MVA 2008 - 2009

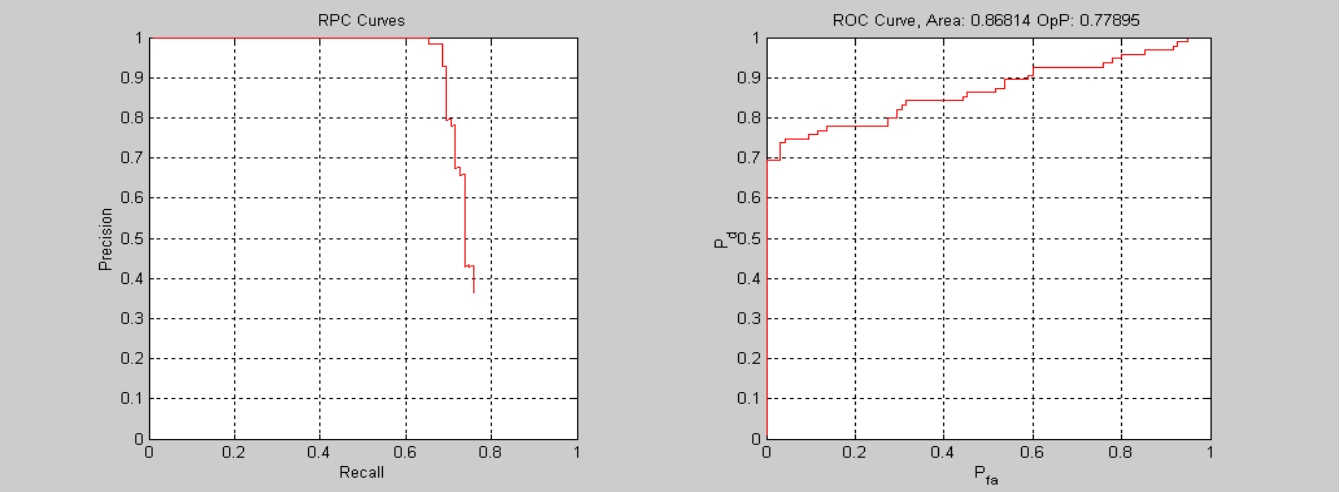
Object recognition and computer vision

Assignment 4: Pictorial structures

Antoine Labatie

Part I

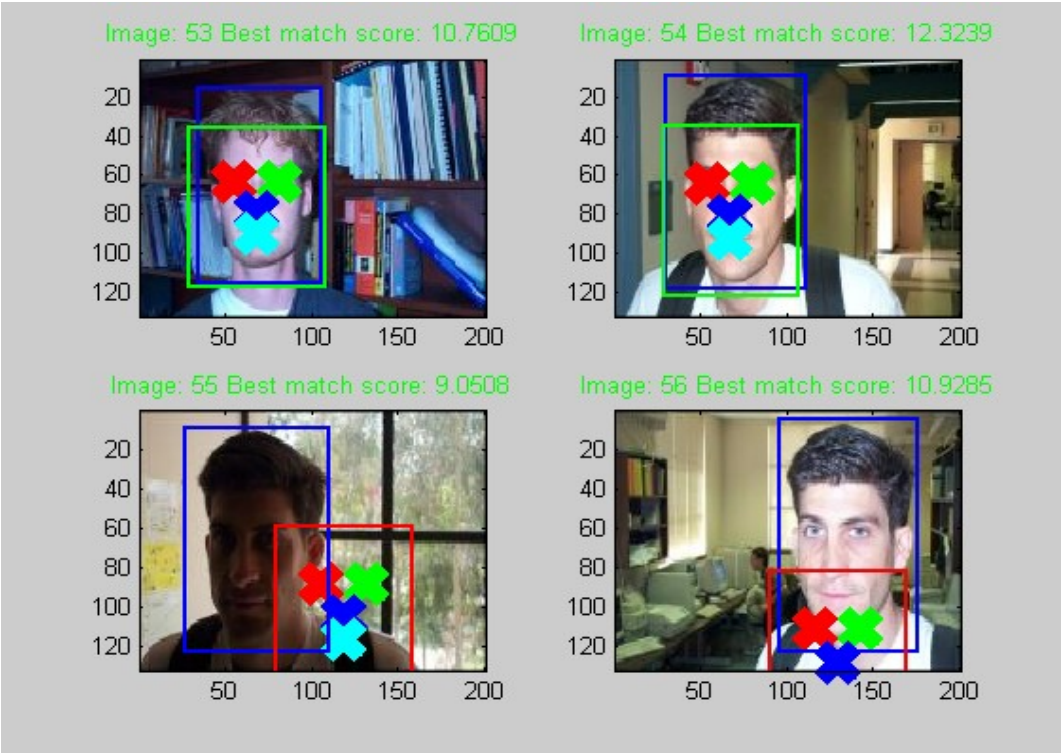
Here are the RPC and ROC curves obtained with the default parameters and the 4 part models.

