Pattern Recognition practical 1

Maikel Withagen (s1867733) Steven Bosch (s1861948)

September 2015

1 Assignment 1

1.1

Consider it done

1.2

To compute the pair-wise correlation coefficients we used the following command:

Input

```
load('lab1_1.mat')
corrcoef(lab1_1)
```

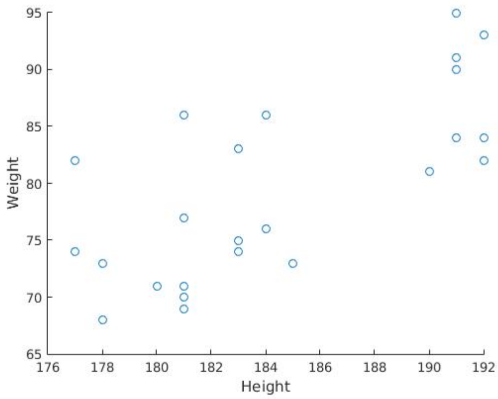
This yields us the following table of correlation coefficients:

Table 1: *Pair-wise correlation coefficients*

	Length	Age	Weight
Length	1	-0.0615	0.7156
Age	-0.615	1	0.5142
Weight	0.7156	0.5142	1

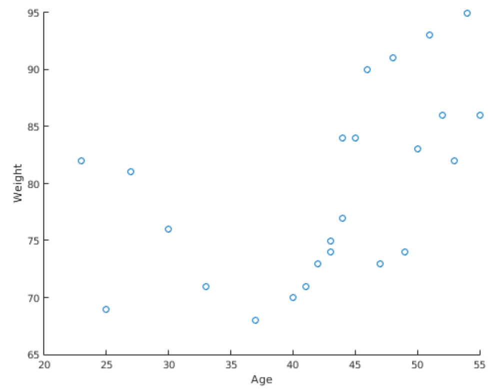
1.3

The two features for which the correlation is the largest are the first and third column, respectively the height and the weight.



(a) Scatterplot of weight to length

The two features for which the correlation is the second largest are the second and third column, respectively the height and the weight.



(b) Scatterplot of weight to age.

Figure 1

From a scatterplot alone it is hard to draw conclusions about any possible relationships between the different features. We do get indications though; figure 1a shows that there is likely to be a correlation between the weight and the height. An increase in weight seems to correspond to a (somewhat linear) increase in height. A similar kind of relationship can be seen in figure 1b, between the factors weight and age.

2 Assignment 2

2.3 Sets S and D

The following subsections show the code we used to acquire the 1000 Hamming distances for set S and D.

a

Code for set S

```
hd_s = zeros(1,1000);
for i = 1:1000
person = randi([1,20]);
row1 = randi([1,20]);
row2 = row1;
while(row1 == row2)
row2 = randi([1,20]);
end
load(sprintf('person%02d.mat',person));

hd_s(i) = sum(abs(iriscode(row1,:)-iriscode(row2,:)));
end
hd_s = hd_s/30;
```

b

Code for set D

```
hd_d = zeros(1,1000);
for i = 1:1000
person1 = randi([1,20]);
row1 = randi([1,20]);
row2 = randi([1,20]);
person2 = person1;
while(person1 == person2)
person2 = randi([1,20]);
end
load(sprintf('person%02d.mat',person1));
x = iriscode(row1,:);
load(sprintf('person%02d.mat',person2));
y = iriscode(row2,:);
hd_d(i) = sum(abs(x-y));
end
hd_d = hd_d/30;
```

