ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Visual Computing
AS 2014

Prof. M. Gross (Computer Graphics Lab)
Prof. M. Pollefeys (Computer Vision Group)
C. Öztireli, A. Cohen, C. Häne,
Y. Aksoy, A. Milliez, C. Schumacher

Final Exam

4 February 2015

First and Last name: ______________________________

ETH number: ______________________________

Signature: ______________________________

General Remarks

- At first, please check that your exam questionnaire is complete (there are 17 pages total).
- Remove all material from your desk which is not allowed by examination regulations.
- Fill in your first and last name and your ETH number and sign the exam. Place your student ID in front of you.
- You have 3 hours for the exam. There are 12 questions, where you can earn a total of 180 points. You don't have to score all points to obtain a very good grade.
- Start each question on a separate sheet. Put your name and ETH number on top of each sheet. Only write on the question sheet where explicitly stated.
- Please do not use a pencil or red color pen to write your answers.
- You may provide at most one valid answer per question. Invalid solutions must be canceled out clearly.

	Topic	Max. Points	Points Achieved	Visum
1	Morphological Operators and Texture	11		
2	Optical Flow	19		
3	Filtering	15		
4	Edge Detection	9		
5	Hough Transform	6		
6	Fourier Transform	22		
7	JPEG Compression	2		
8	Principal Component Analysis	6		
9	Light and Colors	20		
10	Transformations	28		
11	Rendering and Geometry	34		
12	Animation and Physics	8		
Total		180		

Grade: ....

Question 1: Morphological Operators and Texture (11 pts.)

a) Define opening and closing of an image I by a structuring element S. Give the set formulation for any basic operators you may use. **3 pts.**

b) A pair of images is shown in Figure 1. Image A was processed by a morphological operation to produce image B. A structuring element S of size 3 x 3 was used in this transformation. Choose the operation used from the list provided. **2 pts.**

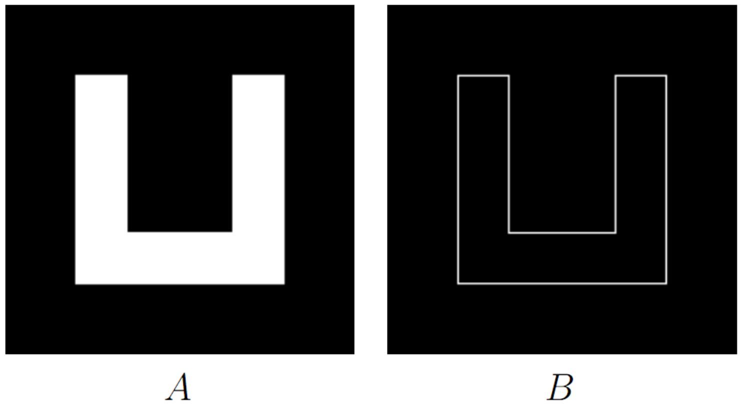


Figure 1

(a) $B = (A \ominus S) \cap A^c$
(b) $B = (A \oplus S)^c \cup A$
(c) $B = (A \oplus S) \cap A^c$
(d) $B = (A \ominus S)^c \cup A^c$

c) One way of representing textures is by using oriented pyramids.

i) Briefly give an outline of this method. **1 pt.**

ii) Why do we use a pyramid of the image instead of different sized filters in order to get the scale information? **1 pt.**

d) Why is it that histograms are not a good choice for characterizing textures? Can you think of a more adapted tool to characterize textures based on the same principle as histograms? **2 pts.**

e) Texture synthesis by using image-based methods can be fairly simple and efficient. One of the methods seen in class is the Chaos Mosaic method.

i) State the three main steps of this method. **2 pts.**

ii) For which kind of textures does this method work the best? When does it fail? Give an example of each kind of texture. **2 pts.**

Question 2: Optical Flow (19 pts.)

