

Homework #4

Issued: 02/19/2024

Due: 03/29/2024

Problem 1: Texture Analysis

(a) Texture Classification - Feature Extraction

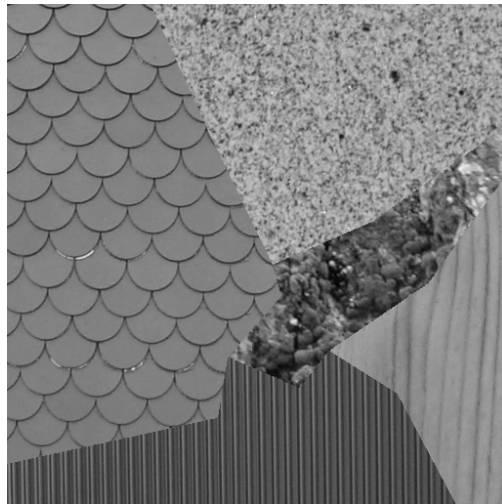
Actually it is hard to say which one has the strongest discriminant power: some of them seems can classify correctly in some cases, but also has a lot of false cases. For example, S5*R5. However, it is easy to find that the weakest is L5*L5. The feature remains very high and varies little in all 36 pictures.

(b) Advanced Texture Classification — Classifier Exploration

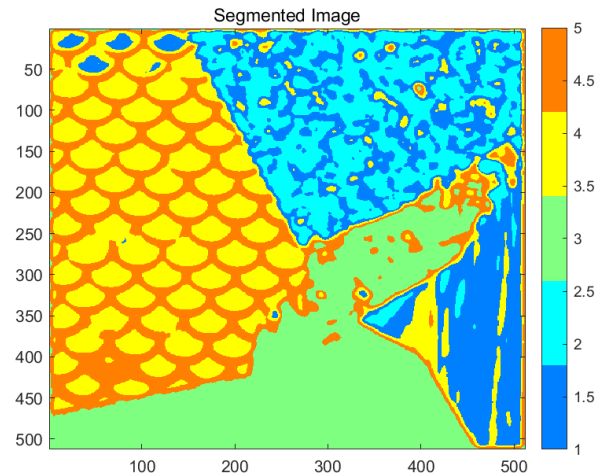
- 1. Unsupervised Learning** In the 25-D energy features' kmeans unsupervised learning, test cases are divided into 4 groups: {1,9},{2,3,4,5,6,7, 12},{8,11} and {10}. In this cases, group 1 is classified as Brick, group 2 is classified as Grass, group 3 and 4 should be calssified as Blanket or Stones. The final error rate for 25-D energy k-means learning is 50% Then we go to the pca 3-D features' unsupervised leraning. The test cases are divided into 4 groups: {1, 8, 9, 11},{2,4,6,7,12}, {3,5} and {10}. In this case, Group 1 should be classified as Brick, Group 2 as Grass, Group 3 as Stones and Group 4 as Blanket. The finial error rate for 3-D kmeans unsupervised learning is 41.67%, which is a bit lower than without pca transferred. In this case, the application of pca dimension reduction decreases the error rate by one case.
- 2. Supervised Learning** In supervised learning, when dimension reduction is not applied, the error rate is 25%. Picture 2,3,4,5,6,7,8,9 and 12 are classified percisely. However, when applied pca to reduce dimensions, the error rate increases to 58.3% , with only picture 5,6,7,10 and 12 classified correctly. When compaired to unsupervised learning, using 25 dimension in supervised learning is the learning ide with the highest accuracy, while using pca in supervised learning has the lowest.

Problem 2: Texture Segmentation

(a) Basic Texture Segmentation



(a) Original Mosaic Image



(b) Texture Segmentation Image

Figure 1: Texture Segmentation

(b) Advanced Texture Segmentation

In this case, we applied pca dimension-reduction to reduce the interruption of unimportant features. Besides, a post-processing technique that could stop taking small regions into consideration. Besides, edge enhancement is also applied. However, the segmentation still could not recognize the pattern on the left part of the Mosaic Image, where dark lines on light ground forms the texture. It still recognizes the lines as a new texture. An idea to solve the problem is to use more improved algorithm, including SIFT or CNN learning.

