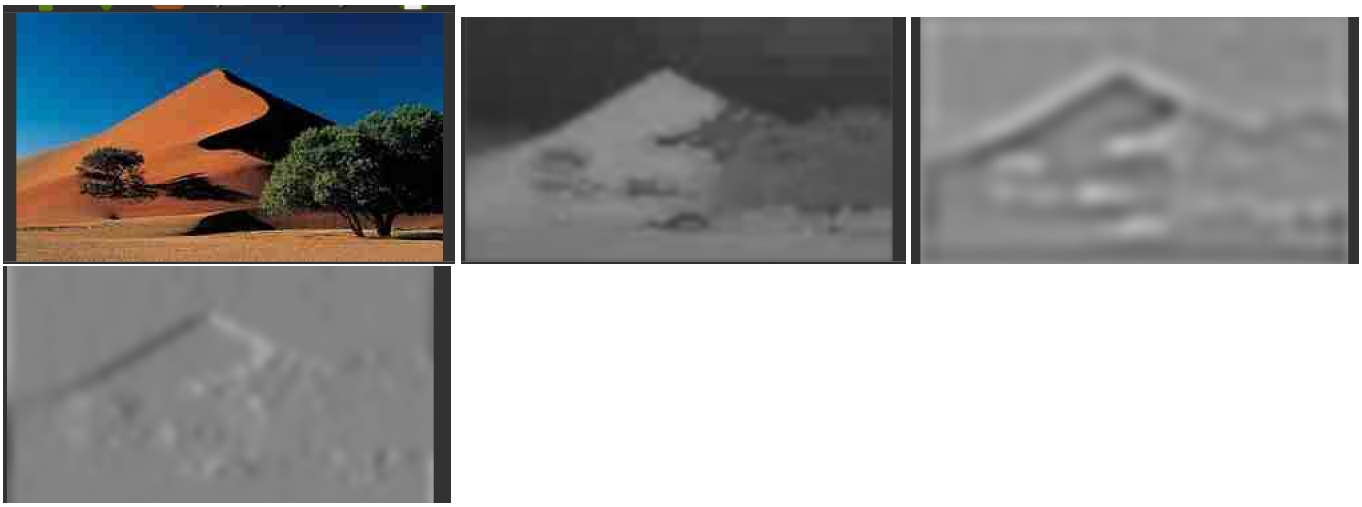
Homework 4 Writeup

Jixiang Li

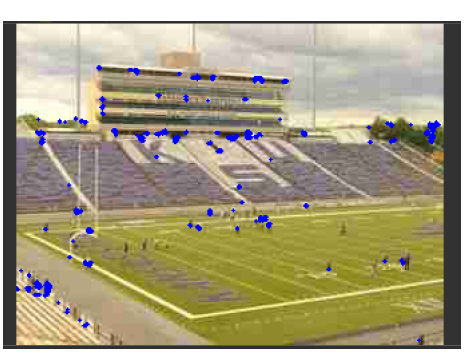
Collaborators: Joy Song

2.1



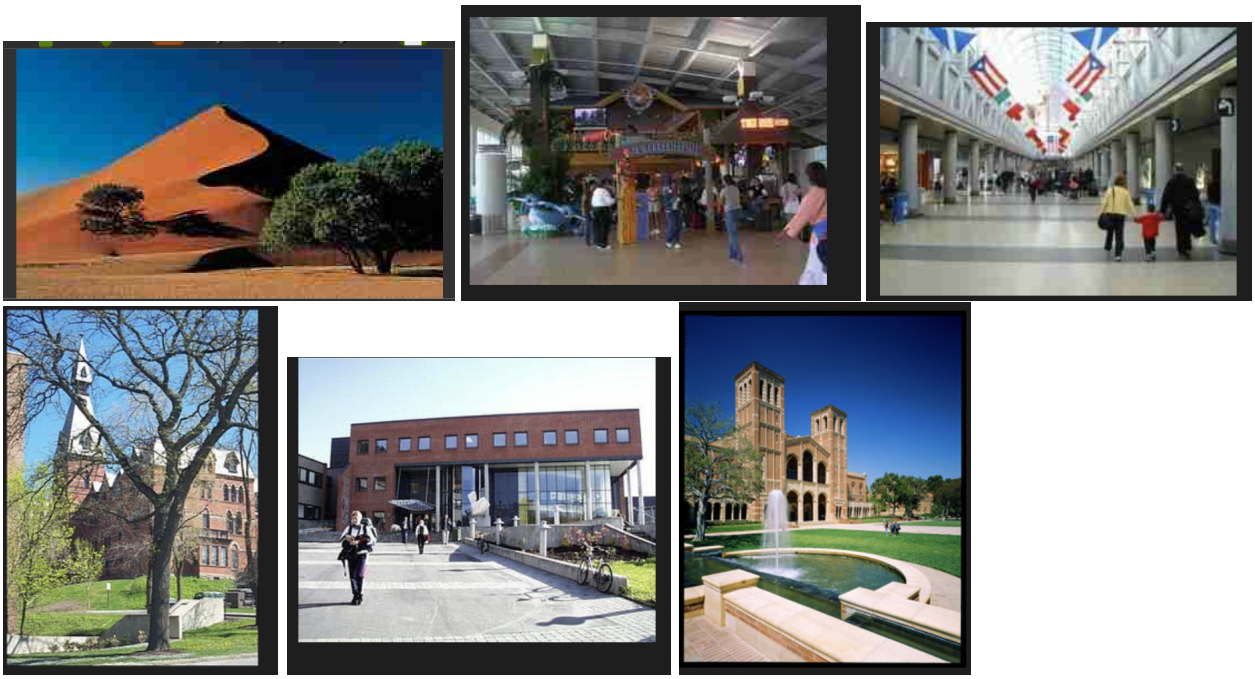
CIE LAB is a 3d color space in which the differences in value correlate to the amount of change we as humans can see the lab or L*a*b in CIE lab stands for 3 different values used to calculate the color and color differences. L stands for lightness on a scale from 0 to 100 which goes from white to black. A and B both represent chromacity. Unlike L however, neither are limited on a scale. A ranges from green to red while B ranges form blue to yellow. Because CIE Lab tries to measure the difference in color, it is better suited for bag of words, where we try and identify small features in images, where slight changes in color may make a big difference. For artifacts, longer straight lines seem to be detected more easily. you can see that in this example because the big sand dune is more visible than things like the tree.