a) The Lucas-Kanade algorithm can be used to estimate the optical flow of an image sequence. Nevertheless, it doesn't work for all cases. State the 3 conditions that have to be fulfilled so that this method works **3 pts.**

b) You have captured a video at 25 frames per second of a car moving at 18 kilometers per hour. The side of the car is parallel to the image plane and the car is moving straight. The car is 2.4 meters long, but in your video it is 192 pixels long. Assume that your optical flow algorithm breaks down for pixel displacements that are larger than 1 pixel. By using image pyramids in a coarse-to-fine approach you can still use your algorithm. Explain this method, how you would proceed and how many pyramid levels you would need (the original image counts as one level and in each level the image size is half the size of the level before). **4 pts.**

c) For some cases, the optical flow field of an image sequence can be explained by a global parametric motion model. Assume that the global flow occurs because of 3D rigid motion. This means that the motion can be modeled by a rotation R and a translation T. In other words, for any 3D point $[X, Y, Z]$ of the scene at the time t, its new position $[X', Y', Z']$ on the next frame at time $t+1$ can be calculated using the following formulation:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T$$

Let $R = [R_X, R_Y, R_Z]$ be the rotation applied, and $T = [T_X, T_Y, T_Z]$ the translation. If the angles of rotation corresponding to each axis are small enough, then the above equation can be approximated by:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & -R_Z & R_Y \\ R_Z & 1 & -R_X \\ -R_Y & R_X & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

i) Give an example of an image sequence to which this model could be applied. **1 pt.**

ii) Under orthographic projection, derive the basic equations of the motion field:

$$u = v_x = -R_Z y + R_Y Z + T_X$$

$$v = v_y = R_Z x - R_X Z + T_Z$$

Where $[x, y]$ are the 2D coordinates of the projected 3D point $[X, Y, Z]$. Assuming very short time between frames, both 2D and 3D velocities can be approximated by the difference between the new point position and the old point position. **3 pts.**

iii) How would the results from the previous question change under perspective projection? Give an idea of how the motion field would look like in this case (you don't need to simplify the equations until they look nice) **3 pts.**

iv) Assuming a planar scene and orthographic projection, derive the new equations for the flow field.

$$u = v_x = a_1 x + a_2 y + b_1$$

$$v = v_y = a_3 x + a_4 y + b_2$$

State the expression of each parameter a_i **2 pts.**

d) Optical flow has many uses and application. State 3 possible applications and explain in no more than 2 lines their general outline. **3 pts.**

Question 3: Filtering (15 pts)

a) We are given a filter $K(i,j)$ with $i \in [-N,N]$ and $j \in [-M,M]$

$$K(i,j) = i^2 j + i^4 j^5 + i^2 j^5 + i^4 j$$

Is this filter separable? Explain why or why not. **3 pts.**

b) What is the advantage of a separable 3×3 filter over a non-separable one? **1 pts.**

c) What is an integral image (also called summed area table)? How is it computed? **2 pts.**

d) Give an example of a 2D box filter. **2 pts.**

e) Explain how an image can be box filtered in such a way that the running time is independent of the filter size. **4 pts.**

f) Given filter $K = [0,0,0;0,0,1;0,0,0]$ (Matlab notation). What effect has this filter on an image when convolving the image with the filter. Is the filter shift invariant? Explain why or why not. **3 pts.**

Question 4: Edge Detection (9 pts)

a) Looking at the first and second derivative of an image how can you tell that there might be an edge? How could you only select strong edges? **3 pts.**

b) Using filtering, image gradients can be computed. Which image filters are needed? How often do you need to filter the image. Give the formulas to compute the gradient magnitude and the gradient orientation. **3 pts.**

c) In the lecture the Canny edge detection algorithm has been presented in 5 steps. Give a list of the 5 steps.

3 pts.

Question 5: Hough Transform (6 pts)

a) For the Hough transform a line is parametrized as $x\cos(\theta) + y\sin(\theta) = \rho$ instead of the standard parameterization $y = mx + q$. Which problem is avoided by using the (θ, ρ) parameterization. **2 pts.**

b) The Hough transform can also be used to detect circles. Which parameterization would you use? **2 pts.**

c) If we would use the standard hough transform for line detection to detect line segments instead of lines. What problem would occur? **2 pts.**

Question 6: Fourier transform (22 pts.)

a) The Fourier transform $X(f)$ of a signal $x(t)$ is defined as

$$\mathcal{F}\{x(t)\} = X(f) = \int_{-\infty}^{\infty} x(t)e^{-2\pi i f t}dt$$

i) For scalar a, signals $x(t)$ and $y(t)$ and their Fourier transforms $X(f)$ and $Y(f)$, show that the following equations hold. **3 pts.**

$$\mathcal{F}\{ax(t)\} = aX(f) \tag{1}$$

$$\mathcal{F}\{x(t) + y(t)\} = X(f) + Y(f) \tag{2}$$

ii) Derive Fourier transform of $x(t - t_0)$ in terms of Fourier transform of $x(t)$. **6 pts.**

iii) A sampling function s(t), *i.e.* an impulse train with period T, and its Fourier transform $S(f)$ can be represented as:

