Prof. Dr. Thomas Schultz
Mohammad Khatami (khatami@cs.uni-bonn.de)
Ikram Jumakulyyev (ijumakulyyev@cs.uni-bonn.de)

University of Bonn
B-IT/Computer Science II
April 24, 2020

Summer term 2020

Visual Data Analysis

Assignment Sheet 2

Solution has to be uploaded by May 3, 2020, 23:59
to https://uni-bonn.sciebo.de/s/O7lPWUYox8qOuvp with password `vda.2020`

Please bundle the results (as PDF) and scripts (*.py/*.ipynb files) in a single ZIP file. Submit each solution only once, but include names and email addresses of all team members in the PDF and each script. Name the file `vda-2020-xx-names.zip`, where `xx` is the assignment sheet number, and `names` are your last names.

If you have questions concerning the exercises, please subscribe to our mailing list via the lecture webpage and write to: vl-scivis@lists.iai.uni-bonn.de.

Exercise 1 (Working with HSV Color Space, *15 Points*)

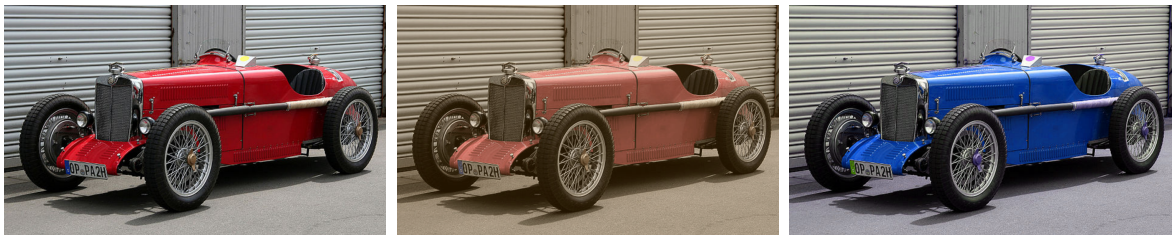

Figure 1: The input image (left), and results of two of the transformations we are asking you to perform on it.

We discussed different color spaces in the lecture. In this exercise, you will use them to perform different operations on the color image `oldtimer.png` which is provided on our lecture homepage. Please use Python for your solution. It is recommended to use routines from `matplotlib`, which allows you to do most subtasks in very few and simple lines of code. However, *do not* use an image processing program like Photoshop. Your tasks are as follows:

a) Read the given `oldtimer.png` image. (1P)
b) Convert the color image to HSV color space, and create a grayscale version by retaining only the value. (3P)
c) Think about another way to convert color image to grayscale. Its result should look reasonable (with bright regions appearing bright), but it should differ from the one you obtained from the HSV color space. Implement and briefly justify your method. (3P)
d) Output a version of the image in which saturation has been reduced by 50%. (2P)
e) Create a simple "aged photograph" effect by globally blending (via simple weighted averaging or some other blend mode of your choice) the half-desaturated image from the previous step with some amount of brown. (2P)
f) Rotate the hues of the original image so that the car obtains a different color. Provide two images in which the car is blue, and has another color of your choice. (4P)

Exercise 2 (Face-based Luminance Matching, *10 Points*)

As a regular part of the exercises, we will ask you to read research papers from the visualization literature and extract specific information from them. Our first example is the paper `kindlmann-luminance-2002.pdf`, which you can find on the lecture webpage. Please read it and answer the following questions, using 1-2 sentences for each. Answer them in your own words. **We will not grant even partial credit for copy-pasted text.**

a) What is the Helmholtz-Kohlrausch effect? (2P)

b) Why are the authors proposing to use images of faces? (2P)

c) To what alternative method do the authors compare their newly proposed one in the user study? (2P)

d) Based on the result of the user study, what is the advantage of the newly proposed method? (2P)

e) Why do the authors have to know the monitor gamma while creating a colormap based on the result of the user study? (2P)

Exercise 3 (Color in Visualization, *10 Points*)

a) Design a color map for different files on a computer. Each file should be represented by a color that conveys both the file type (text / image / audio / video / other file) and the file size (amount of memory). Decide on a suitable color scheme and create a color legend that illustrates it. Justify your choices. (4P)

Hint: You do *not* have to write a program that searches your computer for files and visualizes them, it's enough to create the corresponding color legend. For this task, you may use any software of your choice, e.g., a Python script or an image processing program.

b) In Section 2.2 of the lecture slides, equations are provided to convert from RGB to HSV color space. Derive the inverse mapping, which should convert from HSV to RGB. Please specify intermediate steps. (4P)

Hint: We we give full credit for correctly treating the case in which $H \in [60°, 180°]$, since all other cases can be dealt with in analogy.

c) Name an application for which you would prefer HSV color space over CIEluv, and one for which you would prefer CIEluv over HSV. Briefly justify why. (2P)

Exercise 4 (Plotting Categorical Data, *15 Points*)

In the previous sheet, we asked you to start reading the official seaborn tutorial at https://seaborn.pydata.org/tutorial.html. Today, please continue with the second chapter ("Plotting with categorical data"). Afterwards, please answer the following questions, making use of additional resources if needed:

a) What is an advantage of stripplots over swarmplots? What is an advantage of swarmplots over stripplots? (2P)

b) What is meant by IQR? (2P)

c) What would be a reason to use violinplots instead of, or in addition to boxplots? (2P)

Please use pandas and seaborn to inspect the numerical attributes of a chronic kidney disease dataset, which is available as an Excel file on the lecture homepage:

d) Read the dataset using pandas `read_excel` function. (1P)

e) Use `pandas.melt` to transform the data from "wide" to "long" format, using `class` (indicating `ckd` for chronic kidney disease or `notckd` for its absence) as the identifier variable. (2P)

f) For each numerical attribute, such as age or blood pressure, create two boxplots side-by-side. One should show the attribute's distribution among patients suffering from chronic kidney disease, the other one from patients who do not suffer from the disease. *Hint:* Due to the different numerical ranges, you will have to disable sharing of y axes between plots of different attributes. (4P)

g) Based on viewing the plots, name an attribute that appears to be highly indicative of chronic kidney disease, and one that seems to be mostly unrelated to it. (2P)

Good Luck!