2.2

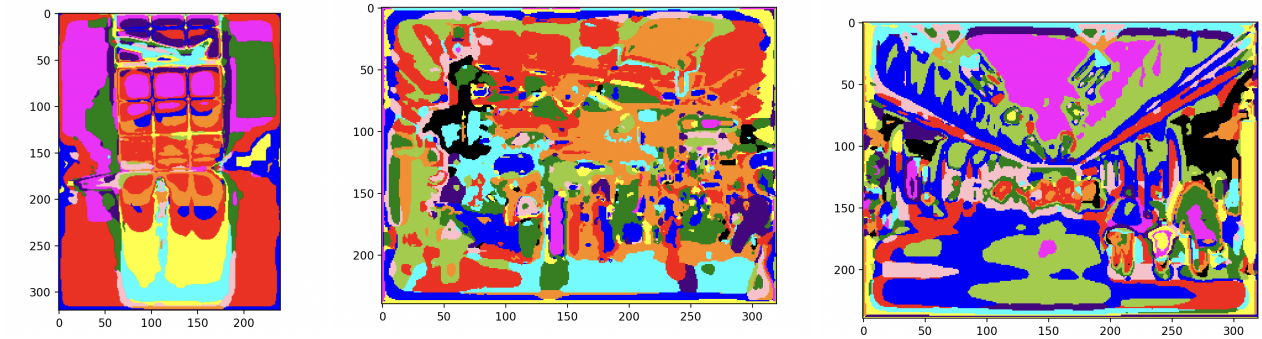


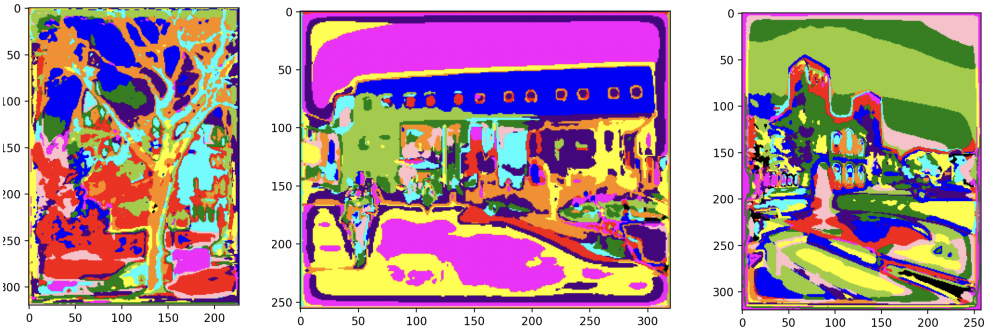
3.1

Original Pictures

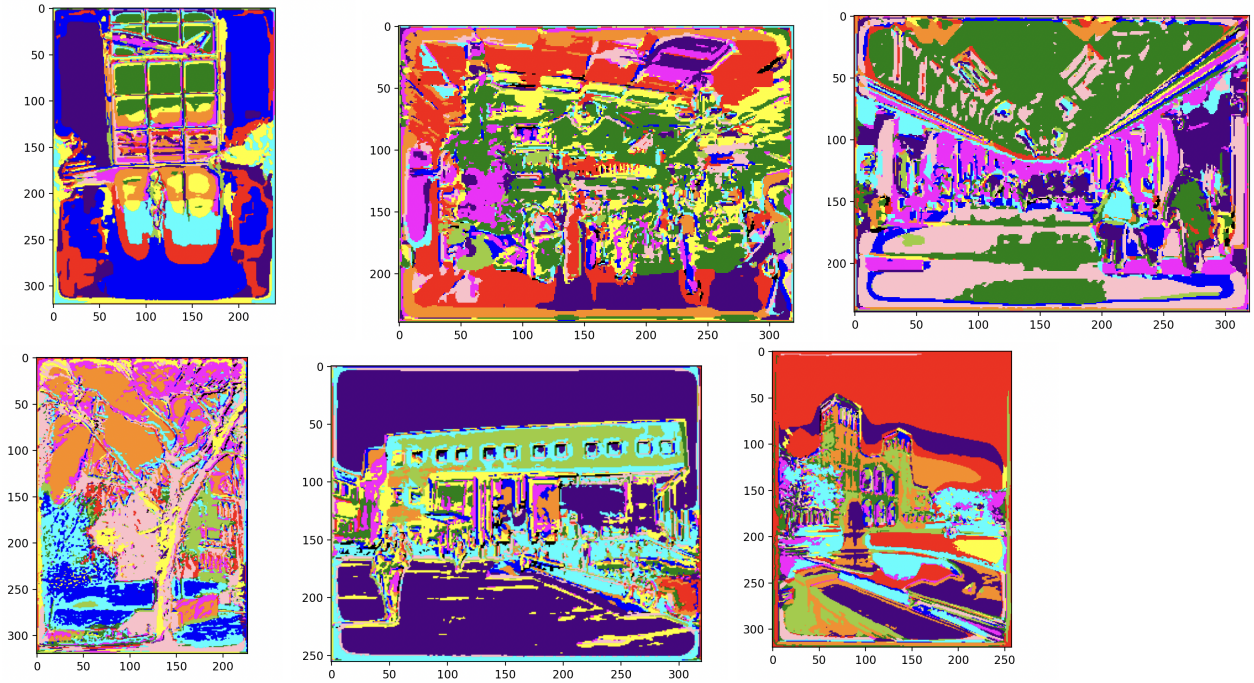


Random points





Harris



Yes the visual words are capturing semantic meaning. I think the harris word maps are much better as they seem to more detailed alongside edges. For example in the first image of the airport with the kid in the middle, the outline of the kid is much more defined in the Harris version. The harris versions also visibly have more sections; the chunks in the random point versions are much larger, which means less semantic meaning.

4.2

```
random euclidean matrix:
 [[11  5  5  1  1  4  2  3]
 [ 1  8  4  0  4  2  2  1]
 [ 4  3  8  3  3  1  1  0]
 [ 1  0  1 11  3  2  6  1]
 [ 0  2  1  1  5  1  1  0]
 [ 0  2  0  1  0  4  1  0]
 [ 1  0  0  2  4  3  4  1]
 [ 2  0  1  1  0  3  3 14]]
random euclid metric: 0.40625

random chi matrix:
 [[11  4  4  0  0  2  2  3]
 [ 1 12  5  0  4  1  2  0]
 [ 5  2 11  4  3  3  4  0]
 [ 0  0  0  9  1  2  2  2]
 [ 0  1  0  0  9  1  1  0]
 [ 0  1  0  2  0  6  1  0]
 [ 0  0  0  4  3  3  7  1]
 [ 3  0  0  1  0  2  1 14]]
random chi metric: 0.49375

harris euclidean  matrix:
 [[ 8  2  8  3  2  2  3  1]
 [ 1 10  3  5  1  3  0  1]
 [ 6  3  7  0  3  1  1  0]
 [ 2  1  1  4  2  4  7  3]
 [ 0  2  1  1  8  1  3  0]
 [ 0  1  0  2  0  4  1  0]
 [ 1  1  0  4  4  3  4  1]
 [ 2  0  0  1  0  2  1 14]]
harris euclidean metric: 0.36875

harris chi matrix:
 [[14  3  3  3  0  0  1  3]
 [ 2  9  3  2  2  3  0  0]
 [ 1  4 12  0  3  2  2  0]
 [ 0  0  0  8  0  2  3  1]