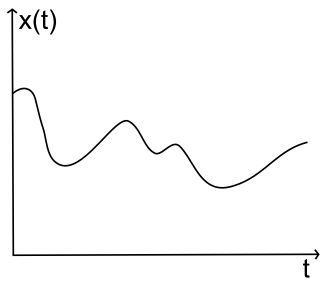
$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \qquad\qquad S(f) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T}\right)$$

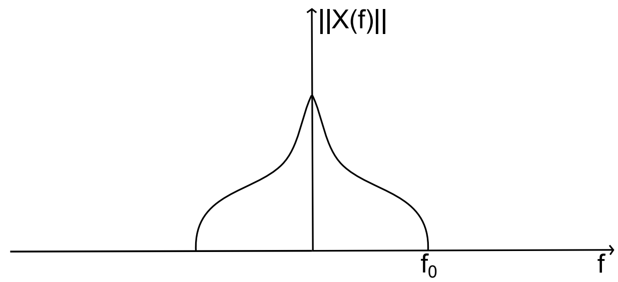
where $\delta(\cdot)$ is the Dirac-delta function.
A continuous signal $x(t)$ can be sampled by multiplying it by the sampling function:

$$x_s(t) = x(t)s(t)$$

Using the convolution theorem, derive the Fourier transform of $x_s(t)$ in terms of the Fourier transform of $x(t)$. **6 pts.**

b) Assume that we are given a continuous band-limited one dimensional signal $x(t)$ and magnitude of its Fourier transform $X(f)$ as shown below.





We sample $x(t)$ using three different sampling intervals T_1, T_2 and T_3 ($T_1 > T_2 > T_3$) to get the sampled signals $x_1(t)$, $x_2(t)$ and $x_3(t)$ in continuous domain. Sampling is done by multiplying $x(t)$ with unit impulse (or Dirac delta function) trains:

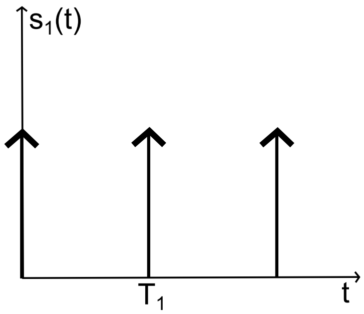
$$x_1(t) = x(t)s_1(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_1)$$

Illustrations of the three sampling functions $s_1(t)$, $s_2(t)$ and $s_3(t)$ are given below. We know that sampling frequency of $s_2(t)$ is equal to Nyquist frequency of $x(t)$.

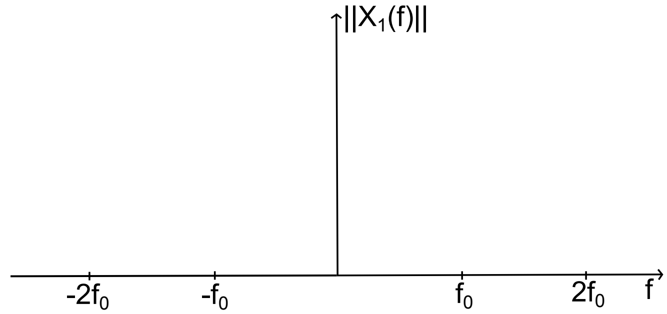
i) Roughly sketch the magnitudes of the Fourier transform of three sampled signals on corresponding plots provided below. **4 pts.**

ii) Express sampling interval T_2 of sampling function $s_2(t)$ in terms of f_0. **1 pt.**

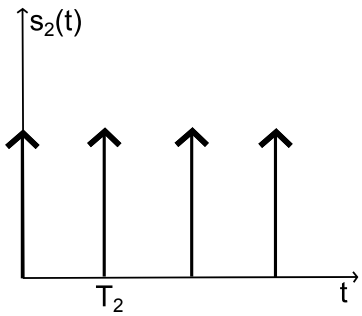
iii) Which one of the sampling functions cause aliasing? Without changing the sampling frequency, propose a method to avoid aliasing. **2 pts.**



