Practice 5
Face Detection and Eigenfaces

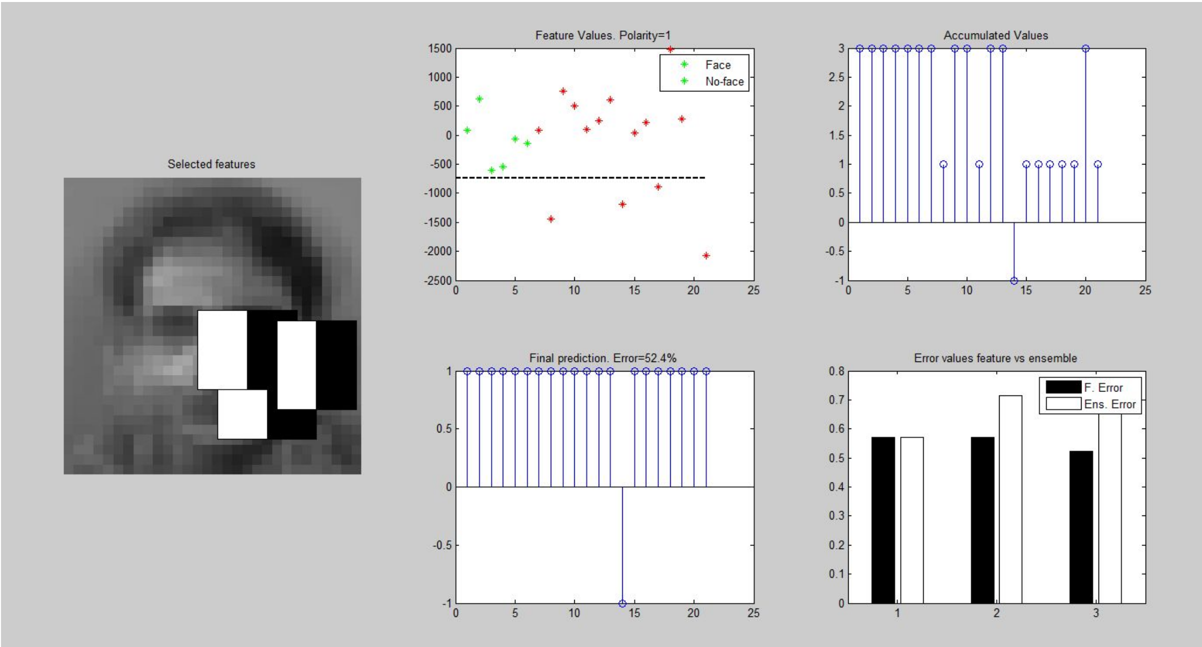
Aarón Negrín and Miriam Bermúdez

Exercise 4.2

Using the given code you need to:

a) **Select three features and explain the results provided by the function HaarFeaturesDemo**
b) **From the results given for each feature, can you point some good criteria to select a feature?**
c) **Which criteria is used to select the threshold?**
d) **Why combining features is not always good? Can you explain why?**

a) At the first place we present the output given by haarFeatureDemo, we can see a collection of figures which we are going to explain one by one



i) The figure at the left represents mean of the images we are using to work on and the three Haar features selected for this example.
ii) The figure labeled with the title ‘Feature Values’ shows the feature value calculated for every image in the dataset. Moreover the dotted line represents the threshold value generated for the last introduced feature.
iii) At the top right we can appreciate the ‘Accumulated Values’ figure which represents for each image in the dataset the sum of votes of ‘being a face’ (+1) of ‘not a face’ (-1), this values are generated once we have applied all the features. So this figure represents the strong classifier.
iv) At the bottom left we can appreciate the error in the face detection, all the images in the dataset are marked as a face or not a face, this plot shows the hits of face/not face
v) The bottom right plot shows the error of a certain feature individually or consider all the features together