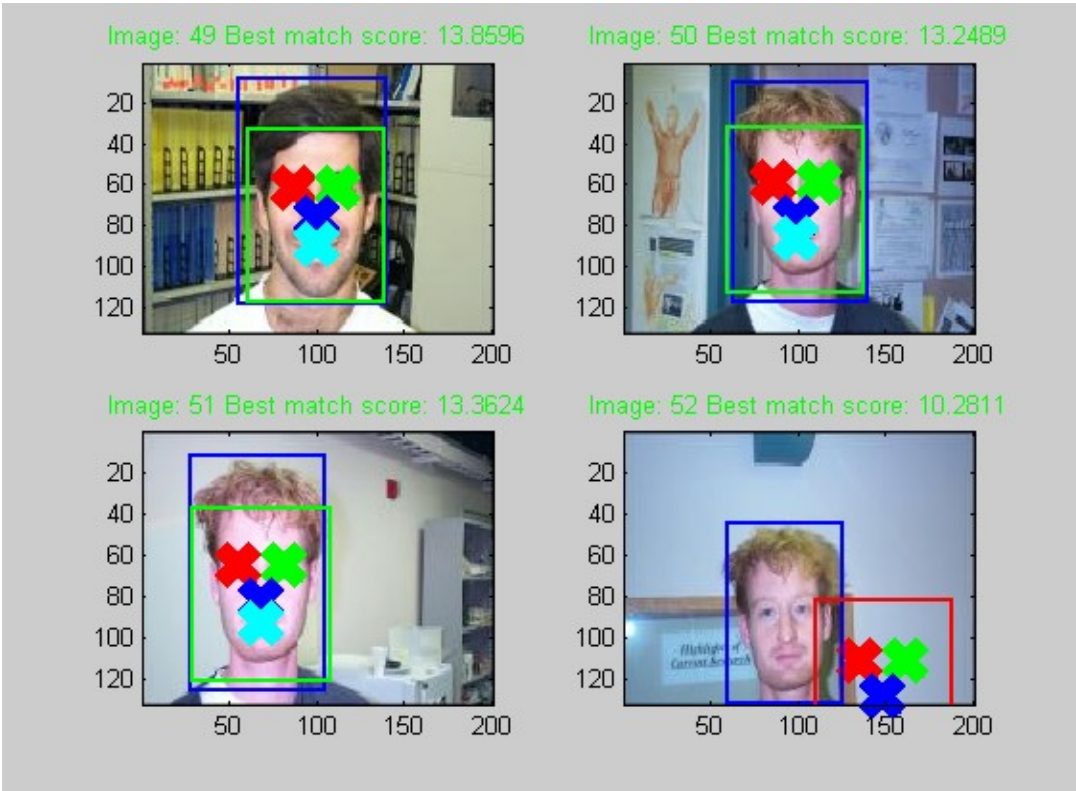


We see in particular that the precision rate decreases rapidly at one point and probably starts to detect faces even if there is not any in the image.

Here are two examples of false detection. For the first image on the bottom left it is probably because of the change in luminosity. Indeed the features are detected on the window which has an intensity closer to the features in the training images.

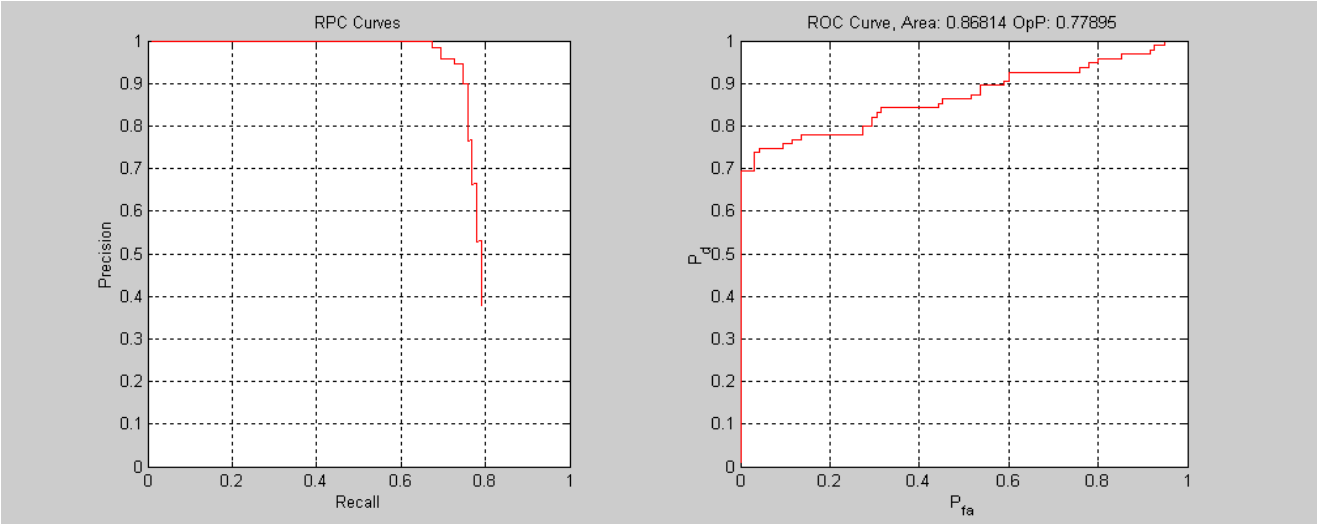
In the second image it may come from the change in the orientation of the face. On the second image we see a false detections due to a difference of scale for the face between the image and the training examples.