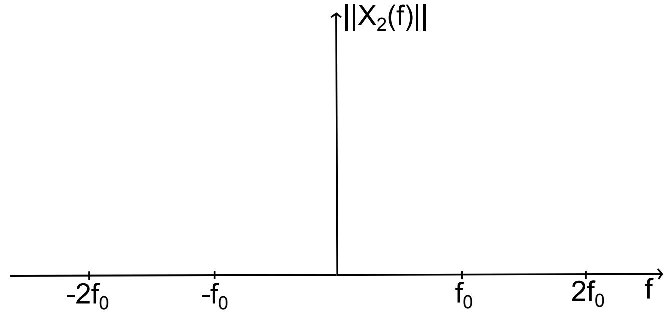
s1(t)
T1
t



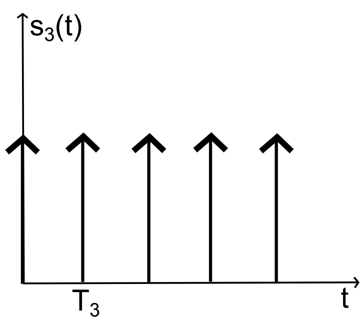
||X1(f)||
-2f0
-f0
f0
2f0
f



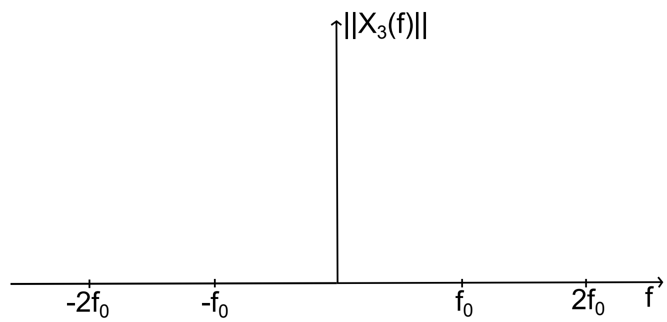
s2(t)
T2
t



||X2(f)||
-2f0
-f0
f0
2f0
f



s3(t)
T3
t



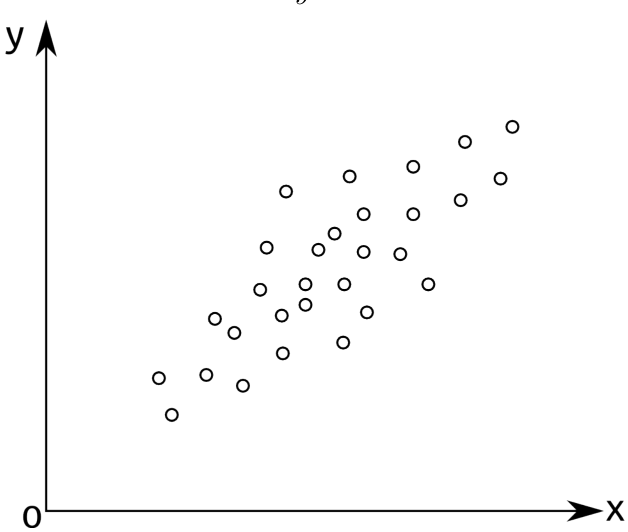
||X3(f)||
-2f0
-f0
f0
2f0
f

Question 7: JPEG Compression (2 pts.)

Given the uncompressed (top) and JPEG-compressed (bottom) images below, shortly explain the reason of the quality difference between them considering Discrete Cosine Transform used in JPEG compression.

Question 8: Principal Component Analysis (6 pts.)

You are given the following dataset S with $n = 2$ dimensional inputs, meaning that each sample has 2 features: x and y. We run PCA on this dataset.



a) Draw the principal components you expect from PCA on the figure. Clearly indicate the first principal component. **2 pts.**

b) What does the first principal component represent in terms of the distribution of the samples? **2 pts.**

c) Explain for what purpose we can use PCA in terms of representation of the samples. **2 pts.**

Question 9: Light and Color (20 pts.)

a) RGB and CMY Color Spaces

i) Explain two main differences between the RGB and CMY color spaces.