2.4 Histogram

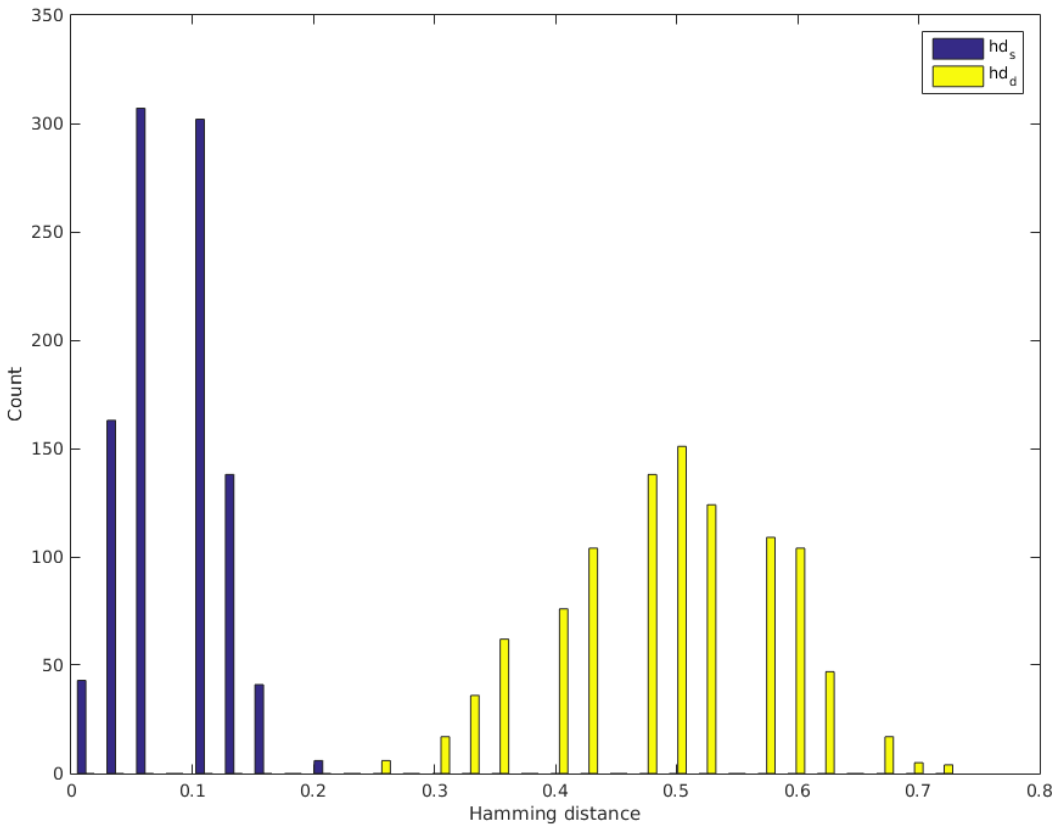


Figure 2: Histogram of sets S and D.

Figure 2 shows the histogram of sets S and D. The figure shows that the two distributions overlap very little, most of it around $hd = 5, 6$.

2.5

The means and variances of both of the sets are the following:

Set	Mean	Variance	Standard deviation
S	0.0825	0.0016	0.0398
D	0.4946	0.0079	0.0886

Table 2: Means and variances of both of the sets

The prior probability of two bits being the same between persons the following is: $1 - 0.4946 = 0.5054$. Using the formula $n = p(1-p)/\sigma^2$ we get $n = (0.4946 * (1 - 0.4946))/(0.0886^2) \approx 31.84$ statistically independent bits that are needed to encode an iris pattern. This means that the current number of bits in the iris vectors is insufficient.

2.6

We used the following Matlab code to create the graph in figure 3:

```
hold off;

% Making the histogram
hd = [hd_s;hd_d].'
hist(hd,30); xlabel('Hamming distance'), ylabel('Count');
hold on;

% Create scaled Gaussian plots
x2 = sum(hd_s(:) == mode(hd_s));
normd = normpdf([0:0.01:1],mean(hd_s),std(hd_s));
plot([0:0.01:1],normd*(x2/max(normd)));

x1 = sum(hd_d(:) == mode(hd_d));
normd = normpdf([0:0.01:1],mean(hd_d),std(hd_d));
plot([0:0.01:1],normd*(x1/max(normd)));

legend('hd_s', 'hd_d', 'hd_s normd', 'hd_d normd');
```

