TxMM A2 - Multimedia Clustering

Students: \*student names and student numbers\*  
A2-MM Group: \*group number\*

*Note: make sure to hand in a pdf of this file*

# First impression

**Question 1:** Look at the images. Make three observations concerning the objects and scenes that you see depicted in the images (in computer vision, what people see in images is referred to as the “semantic content” of the image or "semantic properties").

**Answer:**

**Question 2:** Now, look at the images again, and make three observations that are not related to objects and scenes, but rather are related to the style and quality of images.

**Answer:**

# Clustering on simple features

**Task 1:** Implement the function average\_color in the cell below using the information provided in the comments.

(Hint: Remember that the original image can be accessed via image\_dict['original'], and that you can convert this original image to a numpy array. (As a sanity check, make sure that you understand why the vector representation consists of three components in this case.))

**Your code:**

def average\_color(image\_dict):

'''

A function to compute the average RGB value of an image.

First, average over rows to obtain an average value per column.

Then, average over the resulting values to obtain one average value per color

channel.

:param image\_dict: The dictionary containing the loaded image

:return: A 3-dimensional np array: 1 average per color channel

'''

raise NotImplementedError() # Your code goes here

**Question 3:** Subsequently look at the images with indices 102, 153, 245, 439 and 555 and their average color by changing the index in the cell above. Then discuss and decide: Is this a good representation for color images or not (and why)? Describe a potential issue that you notice.

**Answer:**

**Question 4:** Inspect the image montages. Do all clusters make equal sense to you, intuitively? Look at the 3-dimensional scatter plot again. Do you see a reason why/why not?

**Answer:**

**Task 2:** Change the color histogram function below so that it uses 32 bins per color channel. Then run the second cell and inspect the results.

**Your code:**

def color\_histogram\_32bins(image\_dict):

'''

Compute the normalized color histogram binned into 32x32x32 bins from the RGB image.

:param image\_dict: The dictionary containing the loaded image

:return: A 32768-dimensional np array

'''

# extract a 3D color histogram from the RGB color space

im = image\_dict['cv2']

hist = cv2.calcHist([im], [0, 1, 2], None, [3,3,3], [0, 256, 0, 256, 0, 256])

# normalize the histogram

hist = cv2.normalize(hist,hist)

# return the flattened histogram as the feature vector

return hist.flatten()

**Question 5:** Do you think that more bins helped us in the clustering? Why?

**Answer:**

**Task 3:** Perform clustering with 12 clusters on the chist\_32bins\_feature\_vectors and display the montages for these clusters in the cell below.

**Your code:**

y\_kmeans\_chist\_12 = None # Your code goes here

show\_images\_in\_clusters(y\_kmeans\_chist\_12, sample\_pathnames, sample\_images)

y\_kmeans\_chist\_24 = None # Your code goes here

show\_images\_in\_clusters(y\_kmeans\_chist\_24, sample\_pathnames, sample\_images)

**Question 6:** Which are the major cluster changes when increasing the number of clusters to 12 and then to 24? What does this tell us about our dataset's images?

(Hint: You don't need to point out every small detail. Please try to answer this question at a high-level.)

**Answer:**

**Question 7:** Looking at the semantic content of the clusters, which of the choices of the number of clusters do you find to be best-suited for our dataset using the color histogram feature vectors? Why?

**Answer:**

**Question 8:** Do you think that all images are in the 'correct' cluster? Do you think 8 is a good number of clusters for HOG features, considering the images' semantic content? Why?

**Answer:**

# Clustering using neural representations

**Task 4:** Now systematically vary the number of clusters in the cell below. Try out both more and fewer clusters. Then answer the questions below. (*No need to copy anything here*)

**Question 9:** Which numbers did you try out? Which number of clusters worked best for you? How did you make this decision?

**Answer:**

**Question 10:** What does this tell us about the neural representations?

**Answer:**

**Question 11:** Reflect on the "clustering bias". For each of the four image representation approaches, answer the following four subquestions:

1. Which semantic properties play a decisive role when clustering the images in the dataset using this representation? By which semantic properties are the images grouped? Remember "semantic properties" refer to what people ("human users") see in images.
2. Which semantic properties play no decisive role?
3. Why? What causes the phenomena you observed in the first two subquestions?
4. When (i.e. for which kind of clustering problem) would you use this image representation?

**Answer:**

## Average color:

1. Semantic properties: Decisive role:
2. Semantic properties: No decisive role:
3. Why?:
4. When to use:

## Color histograms:

1. Semantic properties: Decisive role:
2. Semantic properties: No decisive role:
3. Why?:
4. When to use:

## HOG:

1. Semantic properties: Decisive role:
2. Semantic properties: No decisive role:
3. Why?:
4. When to use:

## Neural representations:

1. Semantic properties: Decisive role:
2. Semantic properties: No decisive role:
3. Why?:
4. When to use:

**Question 12:** At the beginning of the assignment, we told you that the data set contained Paris landmarks. However, now you have carried out a clustering study, you have gained more insight into the semantic content of the images in the data set. What have you discovered?

**Answer:**