2 pts.

b) CIE Experiment

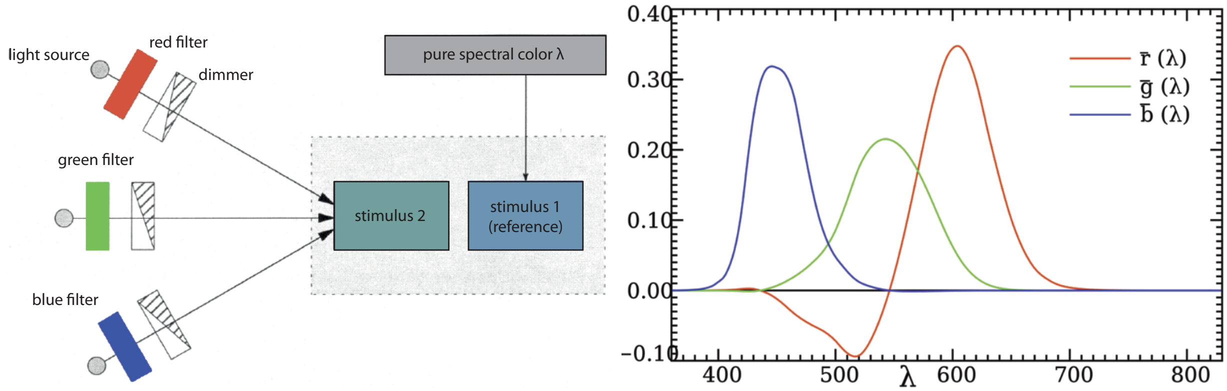


Figure 2: Left: CIE 1931 Experiment Setup. Right: Experiment Results.

i) In 1931 CIE performed an experiment. Figure 2 (left) shows a diagram of the setup. With the help of this diagram explain briefly how the experiment was conducted and the motivation behind it.

4 pts.

ii) In Figure 2 (right) the results of the CIE experiment are shown. For a point (x, y) on one of the curve, please explain briefly what x and y represent.

2 pts.

iii) In the range of 450-550 on the x-axis we observe that the red curve has negative values. Explain briefly how and why this is possible given the fact that negative light does not exist.

2 pts.

c) **CIE Color Spaces**

Consider the 3 following primaries given in the CIE xyY color space:

	x	y	Y
c_1	0.1	0.8	12
c_2	0.2	0.05	10
c_3	0.7	0.25	26

i) Explain briefly the perceptual meaning of the xy and Y axes in the CIE xyY color space.

2 pts.

ii) Figure 3 is an empty CIE xy Chromaticity diagram. Plot c_1, c_2 and c_3 on the diagram.

1 pt.

iii) Draw the isoline of constant saturation passing through c_1 on the diagram.

1.5 pts.

iv) What is the name of the connection between 770nm and 380nm? What do the points on that line represent?

1.5 pts.

v) Provide the transformation formulas from CIE XYZ to CIE xyY and from CIE xyY to CIE XYZ.

2 pts.

vi) Compute the sum of the three primaries c_1, c_2 and c_3 in the XYZ color space. Plot the resulting color on the CIE xy Chromaticity diagram (Figure 3).

2 pts.

