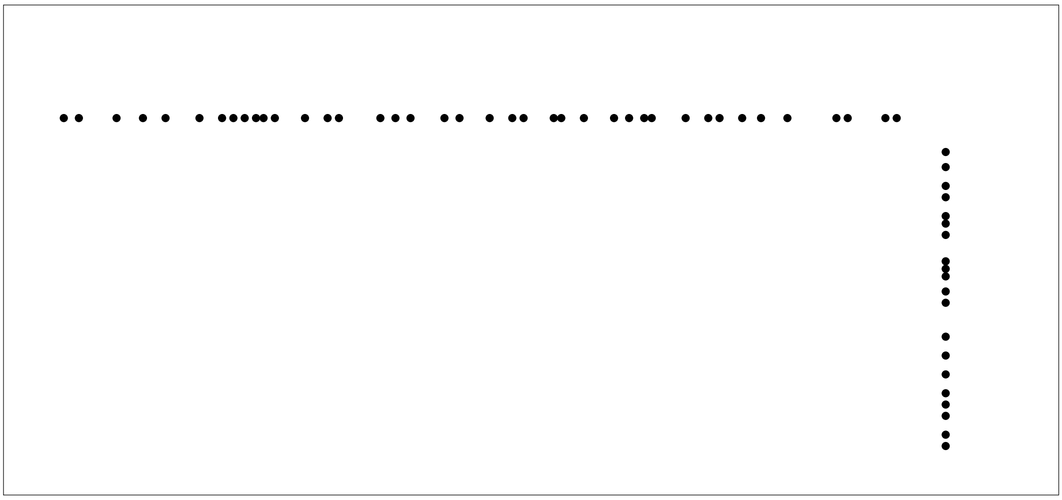
Question 3 (10 points)

The figure below shows in total 60 points in the two dimensional plane. 40 points have been sampled from a horizontal line, 20 points have been sampled from a vertical line.

(a) What would happen if we applied the total sum of least squares algorithm to fit a line to these 60 points? Which orientation would the estimated line have? Justify your answer briefly.

(b) What would happen if we applied the least trimmed sum of squares algorithm (LTS) to fit a line? In which way would the acceptance rate influence the result if we varied it between $\frac{1}{2}$ and 1? Justify your answer briefly.

(c) What is the probability that we obtain the horizontal line if we performed one trial of RANSAC? Derive or explain your solution briefly.



Remark: You may use and modify the figure above to explain your solution.

Remark: The subtasks can be solved independently.

Solution

(a) Since all points have the same influence on the outcome of the total sum of least squares approach we would obtain a line that passes in diagonal manner from left top to right bottom.

(b) LTS ignores those points with the largest residuals. If the acceptance rate is small enough, only points of the horizontal line are considered so that we obtain the horizontal line as result. This happens if the acceptance rate is at most $\frac{40}{60} = \frac{2}{3}$. If the acceptance rate is higher, the resulting line will consider the points on the horizontal line and the topmost points from the vertical line which yields a slighly skewed line. If the acceptance rate is 1 we obtain the same result as in subtask (a).

(c) The probability to select the first point from the horizontal line is $\frac{40}{60}$. The probability that we select the second point as well from the horizontal line is $\frac{39}{59}$. Thus, the total probability that we randomly select both points from the horizontal line and, hence, to obtain the horizontal line is $\frac{40}{60} \cdot \frac{39}{59} = \frac{78}{177}$.

Question 4 (8 points)

Depicted below is a square area of a color image in which color is represented in RGB values ranging from 0 to 255.

(108,100,64)	(113,111,80)
(97,89,80)	(122,100,96)

(a) Calculate the correction terms c_R and c_B for the white balance procedure that was introduced in the lecture based on these four pixels.

(b) Apply the white balance with the correction terms calculated before to a pixel with RGB values (55,80,160). What are the resulting RGB values?

