

BIOMETRICS

Fall 2013

Assignment 3: Texture Classification

Due: October 17, 2013 at 5:00 pm via Courseworks

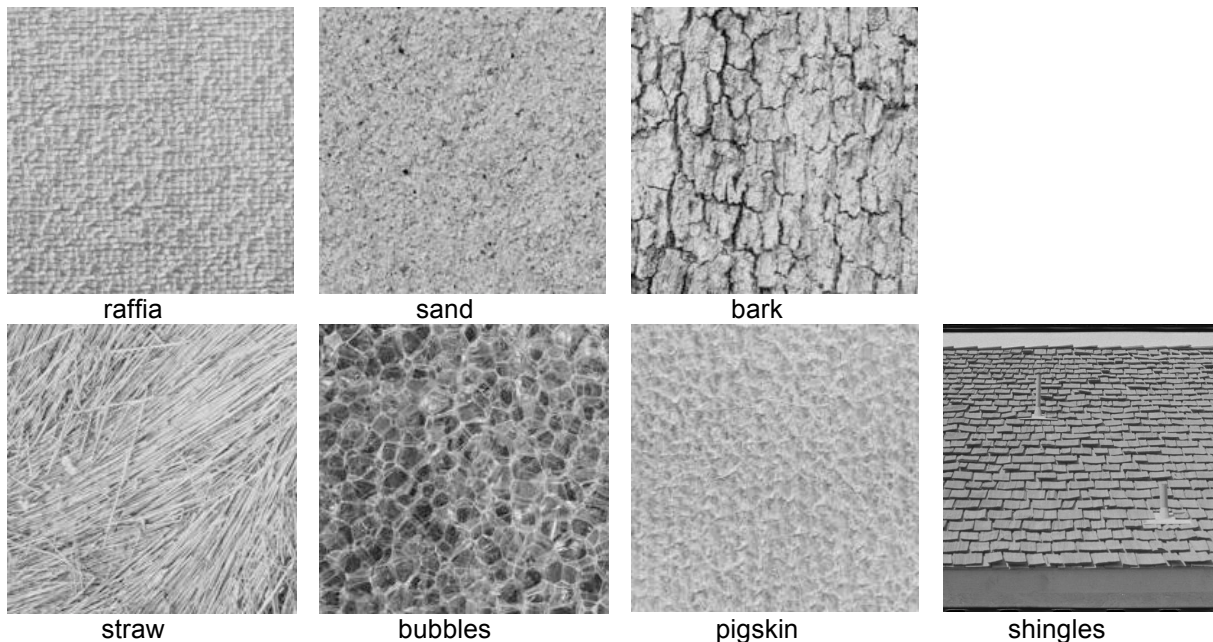


Figure 1: Textures of Natural Objects

Lab Description:

You must select and implement two classification algorithms for discriminating samples taken from 7 different image textures – raffia, sand, bark, straw, bubbles, pigskin, and shingles; the textures are shown in the above figure. The samples are 32x32 pixel windows whose position is randomly chosen in one of the 512x512 pixel textures. The image processing portion of the lab, i.e., the extraction of the feature vectors, has been done for you. You will download a training file (trainfile) containing the sample feature vectors labeled with the class from which they originated. An explanation of the form of the trainfile, including descriptions of how each feature was obtained, is included below.

Both of your classification schemes will be tested on a testing file (testfile) of the same format – except the column in the test file that contains the label for the feature vector will be censored.

You are to consider this lab as a friendly competition between you and the other class members. The winner, the one whose programs correctly identify the most textures, will be announced in class. We regret that no prize money will be awarded.

What to Program:

You will program two classification algorithms. The first (in file `classifier_1.m`) must be a Bayes classifier using maximum likelihood estimation. The second (in file `classifier_2.m`) must be either a (K) Nearest Neighbor classifier or a Decision Tree. Please write your programs in Matlab.

What to Hand In:

Please put all your files into a directory named `hw3_{your_UNI}` (substitute your UNI for `{your_UNI}`, so if your UNI is `ab1234`, the directory should be `hw3_ab1234`). Zip the directory into a file named `hw3_{your_UNI}.zip`, and submit it using the "Assignments" feature of Courseworks. The name of the assignment in Courseworks is "HW3." Simply upload you zipfile as an attachment and submit the assignment. Your submission should include the following:

1. A well-written report on the experimentation done, including an explanation of how you chose which classifiers and features to use. Include a description of the preliminary experiments or analysis you performed while building your classifiers (e.g. scatter plots you used to find discriminative features, or a table cross-validation results of different classifiers you tried). This should be a pdf document, named `report.pdf`.

2. Your well-commented source code. You should have two **functions**, each in its own file, `classifier_1.m` and `classifier_2.m`. They should be callable like this:

```
>> classifier_1(path-of-training-file, path-of-test-file, ...  
    output-path-for-edited-training-file, output-path-for-edited-test-file)  
(and the same for classifier_2). You are certainly welcome to have other functions which  
are called by these two, but these are the two we will run.
```

3. Two edited trainfiles, one for each of your classifiers, where the true class is replaced on each line by the results of your classifier, e.g., either ***raffia***, ***sand***, ***bark***, ***straw***, ***bubbles***, ***pigskin***, or ***shingles***. Please name these files `trainfile_output_1.txt` and `trainfile_output_2.txt`.

4. Two edited testfiles, one for each of your classifiers, where the word ***test*** is replaced on each line by the results of your classifier, e.g., either ***raffia***, ***sand***, ***bark***, ***straw***, ***bubbles***, ***pigskin***, or ***shingles***. Please name these files `testfile_output_1.txt` and `testfile_output_2.txt`.

Note that the four output files should be *exactly* what is produced by running `classifier_1` and `classifier_2`. Also, please be careful to *use exactly the filenames given here*, as the grading process is (partly) automated.