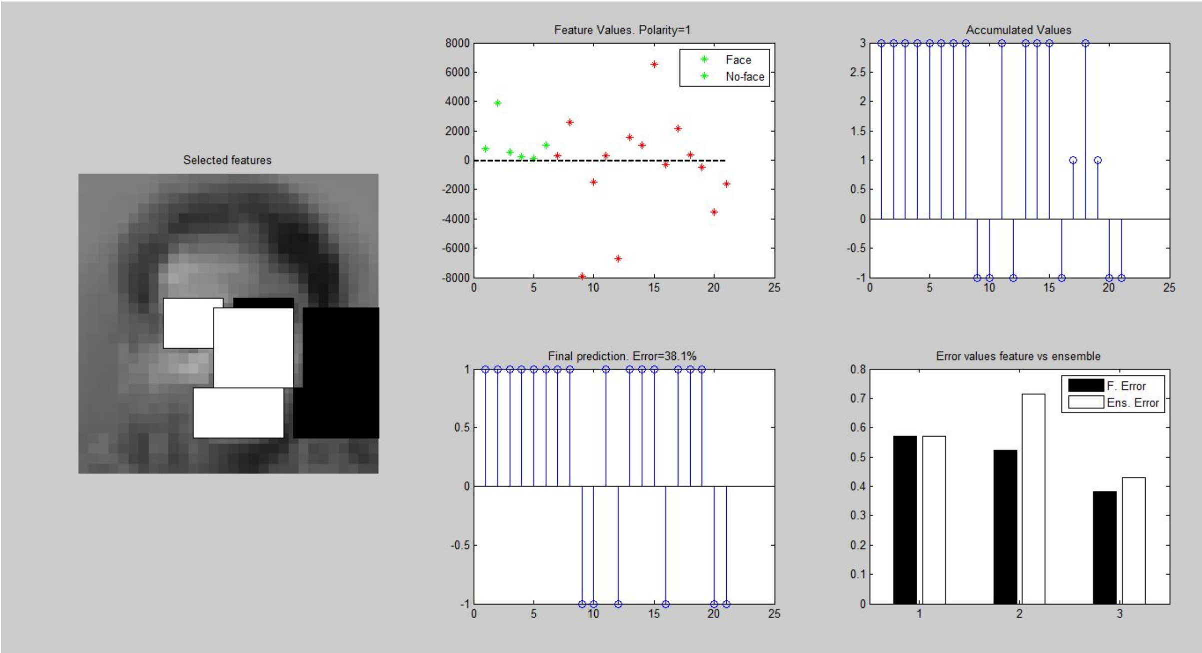
b) A good criteria to select features would be picking up relatively large features and features near of edges, where we can observe a high change in image contrast.
c) In this case the criteria used to select the threshold has been depending on polarity, for instance, for polarity +1 we pick up the minimal value of the ‘white section’ of the feature, the maximal value of the ‘black section’ of the feature and calculate the media
d) Combining features could not be always a good idea because if we find a very good feature adding another one not as good as the previous one could introduce noise into the strong classificator. That is because the additive behaviour of the classifier.

