Practice Exam – Sample Solutions

Note that in the actual exam, no calculators, no books, and no notes allowed.

Question 1: ___ out of ___ points
Question 2: ___ out of ___ points
Question 3: ___ out of ___ points
Question 4: ___ out of ___ points
Question 5: ___ out of ___ points
Question 6: ___ out of ___ points
Question 7: ___ out of ___ points
Question 8: ___ out of ___ points

Total Score:

Grade:
**Question 1: True or False?**

Tell whether each of the following statements is true or false by checking the appropriate box. Do not check any box if you do not know the right answer, because you will lose points for incorrect answers.

<table>
<thead>
<tr>
<th>Statement</th>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) For mean shift segmentation, we need to know the number of regions in advance.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>b) The Hough transform can only be used to detect straight lines.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>c) The median filter is especially well-suited for removing salt-and-pepper noise.</td>
<td>[X]</td>
<td>[ ]</td>
</tr>
<tr>
<td>d) The difference code of a curve does not change when we rotate the curve by 90 degrees.</td>
<td>[X]</td>
<td>[ ]</td>
</tr>
<tr>
<td>e) The Laplacian convolution filter computes the second derivative of the input image.</td>
<td>[X]</td>
<td>[ ]</td>
</tr>
<tr>
<td>f) When we move backward, the direction of optical flow throughout our visual field is perpendicular to the gradient of its magnitude.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>g) Textures with the same gray co-occurrence matrices are always identical.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>h) When we expand a binary image A and then shrink the result, we always obtain image A again.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>i) Local binary patterns are represented by 256-bin histograms.</td>
<td>[ ]</td>
<td>[X]</td>
</tr>
<tr>
<td>j) There is no 2D shape with greater compactness (i.e. a smaller value for P²/A) than a circle.</td>
<td>[X]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
Question 2: Efficient Range Imaging

Imagine that you want to do range imaging through triangulation. In order to be efficient, you decide to use a pattern of eight vertical lines that only needs to be shifted by the distance between neighboring lines in order to measure the depth of all visible points in the scene. In order to avoid associating points with an incorrect vertical line, you take multiple images of the same scene with the pattern in the same position. In each picture, a different subset of lines is projected.

a) What is the minimum number of pictures that you need to take for each projection angle in order to be sure to determine the number of the line that any point belongs to? Why?

You need 4 pictures. Although 3 pictures would suffice to represent $2^3 = 8$ unique on/off patterns, it would mean that one of the lines would be off in all 3 pictures and thus could not be localized.

b) Derive a general formula that tells you for a given number of lines the minimum number of pictures necessary. Explain how you derived the formula.

Given the answer to (a), it follows that $p$ pictures can uniquely identify up to $l = (2^p - 1)$ lines. This gives us:

\[
\begin{align*}
    l &= 2^p - 1 \\
    \Rightarrow l + 1 &= 2^p \\
    \Rightarrow \log_2(l + 1) &= p
\end{align*}
\]

Since we want to know the minimum number $p_{\text{min}}$, we use the ceiling function:

\[
\Rightarrow \lceil \log_2(l + 1) \rceil = p
\]
**Question 3: Polyline Splitting**

Describe the algorithm for polyline splitting to approximate the shape of a contour by a sequence of connected straight lines. Use diagrams to illustrate the concept.

Here you should put something similar to slides 2 to 5 from March 6.
Question 4: Slope Density Functions

(a) Draw the slope density function of the following contour (let us say that we are going in counterclockwise direction and an angle of 0° means downward):

(b) Describe how this function changes if you mirror the contour along a horizontal axis, i.e., turn it upside-down. You can describe it in words or draw another diagram.

It is like a 90° counterclockwise rotation, so everything is shifted by 90° to the right (and wrapping around):
Question 5: Errors in Depth Perception

Having two eyes, we are able to perceive the depth (z-distance) of an object through binocular disparity information. As you know, we can also give our computer vision system two cameras and let it do the same thing. The question is: How accurate is its estimation of depth, and on what factors does this accuracy depend?