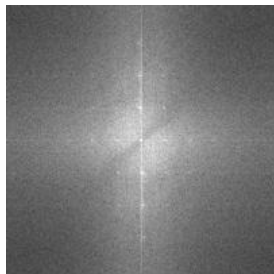
Grading:

The lab will be graded out of 60 marks. 40 marks are based on your classifier's performance on the testfile. This part of your grade is calculated mechanically as accuracy x 40. We'll use the better of your two classifiers. 15 marks are for your report. In your report, be sure to justify the choices you made in building your classifiers and choosing your features. You must include some sort of convincing reason to believe that your classifier is accurate (a good way to do this is to include cross-validation results; a bad way is to just report your accuracy on the training set). The remaining 5 points are for details; minor logic errors in your program, formatting errors in the output, etc. will result in deductions here.

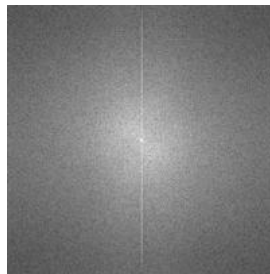
Test File Format:

Columns

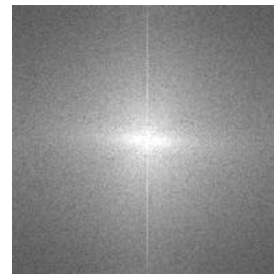
1. Sample class
2. Mean in the white section of Fig. 3.a. of its power spectrum shown in Fig. 2.
3. Mean in the white section of Fig. 3.b. of its power spectrum.
4. Mean in the white section of Fig. 3.c. of its power spectrum.
5. Mean in the white section of Fig. 3.d. of its power spectrum.
6. Mean of the sample.
7. Variance of the sample.
8. Mean of the square of the sample's horizontal gradient.
9. Variance of the square of the sample's horizontal gradient.
10. Mean of the square of the sample's vertical gradient.
11. Variance of the square of the sample's vertical gradient.
12. Mean of the sample's vertical gradient x horizontal gradient.
13. Variance of the sample's vertical gradient x horizontal gradient.
14. Variance of the sample's Laplacian.
15. Variance of the smoothed sample's Laplacian.



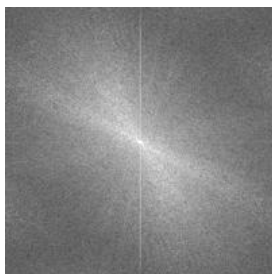
raffia



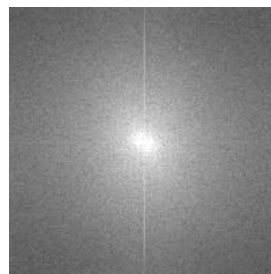
sand



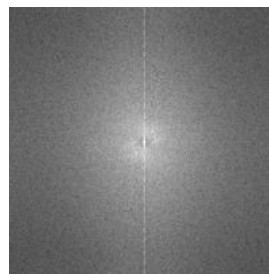
bark



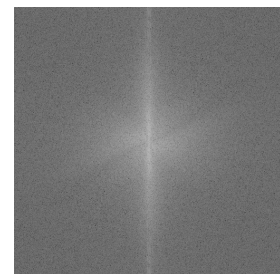
straw



bubbles



pigskin



shingles

Figure 2: Power Spectra of Images

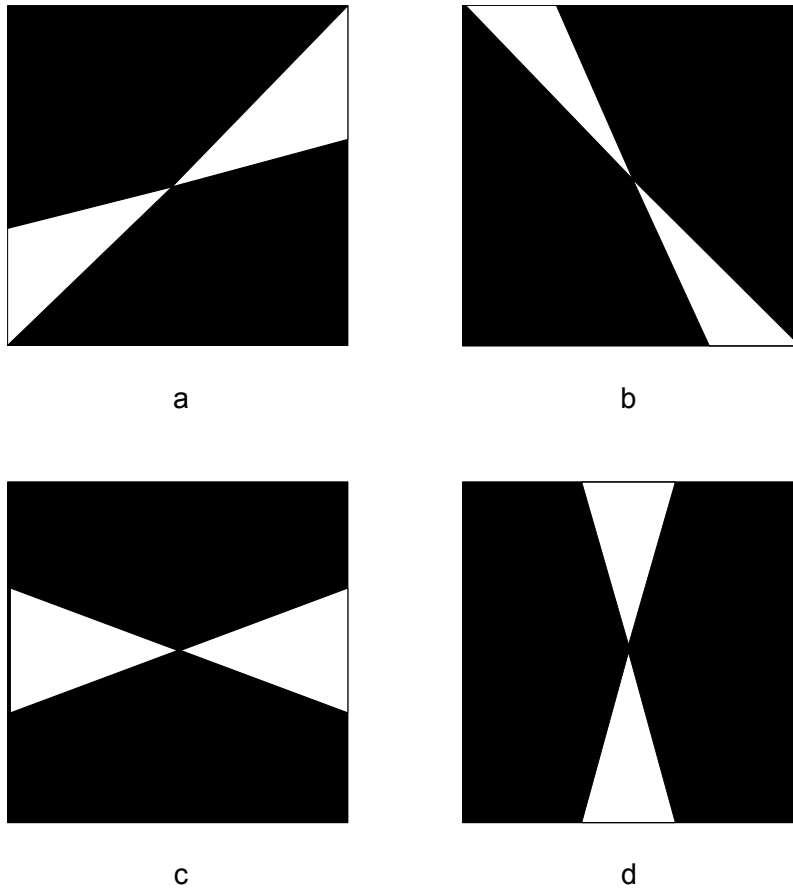


Figure 3: Power Spectrum Masks