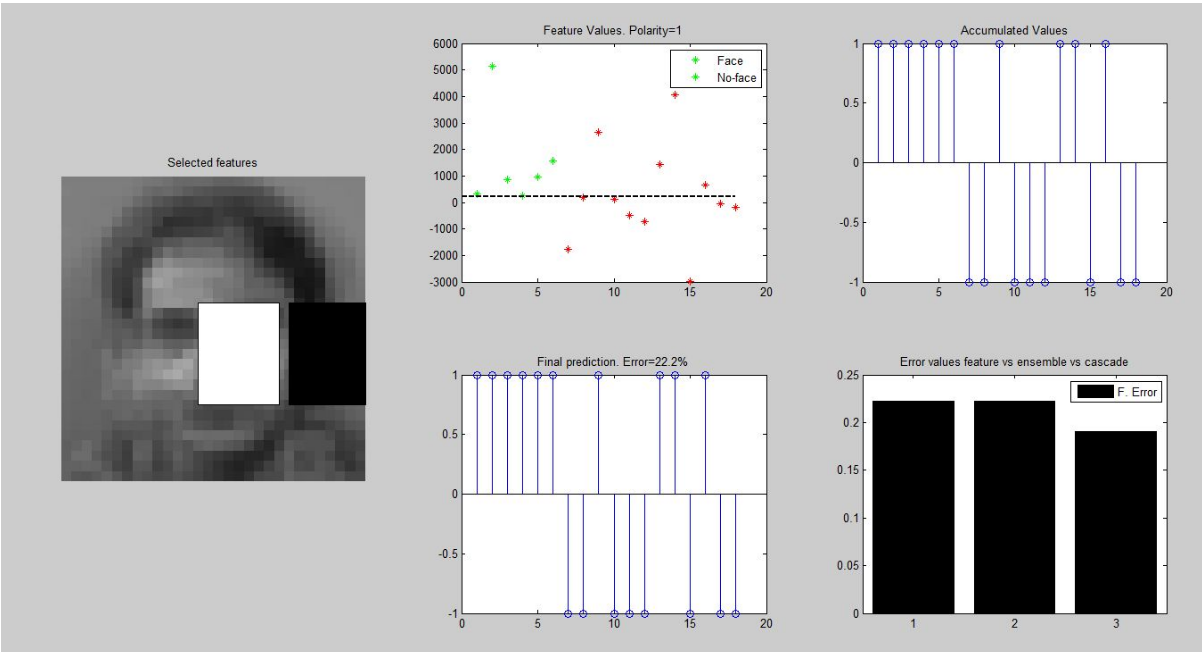
Exercise 4.3

a) Create a function called ex43 which uses the method haarFeatureDemo to train a three levels cascade with two weak detectors per level
b) Repeat the previous experiment with the following behaviours:
   i) One level cascade + three weak detectors
   ii) Three level cascade + one weak detector per level
   Trying to select the same features and answer.
   The error is the same in both case?
   Despite the features are the same, are the weak detectors the same?

a) Check ex43.m

b) The source code is in ex43b.m



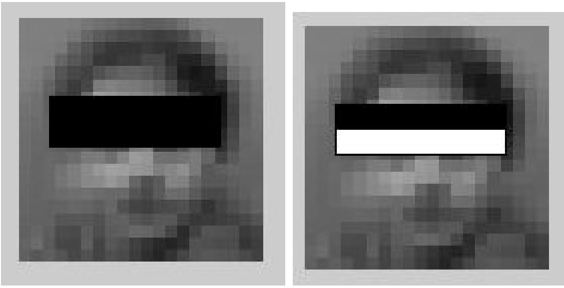


Despite we are using almost the same three Haar Features we can observe a better performance when using 3 levels and only one feature per level.
Moreover, on higher levels we check the non-faces results only for that reason we can assume that in general the weak classifiers are not the same for dimensionality reasons.