Solution

(a) The average red, green, and blue values are

$$\begin{aligned} \bar{r} &= \frac{1}{4}(108 + 113 + 97 + 122) = 110 \\ \bar{g} &= \frac{1}{4}(100 + 111 + 89 + 100) = 100 \\ \bar{b} &= \frac{1}{4}(64 + 80 + 80 + 96) = 80 \end{aligned}$$

Hence, we obtain

$$\begin{aligned} c_R &= \frac{\bar{g}}{\bar{r}} = \frac{100}{110} = \frac{10}{11} \\ c_B &= \frac{\bar{g}}{\bar{b}} = \frac{100}{80} = \frac{5}{4} \end{aligned}$$

(b) We obtain the new RGB values by multiplying the red value with c_R and the blue value with c_B. Hence we obtain

$$RGB_{new} = (50,80,200)$$

Question 5 (5 points)

Let g be a binary image, i.e. all pixel values are either 0 or 1. Furthermore, let h_1 be the binary image that we obtain by applying the erosion operator to g and h_2 the binary image that we obtain by applying the dilation operator to h_1. Prove or disprove that $g = h_2$. For both operations, the erosion and the dilation, we use the 8-neighborhoodship.

Solution

The statement is not true. Counterexample:

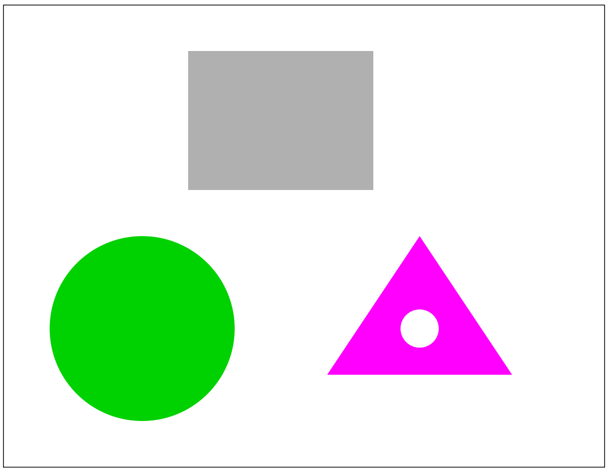
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[erosion]{} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[dilation]{} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Question 6 (8 points)

We want to segment a color image with a level set method. The image is represented in the HSV color space where $s(u,v) \in [0,1]$ refers to the saturation at pixel position (u,v). The level set evolution should shrink the foreground segment in low saturated areas and stop shrinking in highly saturated areas.

(a) Provide an evolution term $\frac{\partial \vec{x}}{\partial t}$ that implements this idea. You may assume that a signed distance function ϕ is given.

(b) The figure below shows an image with white background and three colored objects, a gray rectangle, a green circle, and a magenta triangle with a white hole. Assume we start the level set evolution with a foreground segment that contains the whole image, i.e. the boundary of the foreground segment equals the image boundary. How does the boundary of the foreground segment look like after convergence of the level set evolution? Draw the boundary into the image below.

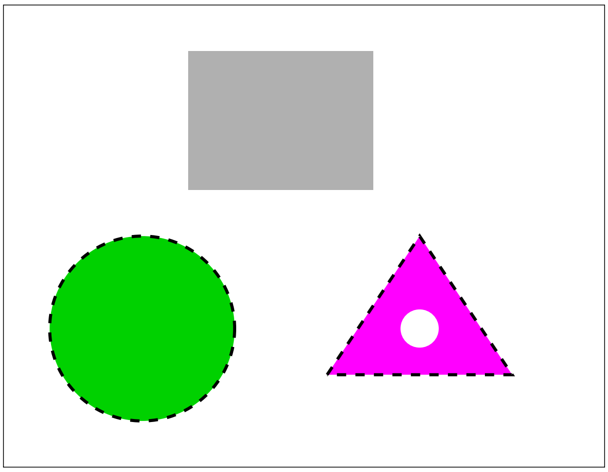


Solution

(a) Let $\alpha > 0$ be a constant that controls the amount of shrinking and ϕ the signed distance function. Then, the idea is modeled by

$$\frac{\partial \vec{x}(u,v)}{\partial t} = -\alpha \cdot (1 - s(u,v)) \cdot \frac{\nabla \phi(u,v)}{||\nabla \phi(u,v)||}$$

(b) Since the magenta and green areas are highly saturated while the gray area is unsaturated, the boundary will enwrap the green circle and the magenta triangle as depicted below. The white circle interior of the triangle will not generate a boundary although it is unsaturated since the decision boundary will never reach this area during evolution. The picture below shows the resulting boundary with a dashed line.