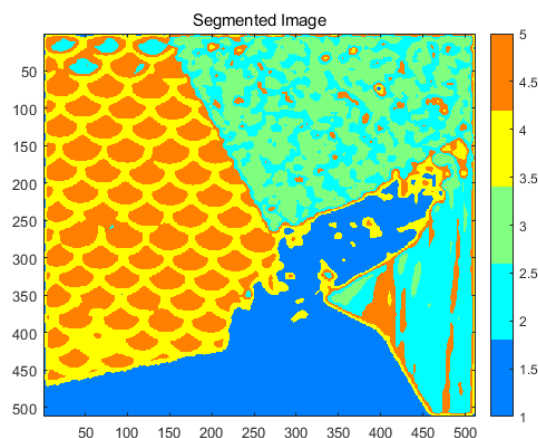


Figure 2: Segmentation Image after Improved

1 Problem 3: SIFT and Image Matching

(a) Salient Point Descriptor

1. From the abstract, we learn that the SIFT could prevent interrupt from image rescaling and rotation, and has robustness againsts substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. Among them, image rescaling, rotation and distortion are geometric modifications.
2. By taking Difference of Gaussian into consideration, each feature points are compared not only to the pixels next to them on the same scale, but also the points around it before and after rescaling. In this method, SIFT achieves Robustness of rescaling. For rotation, SIFT adapts a method that could calculate a histogram of orientation. So it could pick a dominant orientation no matter how the image is rotated.
3. SIFT enhances its robustness to illumination change by normalizing the feature vectors to unit length, which cancels out the effects of changes in image contrast. Also, it reduces the influence of large gradient magnitudes caused by non-linear illumination changes through thresholding the values in the unit feature vector and renormalizing it to unit length. These methods emphasizes the distribution of orientations rather than matching the magnitudes for large gradients, making SIFT descriptors more robust to illumination variations.
4. Laplacian of Gaussian requires more calculation, since smoothed images, L , need to be computed in any case for scalespace feature description, while D can be computed by simple image subtraction, fastening the calculation and remains as a good approximation of the points.
5. The size is 128.

(b) Image Matching

Key Point with max scale in Cat1

closest neighboring key point in Cat₃

(a) Key point with max scale on Cat 1 and its neighbor in Cat3

Cat1



Cat3



(b) Key points pairs on cat 1 and cat 3

Key Point with max scale in Cat3

closest neighboring key point in Cat₂

(c) Key point with max scale on Cat 3 and its neighbor in Cat2

Cat3

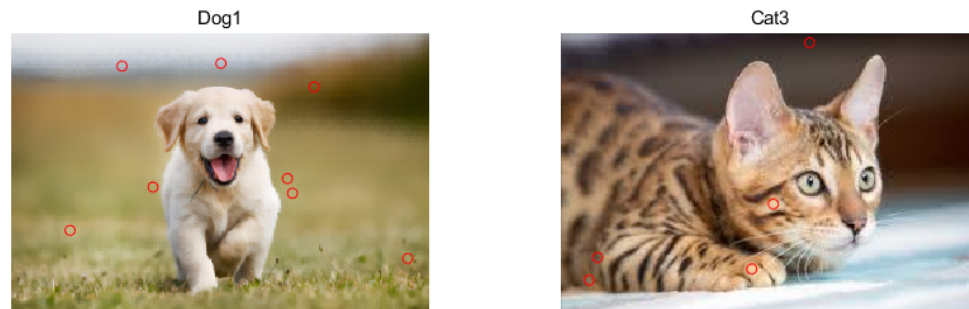


Cat2

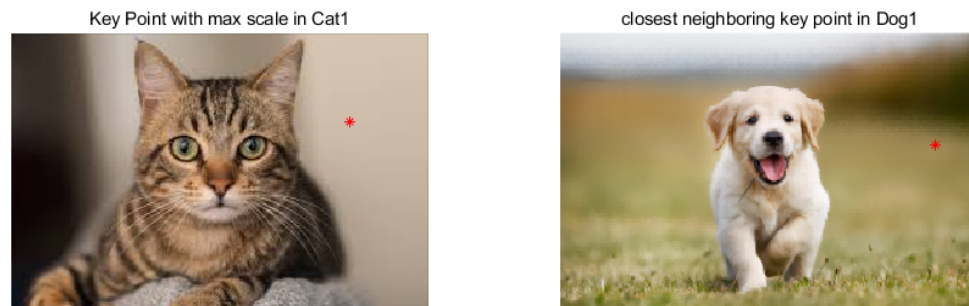




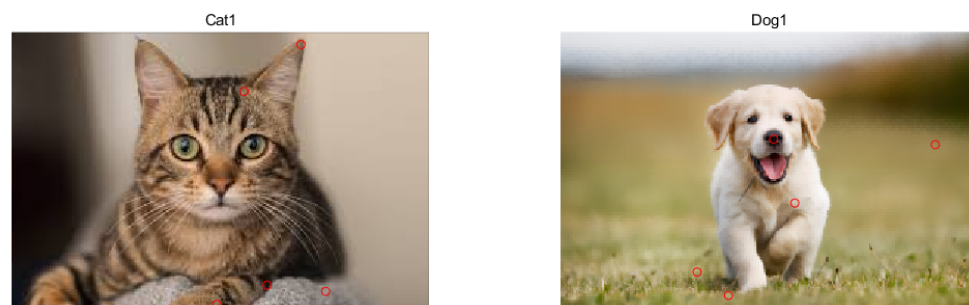
(a) Key point with max scale on Dog1 and its neighbor in Cat3



(b) Key points pairs on cat 3 and dog1



(c) Key point with max scale on Cat 1 and its neighbor in Dog1



(d) Key points pairs on cat 1 and dog1

Figure 4: 3(b) SIFT key points of fig 4.

(c) Bag of Words

The outcome:

Similarity of Cat3 codewords to other images:

Similarity to Cat1: 10.6771

Similarity to Cat2: 9

Similarity to Dog1: 175.2199

We find the SIFT could classify the bodies of the picture very well: at least it could tell dog from a ginger-colored cat with black strips.