Part II

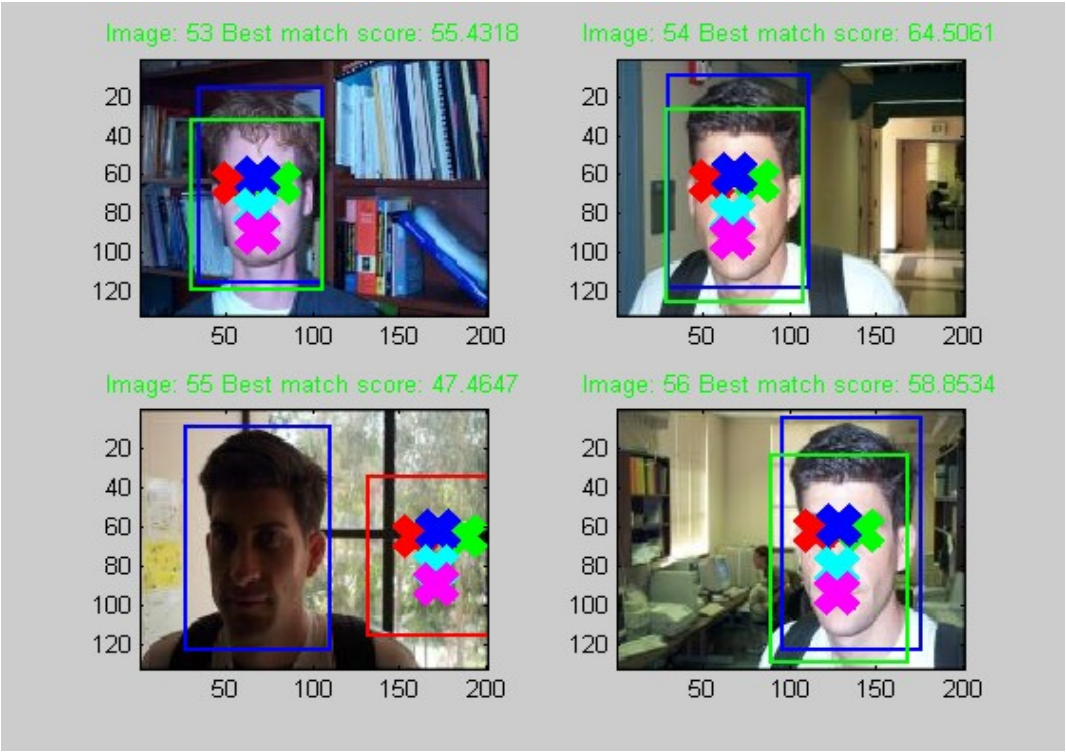
Here I changed two parameters: the number of parts which I changed to 5 and the coefficient *"Recog.Shape_Appearance_Weighting"* which I changed from 5 to 10 giving more weight to the appearance term relative to the spatial term. I obtained RPC and ROC curves which are a little better.