 [ 0  3  2  0 10  1  2  0]
 [ 0  1  0  4  0  7  1  0]
 [ 0  0  0  3  5  3  8  0]
 [ 3  0  0  0  0  2  3 16]]
harris chi metric: 0.525
```

```
harris chi matrix:
 [[13  1  2  2  1  3  3  2]
 [ 2 15  4  1  2  4  0  0]
 [ 2  3 13  0  3  2  1  1]
 [ 1  0  0  8  0  2  1  0]
 [ 0  0  1  1 12  1  4  0]
 [ 0  0  0  2  0  3  0  0]
 [ 0  1  0  5  2  3 11  0]
 [ 2  0  0  1  0  2  0 17]]
harris chi metric: 0.575
best k: 40
```

The two dictionaries don't differ as much as I thought it would, which was surprising, as I thought the harris detector would perform much better. The random euclid actually performs better than the harris euclid which I thought was really surprising, but harris outperforms random with chi squared. The chi squared distance metric outperformed euclidean in both cases This is most likely because the euclidean distance does not work as well in dimensions larger than 2, while chi squared is less prone to high dimensions, making chi squared more suitable for the given task since there are multiple layers of filters.

The best value of k in this case was 40. A larger value of k is not always better. A large k value can significantly increase the the time it takes to compute as well as potentially creating too much noise. In this example if k was really large, say 100, then we could potentially be considering points where the distance metric is large, which could result in inaccurate predictions, since those points would be considered noisy. Ties are broken by looking at the order in which they appeared in the original array.