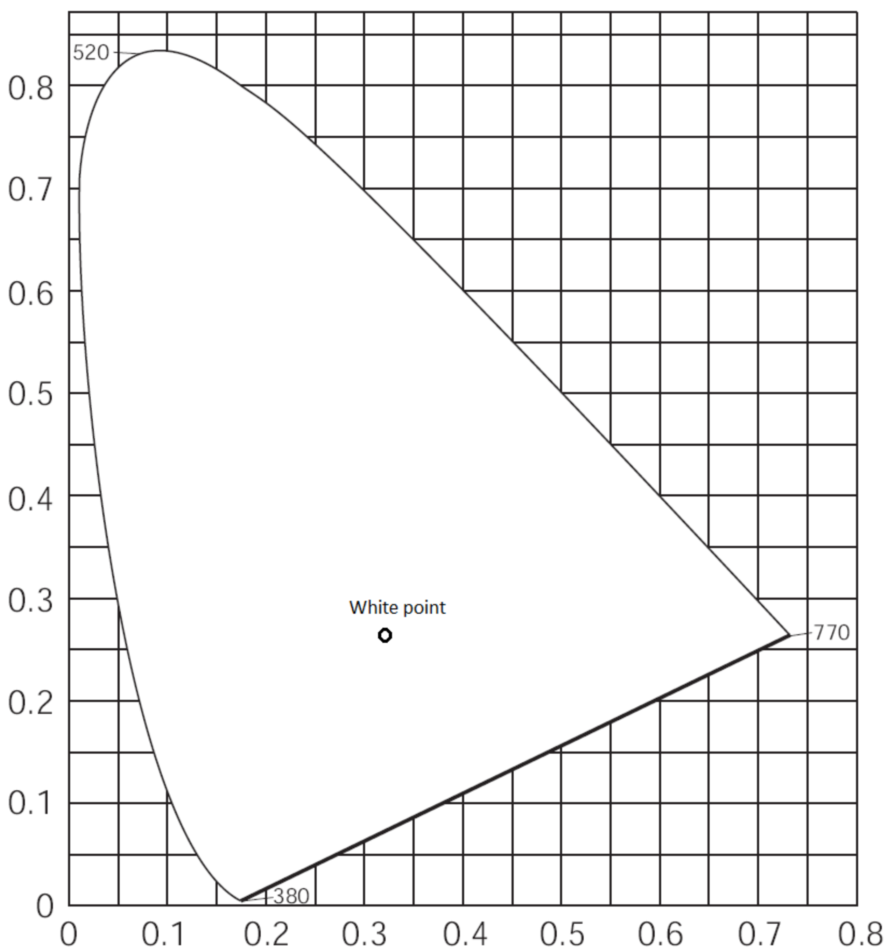


Figure 3: CIE xy Chromaticity diagram

Question 10: Transformations (28 pts.)

a) **Matrices**

i) In 3D, give the homogeneous coordinates matrices describing the following transformations :

- A rotation by $\pi/4$ around $(0,1,0)^T$.
- A translation by $(1,2,3)^T$.
- A scaling by a factor 4.

3 pts.

ii) The following matrices describe 3D transformations. For each, describe that transformation.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3 pts.

b) **Matrices**

Let $\mathcal{P}$ be a plane with a normal vector $\mathbf{n}$. $\mathcal{P}$ is transformed to $\mathcal{P}'$ using the transformation matrix $\mathbf{M}$. This means that $\forall P \in \mathcal{P}, \exists! P' \in \mathcal{P}' / P' = \mathbf{M}P$.

i) Prove that $\mathbf{Mn}$ is not necessarily a normal to $\mathcal{P}'$.

2 pts.

ii) Prove that $\mathbf{n}' = (\mathbf{M}^{-1})^T\mathbf{n}$ is normal to $\mathcal{P}'$.

2 pts.

iii) Does a rotation preserve orthogonality?

1 pt.

c) **Quaternions**

i) Let $P = (2,0,0)^T$. Describe this point as a quaternion $\mathbf{p}$. **1 pt.**

ii) Let R_1 be the rotation by $60°$ around $(0,0,1)^T$.
Describe this rotation as a quaternion $\mathbf{q_1}$. **1 pt.**

iii) Let R_2 be the rotation described the the following homogeneous coordinates matrix.

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Describe this rotation as a quaternion $\mathbf{q_2}$. **1 pt.**

iv) We define P_1 and P_2 as the transformations of P by R_1 and R_2 respectively.
Let $\mathbf{p_1}$ and $\mathbf{p_2}$ be the quaternions describing P_1 and P_2 respectively.
Using quaternion operations, compute $\mathbf{p_1}$ and $\mathbf{p_2}$.
What are the coordinates of the points P_1 and P_2? **5 pts.**

We now want to average the rotations R_1 and R_2.

v) Compute $\mathbf{q_{avg}} = \frac{1}{2}(\mathbf{q_1} + \mathbf{q_2})$.
Is $\mathbf{q_{avg}}$ a valid rotation? Explain your answer. **1 pt.**

The spherical interpolation between two quaternions $\mathbf{q_a}$ and $\mathbf{q_b}$ using a parameter $t \in [0;1]$ is defined as such:

Let θ be a positive angle satisfying $\mathbf{q_1} \cdot \mathbf{q_2} = cos(\theta)$ (where $\cdot$ denotes the quaternion dot product).

The spherical interpolation is given by the following formula:

$$Slerp(\mathbf{q_a}, \mathbf{q_b}, t) = \frac{sin((1-t)\theta)}{sin(\theta)}\mathbf{q_a} + \frac{sin(t\theta)}{sin(\theta)}\mathbf{q_b}$$

vi) Compute $\mathbf{q_{slerp}} = Slerp(\mathbf{q_1}, \mathbf{q_2}, 0.5)$.