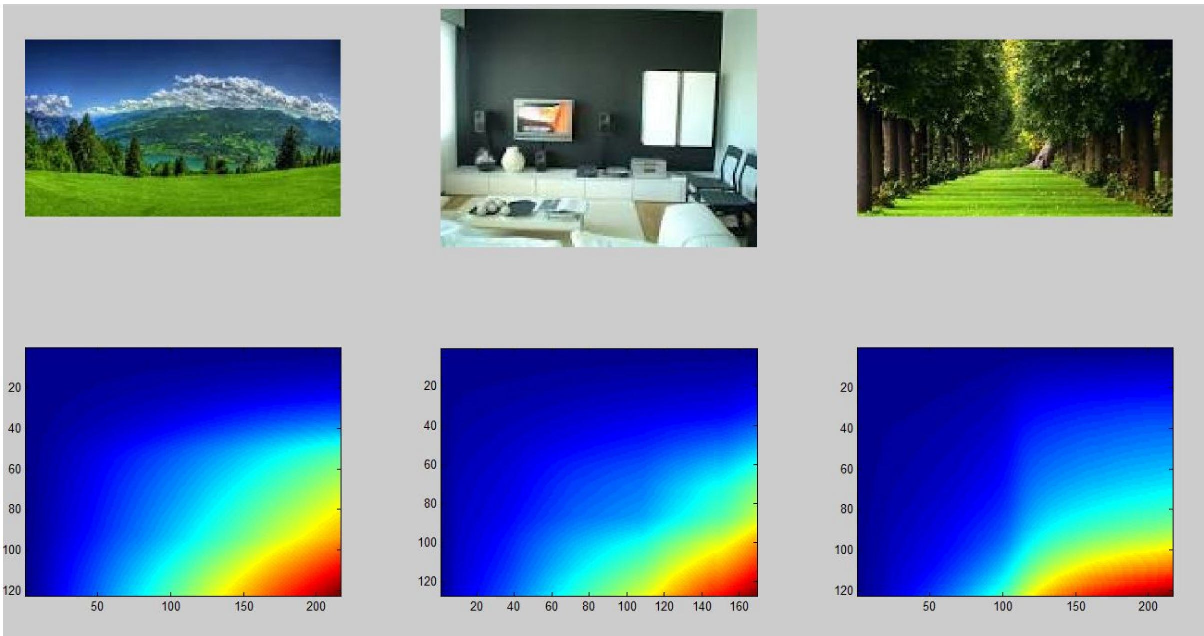
Exercise 4.4

a) **Why a rectangle is defined with 5 parameters? Does this feature make sense?**
b) **Show the integral image for two or three examples and explain what they have in common. Justify your answer**

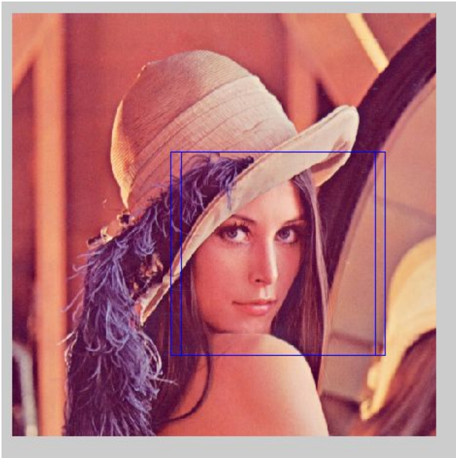
a) Check ex44a.
These features provides an additional parameter which if it is -1 it means it will be an entire rectangle considered as an inhibitor zone.
On the other side a positive value gives us the activator zone for that feature.
The given features makes sense because it can provide good quality of information about the contrast in the zone of the forehead and eyes



b) Check ex44b.
As we know the integral image represents, for a certain pixel (x,y), the sum of all intensities for the rect [0 0 x y], considering our origin at the top left of the image.
So, as we can appreciate here in the given heatmaps while the value for the integral image near of (0,0) is very low, when we take values near to the size of the image, the integral value takes its highest level.



c) Check ex44c



Part B

1. **What is the learning process?**
2. **Which method is used to retrieve the ‘nearest’ faces?**
3. **What is the ‘trick’ used in the eigenfaces method?**
4. **Which is the dimension of the reduced space? Compare it with the dimension of the original space**
5. **What happens if you apply the recognition method over the training dataset?**
6. **What happens if you rotate the faces in the test dataset?**

1. It consists first on charging the images we are going to use in a matrix (the ones we are going to learn),then we calculate its media in order to get the covariance matrix but in this case we use the trick explained in question 3 instead of the covariance matrix, after that we get the eigenvalues and eigenvectors, we sort it and we only keep the ones that are more than 0 and finally we calculate the eigenfaces and save them in the columns of the matrix (U) that we will use it the test part.

2. We use the method knnsearch() which finds the ‘K’ nearest neighbors in this case three.

3. It consists on the multiplication of a matrix with it transpose in order to get the eigenvalues easier than with using the covariance matrix.
   That is because the eigenvectors of AA^T are the same of the A^TA

4. The original space is 243x320px and the final dimension is 83x83px.

5. When we use the same file (‘training_data’) for recognition, the similar faces that it tooks are the faces from the same person but in different moods and also it finds itself as a similar face.

6. It finds different faces whether we compare it with the ones it founds at the beginning with the non rotates test faces.