Question 7 (7 points)

Assume a pinhole camera that is described by its matrix of intrinsic parameters

$$A = \begin{pmatrix} 500 & 0 & 1000 \\ 0 & 500 & 750 \\ 0 & 0 & 1 \end{pmatrix}$$

We want to reconstruct a 3d point P. We know, that P is mapped into the image at image position $(u,v) = (1250,500)$. Furthermore, we know, that P is located on a ball with radius $r = \sqrt{17}$ and center point $(m_x,m_y,m_z) = (0,0,1)$ in the camera coordinate system. Derive the 3d position of point P in the camera coordinate system.

Remark: The solution of an equation of the form $ax^2 + bx + c = 0$ is $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$

Solution

Point P is located in a line of sight that we can calculate from its image position as

$$p_z \cdot (1250, 500)^T = p_z \cdot (u,v)^T = A \cdot (p_x, p_y, p_z)^T = (500p_x + 1000p_z, 500p_y + 750p_z)^T$$

Hence, $(p_x, p_y, p_z)^T = p_z \cdot (\frac{1}{2}, -\frac{1}{2}, 1)^T$ with unknown $p_z > 0$. We know, that P is located on the ball, i.e.

$$17 = r^2 = ||(p_x, p_y, p_z)^T - (m_x, m_y, m_z)||^2 = \frac{1}{4}p_z^2 + \frac{1}{4}p_z^2 + (p_z - 1)^2 = \frac{3}{2}p_z^2 - 2p_z + 1$$

We obtain two possibilities for p_z, i.e. $p_z = 4$ and $p_z = -\frac{8}{3}$. Since the observed point is in front of the camera, i.e. $p_z > 0$, we can reject the second solution and obtain $(p_x,p_y,p_z)^T = 4 \cdot (\frac{1}{2}, -\frac{1}{2}, 1)^T = (2, -2,4)$.

Question 8 (6 points)

Are the following statements true or false? Justify your answers briefly.

(a) Support vector machines do not allow to solve non-linear classification tasks.

(b) In the soft margin case of a support vector machine, the number of support vectors is a lower bound of the number of misclassified training examples.

(c) The larger the parameter k in k-fold cross validation is, the more computational effort is required.

Solution

(a) This statement is wrong. The usage of kernel functions different from the dot product kernel allows to generate non-linear classifiers.

(b) This statement does not hold. In contrast, the number of support vectors is an upper bound of the number of misclassified training examples since misclassified examples automatically become support vectors.

(c) This statement is in general true, since the larger k is, the more often the classifier must be retrained.

Question 9 (6 points)

Assume, we trained three different classifiers c_1, c_2, and c_3 for a binary classification task. We evaluated the classifiers on a test set with 50 positive examples and 50 negative examples. The positive examples are enumerated from 1 to 50 while the negative examples are enumerated from 51 to 100. The table below provides the test examples which were misclassified by the three classifiers.

(a) Calculate precision and recall for classifier c_1. You may provide the numbers in terms of fraction numbers.

(b) How many false positives and how many false negatives would an ensemble of the three classifiers c_1, c_2, and c_3 achieve?

classifier	misclassified test examples
c_1	5, 26, 42, 61, 62, 68, 99
c_2	3, 5, 49, 62, 91, 93
c_3	11, 34, 51, 62, 78, 93

Remark: The subtasks can be solved independently.

Solution

(a) c_1 misclassifies 3 positive and 4 negative examples, i.e. the number of false positives is 4 and the number of false negatives is 3 while the number of true positives is 47 and the number of true negatives is 46. Hence, the classifier achieves a precision of $\frac{47}{47+4} = \frac{47}{51}$ and a recall of $\frac{47}{47+3} = \frac{47}{50}$.

(b) Since an ensemble classifier decides according to the majority of its members, the ensemble classifier with three members misclassifies an example if at least two of its members misclassify it. In the given task, this applies to examples no. 5, 62, and 93. Hence, we obtain one false negative (no. 5) and two false positives (no. 62, 93).

Gesamtpunkte: 60