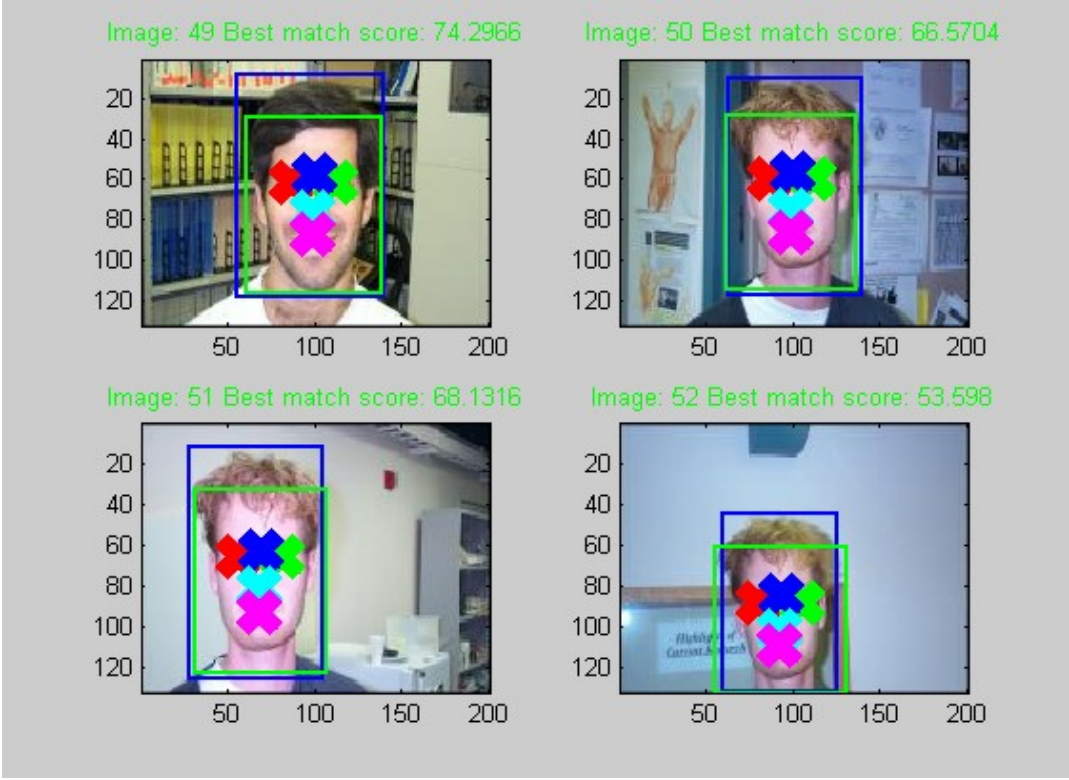
Adding another part in the model seems to make it work better, which is intuitive. However the part probably has to be distinctive and should not depend too much on the orientation of the face or on the differences between individuals (I chose a part between the two eyebrows).

Concerning the weight between appareance and spatial terms, it seemed that the default value is a little low. Increasing it will result in a box with parts closer to the training parts in terms of appearance, but with more possible distance variation between them. It is particulary good if we want to detect faces with some spatial variations due to small changes in scale or viewpoint.

Here is an example of good detection, where it was not the case with the default settings on the bottom right.

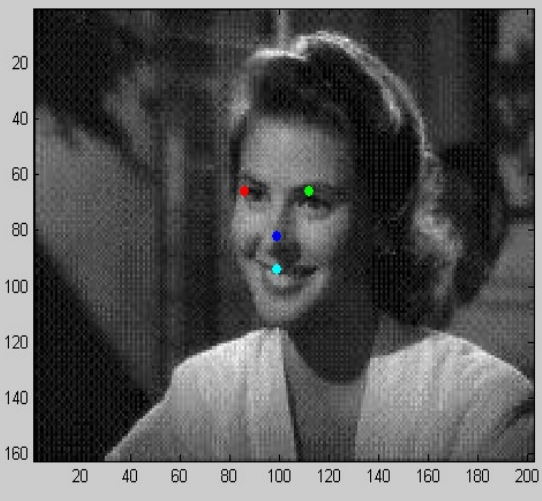


In this case we see that is can adapt to changes in scale.



Part III

Before testing the algorithm on the frames I had to resize them so that their width be equal to 200 which I did manually but I could have used to preprocessing function. Here are the results that I obtain for lambda=1.



So the features are well detected and localized. I tried values of lambda= 0.1, 1, 10 and 100. For all those values the parts were still well localized but the precision varied a little bit. In particular I noticed that the mouth was better localized with a small lambda and, on the contrary, the eyes were better localized with a high lambda. So the value lambda=1 seemed to be a good compromise.

The reason why the eyes are better localized with a high lambda seems to be that they are not horizontal due to the inclination of the face.

I am not completely sure why the localization of the mouth is less precise with a high lambda but I think it is because among the different parts, the eyes are the most easy to detect. For the nose and mouth which have more variability between individuals, allowing more variations will lead to a better maximisation of the appearance which is not necessarily good (though I find surprising that the appearance term is higher in cases where the part is located less precisely).

This means that with a part with more uncertainty we would need to rely more on the spatial model. Here is an example for lambda=100.