The main problem here is the limited accuracy and resolution of the cameras. Let us say that the actual position of an object in the camera image may be up to one millimeter to the left or to the right of its actual position. For example, if the left camera measures $x_l = 5.3$ mm, it means that the object’s actual position could be anywhere between $x_l = 5.2$ mm and $x_l = 5.4$ mm.

(a) Given this camera accuracy, what is the z-range that an object could have (i.e., the minimum and maximum z-distance possible), if the cameras with baseline $b = 10$ cm and focal length $f = 20$ cm measure positions $x_l = 6.1$ and $x_r = 5.1$? If you do not remember the equation, try to derive it; it is not very difficult.

The most extreme cases are $x_l = 6.0$ and $x_r = 5.2$ and $x_l = 6.2$ and $x_r = 5.0$. In the first case, we derive distance $z$ as:

$$z = 10 \text{ cm} \cdot 20 \text{ cm} / 1.2 \text{ cm} = 166.67 \text{ cm}$$

In the second case, we get:

$$z = 10 \text{ cm} \cdot 20 \text{ cm} / 0.8 \text{ cm} = 250 \text{ cm}$$

The z-distance ranges from 166.67 to 250 cm.

(b) What do you think will happen if the object is much further away from the system? Will the error in z-distance measurement (i.e., the z-range) increase or decrease? Why?

The z-distance error will increase, because will greater distance the same variation in the camera image corresponds to greater variation in the z-distance measurement (it can also be seen from the equations above).

(c) What do you think will happen if we keep the object in the same place as in (a), but increase the distance between the cameras, i.e., the baseline $b$? Will the error in z-distance measurement (i.e., the z-range) increase or decrease? Why?

Now the z-distance error will decrease, because the same variation in the camera image corresponds to smaller variation in the z-distance measurement. It’s harder to see in the equations this time, because the baseline, the focal length, and the difference between $x_l$ and $x_r$ will change. But do the math, it’s a great exercise.
(d) If we increase the baseline, another problem becomes more and more difficult. What problem is that?

The stereo matching, i.e., determining the correspondence between points in the left and the right image, becomes more difficult. This is because we are now looking at objects from more distinct angles, and so the two images of the same object look less similar.

**Question 6: Teaching the Artificial Brain**

Describe the basic principles underlying supervised learning in artificial neural networks. For example, how is the network “presented” with the information that it is supposed to learn? How does the learning happen, i.e., what changes in the network, and based on what criteria? Use at least two paragraphs to outline the most important ideas. Do not use any equations.

To teach the network to compute a function, we need to provide information in the form of exemplars, i.e., desired input-output pairs. If, for example, the network is supposed to output “1” for the input (1, 0, 1), then the corresponding exemplar is ((1, 0, 1), 1).

Then we randomly pick one of the exemplars and feed its input part into the input layer neurons of our network. The network then performs all its computations, i.e., each neuron multiplies each of its inputs with the corresponding synaptic weight and applies the output function, usually a sigmoid function, to determine its output. Then we compare the network’s output with the desired output specified in the exemplar. If there is a discrepancy, then we use a rule (e.g., the backpropagation learning rule) to change the weights in such a way that next time the same input is presented, the output will be closer to the desired one. Such rules typically use gradient-descent, i.e., they move against the error gradient in weight space.

After each exemplar was shown to the network and its weights changed accordingly, one epoch is finished. If the overall network error, i.e., the deviation between desired and actual output values, is still larger than a certain threshold, we perform another epoch, and so on, until the error falls below the threshold. Then the network should have learned the function and be able to predict appropriate output values even for inputs that it has never seen before.

(This text is more detailed than what you would need to write in order to get the full score.)
Question 7: Signature, Please!

Use the diagram on the right to sketch the signature of the contour on the left. It does not have to be perfectly scaled. Choose any starting point you like and indicate it on the contour on the left.

Question 8 (Bonus Question): Tricking the Difference Image Technique

In difference images, pixel values greater than zero (or greater than a certain threshold) are thought to reflect moving objects in the scene. However, this does not always have to be the case. There are other factors besides object motion that can cause non-zero values in a difference image. List as many such factors as you can think of.

- Change in lighting
- Noise in the images varying over time
- Motion of the camera