

**CS 675 – Computer Vision – Fall 2007**

**Instructor: Marc Pomplun**

**Practice Exam**

**Duration: 2 hours and 30 minutes**

Notice that in the actual exam, no calculators, no books, and no notes are 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

**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.

Statement	True	False
a) Image understanding is the highest-level problem in computer vision.	<input type="checkbox"/>	<input type="checkbox"/>
b) Let us apply a size filter to an image A, resulting in image A', and then apply the same size filter to A', resulting in image A''. Then it is possible that $A' \neq A''$ .	<input type="checkbox"/>	<input type="checkbox"/>
c) The thinning algorithm can also be applied to three-dimensional structures.	<input type="checkbox"/>	<input type="checkbox"/>
d) The k-means algorithm will always yield the same clustering result when started on the same dataset.	<input type="checkbox"/>	<input type="checkbox"/>
e) Self-organizing maps perform unsupervised learning.	<input type="checkbox"/>	<input type="checkbox"/>
f) When we move forward, the direction of optical flow in our visual field is always perpendicular to the gradient of its amplitude.	<input type="checkbox"/>	<input type="checkbox"/>
g) Two different shapes, even if they cannot be turned into each other by rotation, can have identical signatures.	<input type="checkbox"/>	<input type="checkbox"/>
h) In the Canny Edge Detector, non-maxima suppression has the effect of thinning the detected edges to a width of one pixel.	<input type="checkbox"/>	<input type="checkbox"/>
i) Support Vector Machines can learn to classify objects based on a set of given sample objects for each class.	<input type="checkbox"/>	<input type="checkbox"/>
j) In artificial neural networks, supervised learning is biologically more plausible than unsupervised learning.	<input type="checkbox"/>	<input type="checkbox"/>

## Question 2: Laplace and His Filter

To keep things simple, let us assume that you have the following one-dimensional image:

7	8	7	9	16	18	19	17	18	14	10	5	3	2	2
---	---	---	---	----	----	----	----	----	----	----	---	---	---	---

Your task is to find the edges, i.e., the points of the most rapid change in brightness, in this image. To do this, you have to use a Laplacian convolution filter.

- (a) Unfortunately, you totally forgot what a one-dimensional Laplacian filter looks like. The only thing that you remember is that it computes the second derivative of a given intensity function (the image). So you start thinking: “Well, I know that a simple gradient (first-derivative) filter looks like this:

-1	1
----	---

Maybe I can simply derive the Laplacian filter from this first-derivative filter.”  
Well, nice idea. Please do that and explain exactly how you do it (use next page).

- (b) Now that you have a Laplacian filter, please apply it to the image above and enter the result in the empty image below. If you do not remember the Laplacian Filter and also do not know how to derive it, please feel free to use any method you like to compute the second derivative of the image.

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

- (c) Now that you have the second derivative, please indicate in it (above) where the edges are. Explain how you determine the edges of an image based on its second derivative.



### Question 3: Customizing the Hough Transform

You just started a job at the fast food chain McDonald's. However, they did not hire you for serving fries. The task they want you to do is building a system that can count the hamburgers in the kitchen that have been prepared and are ready to be wrapped, sold, and eaten. The first thing you do is to install a camera at the kitchen ceiling. This camera looks from above onto the kitchen table area containing those hamburgers. You notice that in the camera picture, all hamburgers look like perfect discs with a radius of precisely  $r$  pixels. This  $r$  is the same for all hamburgers (McDonald's is known to be a high-accuracy manufacturer of food).

Explain and describe in detail how you can use the Hough transform to solve this problem. How do you pre-process the camera image before you apply the Hough transform, and why? Define the input and output spaces for the Hough transform, develop the formulas for the mapping from the input space to the output space, and describe how you will process the output in order to obtain the number of hamburgers in the image. Discuss possible problems with this approach and how you would tackle them. Please provide sufficient detail and give reasons for your design decisions.





#### Question 4: Recognizing Objects

Let us assume that you want to write a computer vision program that can distinguish between two types of objects, named Gnorks and Ploffs. After studying these object types, you decide that two visual features are especially important for distinguishing between them, namely their area and their compactness.

Unfortunately, your database contains the data of only 4 Gnorks and 4 Ploffs, but you want to build a reasonable classifier based on only this small amount of data. Let us name the Gnorks  $G_1, \dots, G_4$  and the Ploffs  $P_1, \dots, P_4$ , and write their data as pairs (area, compactness):

$G_1 = (100, 3), G_2 = (400, 7), G_3 = (200, 4), G_4 = (600, 5),$   
 $P_1 = (300, 3), P_2 = (100, 2), P_3 = (400, 1), P_4 = (500, 4).$

Visualize the data in a diagram and write down a classification function  $f(a, c) = t$ , where  $a$  is the area of a given object,  $c$  is its compactness, and the output  $t$  is the object's class, which should be 0 for a Gnork and 1 for a Ploff. Assume that other (unknown) Gnorks and Ploffs are similar to the ones whose data you have, and try to make the function work reasonably well for such unknown inputs under this assumption.



### 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 objects' 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 formula, try to derive it; it is not very difficult.
- (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?
- (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?
- (d) If we increase the baseline, however, another problem becomes more and more difficult. What problem is that?





### **Question 6: Teaching the Artificial Brain**

Describe the basic principles underlying supervised learning in artificial neural networks. Use at least two paragraphs to outline the most important ideas.

**Question 7 (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.