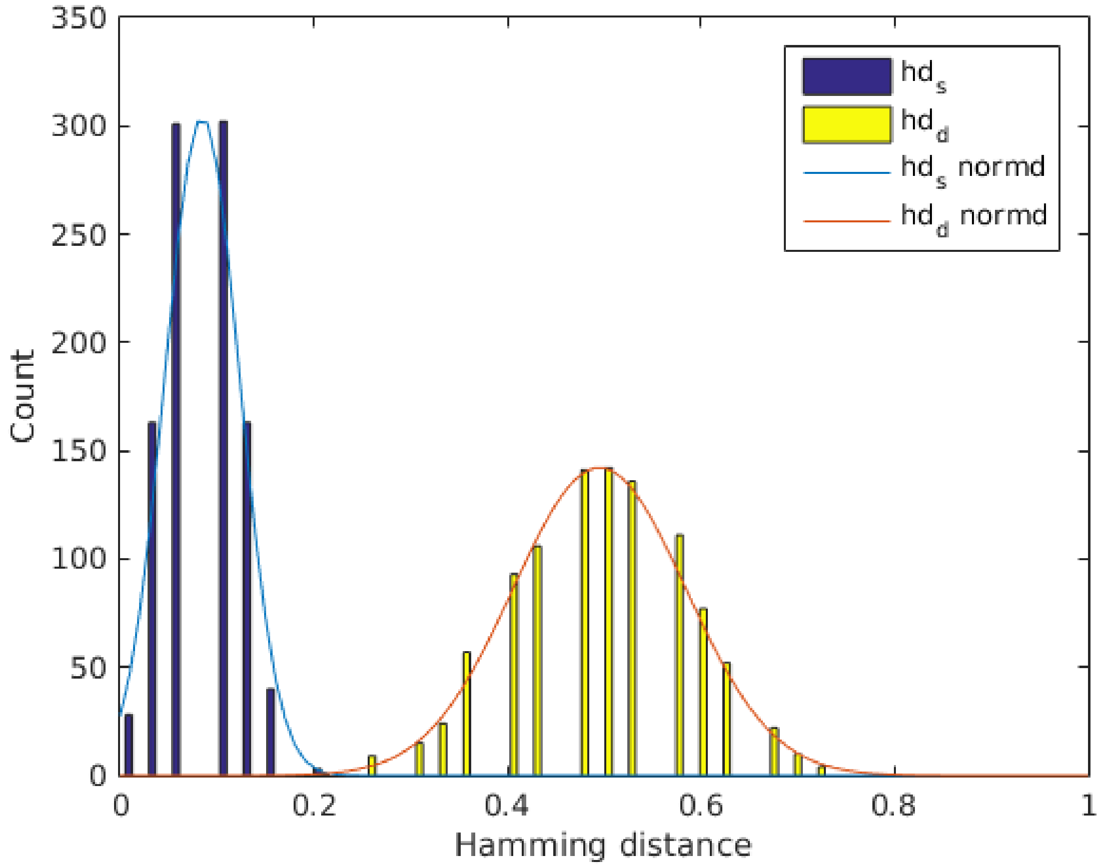


Figure 3: The histograms of sets S and D plotted with two Gaussian functions using the mean and variances.

Figure 3 shows the plot of the combined histograms and gaussian functions (note that the plot is slightly different from the previous subsections, because we reran the script in Matlab). The scaling is done by dividing the maximum value of the D and S sets by the maximum value of the initial Gaussian functions. This way the maximum value of the new Gaussian functions is the same as the maximum value of the D and S sets, so that the the Gaussians are nicely scaled to the histograms.

2.7

a

2.8

3 Assignment 3

3.1 Background

The topic of biometric identification has been receiving a lot of attention lately. The use of fingerprints, facial features, and iris recognition are considered appropriate for identification. Of these three, fingerprints and iris recognition are considered to be more accurate than facial features, while facial featuresare more flexible, e.g. for the use in surveillance scenarios. A possible improvement for the performance of facial feature identification is the combination of information from multiple sources, for instance the ear. Ear images can be obtained in a similar maner to face images, and a recent study suggests that they are comparable in recognition power[1], and a combination of the sources gives a significant improvement over their individual performance.

3.2 Method

3.3 Achieved Results

3.4 Conclusion

References

[1] Kyong Chang, Kevin W Bowyer, Sudeep Sarkar, and Barnabas Victor. Comparison and combination of ear and face images in appearance-based biometrics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9):1160–1165, 2003.