3 pts.

vii) Using quaternion operations, compute the position of P_{slerp}, the transformation of P by $\mathbf{q_{slerp}}$. **2 pts.**

viii) By making a drawing showing P_1, P_2, P_{slerp}, and $\frac{1}{2}(P_1 + P_2)$, explain the advantage of using Slerp to interpolate quaternions. **3 pts.**

Question 11: Rendering and Geometry (34 pts.)

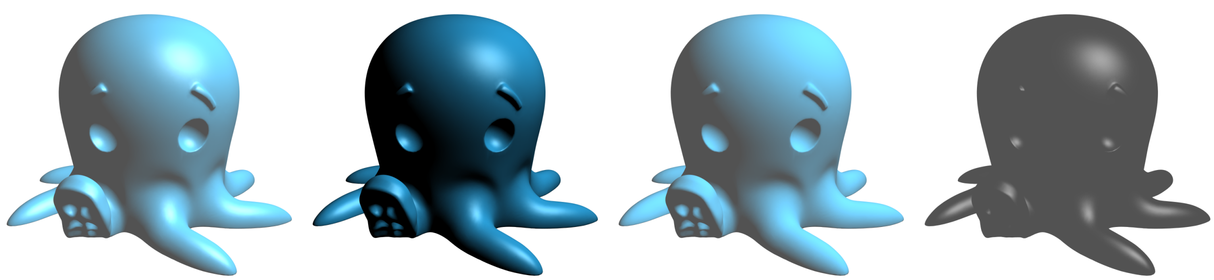
a) **Raytracing**
You would like to render a set of pool balls using raytracing, so you set up scene by placing a camera and multiple spheres that represents your pool balls. To generate your image, you then shoot a ray for each pixel and check for an intersection with the spheres.

i) Assuming the camera is placed at $\mathbf{c} = (0, 0, 10)$, and you shoot a ray defined by the camera position and the direction $\mathbf{d} = (1, 2, -5)$, write down the equation that defines the intersection points between the ray and a sphere with radius 5 placed at $\mathbf{s} = (3, 3, -2)$. The parameter of the ray equation should be the unknown in your equation. **4 pts.**

ii) The pool balls that you want to render are partly reflective. Explain briefly how reflections can be incorporated into a raytracer. **3 pts.**

iii) You realize that your renderer is slow, and the problem seems to be that for every ray, you compute an intersection with every pool ball. Explain briefly how you can avoid computing unnecessary intersections and thus accelerate your renderer, and what steps before and during the rendering you have to add. **4 pts.**

b) **Shading and Shadows**



reference 1) 2) 3)

Figure 4: A rendering with the full Phong model (left) and three renderings with one of the components in the Phong model missing.

i) Figure 4 shows a rendering of an octopus that uses the Phong illumination model, as well as three renderings, each with a different component of the illumination model turned off. First, list all of the components of the Phong illumination model. Then, in each of the renderings 1) to 3), list the components that are still active. **4 pts.**

ii) You render an object with 2000 vertices on a display with a resolution of 800 $\times$ 600 pixels, once using Gouraud shading, and once using Phong shading. For each of these cases, the illumination model (e.g., the Phong illumination model with its three components) will be evaluated a different total number of times in your shaders. Give an upper bound to how many times the illumination model will be evaluated *at most* for each of the two variants. The upper bound should be as small as possible, given the provided information. **2 pts.**

iii) *Light maps* are an efficient way to add complex shading effects to a scene. Instead of computing the shading in real time, an offline renderer (using, e.g., raytracing) is used to compute the lighting and shadows, and for each shaded object, a light map

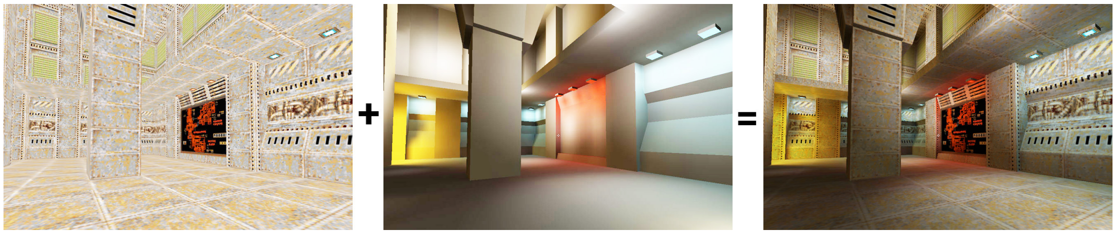


Figure 5: A non-shaded rendering (left) is combined with a light map (middle) to produce the final shaded rendering (right).

