Ethan Chiang 916154228
Kishan Patel 914983103

PS 3

ECS 174
November 12, 2019

I. Short Answer Problems

1. What exactly does the value recorded in a single dimension of a SIFT keypoint descriptor signify?

   A: The values are the magnitude of bins of gradient orientations of the descriptors. And they are part of a histogram that describe each sub-patch in the keypoint descriptor.

2. A deep neural network has multiple layers with non-linear activation functions (e.g., ReLU) in between each layer, which allows it to learn a complex non-linear function. Suppose instead we had a deep neural network without any non-linear activation functions. Concisely describe what effect this would have on the network. (Hint: can it still be considered a deep network?)

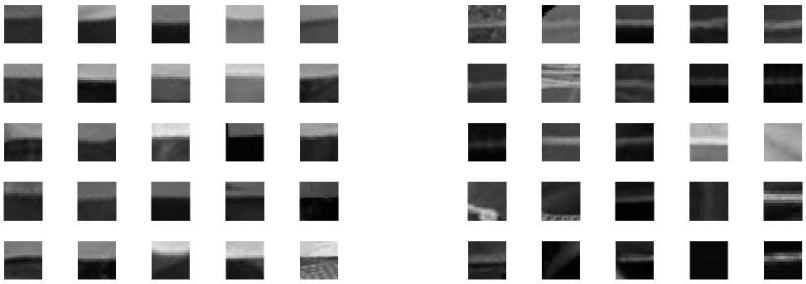
   A: A deep neural network without any non-linear activation functions will results in a deep neural network that is only able to learn a linear function.

II. Programming problem: content-aware image re-sizing

1. Script *raw_descriptor_matches.m* allows user to select a region and then match descriptors in that region to descriptors in the second image based on Euclidean distance in SIFT space. With the two images in *twoFrameData.mat*. We modified the *selectRegion.m*, named *selectRegion_display.m* so that it also display the selected region. The threshold is set to be 0.4 in this case to have a better result. As the result, it didn't make much mistakes to highlight the desire descriptors (ie. the fridge).

2. Script *visualize_vocabulary.m* samples 150 random SIFT descriptors from every 20 provided frames to get at lease the minimum 25 patches per cluster after ruining K-means

with k = 1500. The below is the results top 25 patches for the 1st and 2nd most common visual words.



The most common word is am edge and the second most common is a line. This result makes sense since edges and lines are common elements in real-world environment.

3. To make *full_frame_queries.m* works I have the helper function *getHistogram.m* to pre-generate histogram for every scene. *full_frame_queries.m* will then access the histogram for the target frame, compare it to other frames and assign a score. Then display the top 5 ranking frame as below:

I picked the target frame with characters that also appears in other frames in the same outfit, in order the verify my result easily.

4. *region_queries.m* uses *selectRegion.m* to get the descriptors then create histogram with them. Then we use the same method as part 3 to score and rank all frames with the

target frames to display the top 5 results.

The first two are the result of selecting a piece of clothing that have a unique pattern and the results are great. The three one is a region of painting, it did great but the includes a frame without the painting. The last one is a example of fail case, since the fridge region include rich distinct features and can easily be confused with the shirt with lots of thing things going on.

5. *compare_bow_and_deep.m* uses frames *friends_0000004503.jpeg* and *friends_0000000394.jpeg* to serve as queries. Overall, the pre-trained deep convolutional neural network (CNN) performs better than the SIFT bag-of-words. This is mainly because the CNN is better at learning the whole image than our algorithm use the SIFT space. The SIFT space focus on distinct features but the CNN was trained with the image as a whole and the colors, which is something our algorithm ignores. We can see the effect with the image that contains a woman in purple, the image is fairly plain that is why our algorithm failed to get simpler image. And it also mistaken the lights as the dress the in the other image. Overall, our algorithm is good when a frame contains enough distinct features, but CNN will work even with only a few.


Query Image
friends_000004502.jpeg
friends_000004499.jpeg
friends_000004523.jpeg
friends_000004524.jpeg
friends_000004612.jpeg
friends_000004611.jpeg
friends_000004501.jpeg
friends_000004613.jpeg
friends_000004619.jpeg
friends_000004617.jpeg


Target Image
Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
Rank 8
Rank 9
Rank 10


Query Image
friends_0000000385.jpeg
friends_0000000382.jpeg
friends_0000000381.jpeg
friends_0000000388.jpeg
friends_0000000395.jpeg
friends_0000000386.jpeg
friends_0000001149.jpeg
friends_0000000380.jpeg
friends_0000001150.jpeg
friends_0000000387.jpeg