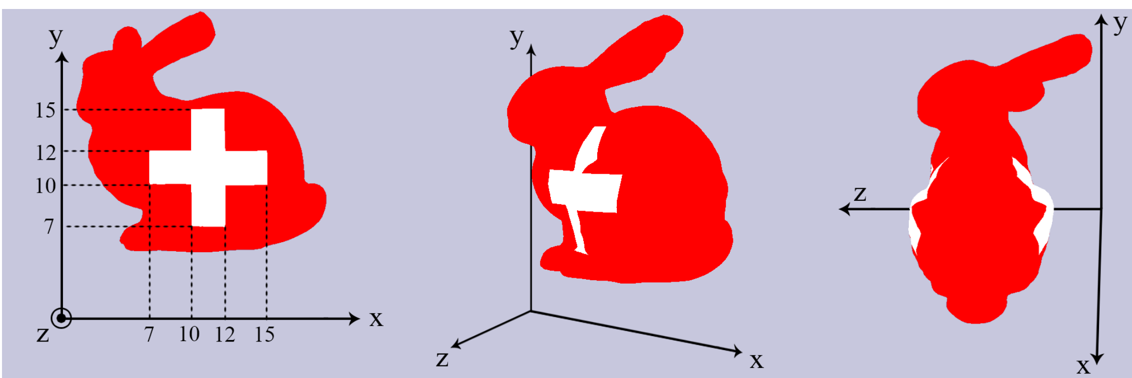
texture is generate from this rendering. During the real time rendering, the light map texture is then laid over the existing texture to produce a shaded scene without any additional shading (see Figure 5).
Briefly explain an advantages and a disadvantages of light maps over real time shading and shadow maps/shadow volumes. **4 pts.**

iv) In real time renderers, *transparency* in is often approximated using alpha blending. However, if several overlapping transparent objects are rendered, objects might disappear. Why is this the case? How can you prevent this from happening? **3 pts.**

c) **OpenGL**

Complete the following shaders to reproduce the "Swiss Bunny" rendering below. The image on the left shows the bunny in object space. **10 pts.**

Hint: You want the color of the fragments to depend on the position of the vertices.



Vertex shader

```
uniform mat4 projectionMatrix;
uniform mat4 modelviewMatrix;

in vec3 inPosition;

// TODO 1

void main()
{
  gl_Position = // TODO 2

  // TODO 3
}
```

Fragment shader

```
// TODO 4

void main(void)
{
  // TODO 5
  gl_FragColor = // TODO 6
}
```

Question 12: Animation and Physics (8 pts.)

a) **Animation**
Characters are often animated using an underlying skeleton, defined as a set of bones and joints. The individual bones of the character move by rigid transformations, generating a motion and also moving the *skin* of the character. A model that uses *linear blend skinning* will typically have non-rigid deformations around the joints of the skeleton, even though the bones only undergo rigid transformations. Briefly explain how this is possible, and how you compute the deformed position $\mathbf{x}'$ of a vertex $\mathbf{x}$ at the joint between two bones with (homogeneous) transformations $\mathbf{T}_1$ and $\mathbf{T}_2$. **3 pts.**

b) **Rigid-Body Physics**
The kinematic state of a rigid body can be described by the the position $\mathbf{x}(t)$ of the center of mass, the linear momentum $\mathbf{P}(t)$, the rotation $\mathbf{q}(t)$ (quaternion), and the angular momentum $\mathbf{L}(t)$. Write down the *ordinary differential equations* (ODEs) that govern the dynamics of a rigid body with a single force $\mathbf{F}_i(t)$ applied to the point $\mathbf{r}_i(t)$. You can assume that the mass m and moment of inertia tensor $\mathbf{I}(t)$ of the rigid body are known at any time. If you use any other quantities than the ones mentioned here, you have to define them first in terms of previously defined quantities. **5 pts.**