

ECSE 415 Project Report

Michael Smith, Haisin Yip, Yang Zhou

April 14, 2015

Part 1

Figs. 1, 2 and 3 are examples of all the possible poses for a given subject. These include yaw rotations from $-45^\circ, -30^\circ, 0^\circ, 30^\circ$ and 45° at 3 different scales: close, medium and far. The images were taken on a white background at a high resolution but downsampled to 150×200 in order to allow for fast processing of the LBP/SIFT algorithms and related functionality. They were later converted to grayscale for processing as algorithms such as SIFT require it.

Figs. 4, 5 and 6 show the different testing poses for a given subject. These include all of the rotations of the training images with the subject facing upwards followed by facing downwards. In addition, different images were taken of each subject without their glasses, with different facial expressions and with a hat on.

Part 2

Fig. 7 shows 10 selected faces with their SIFT features. Note some faces are blurry as the images are low resolution to allow for acceptable computation time although some subjects are far away.

Fig. 8 shows three examples of region histograms for a given training image. We can see that the tendency is for the lower values to have greater frequency.

Fig. 9 shows four histograms. They represent two images; each image has its histogram representing the codewords computed for both LBP and SIFT.

Part 3

Fig. 10 shows five examples of classification using both LBP and SIFT. We see that some images, such as (b) or (c) are correctly classified but others such as (e) are not.

With respect to the performance of the classifier, SIFT performs the best with a recognition rate of 69.6429% with $K = 20$. This is likely due to the fact that SIFT is more robust to the changes in lighting (mostly due to the iPhone's sensitivity to indirect light when pictures at a greater distance were taken) and scale seen in the training images as well as the testing images. LBP performs about as well as flipping a coin, although considering the variations in the images this is to be expected.

Finally, Fig. 11 shows five group photos with the faces detected and tagged. Overall, the performance is quite poor. In (a), while all the faces are detected they are all incorrect. In (b), Haisin is correctly detected on the left but not on the right where the algorithm believes him to be a clone of Haisin. Similarly, Eric is correctly detected in the center but somehow is duplicated on Michael behind him. In (c), Eric was correctly detected albeit with an extremely large bounding box. None of the other members are detected. In (d), the only person to be detected is detected incorrectly. In (e), Gavin is correctly detected but not Haisin, who is misidentified as Michael.

It is likely that the poor performance of the algorithm with respect to the group photos is due to the fact that the lighting conditions and poses vary dramatically, far exceeding the poses and variations in the training images.

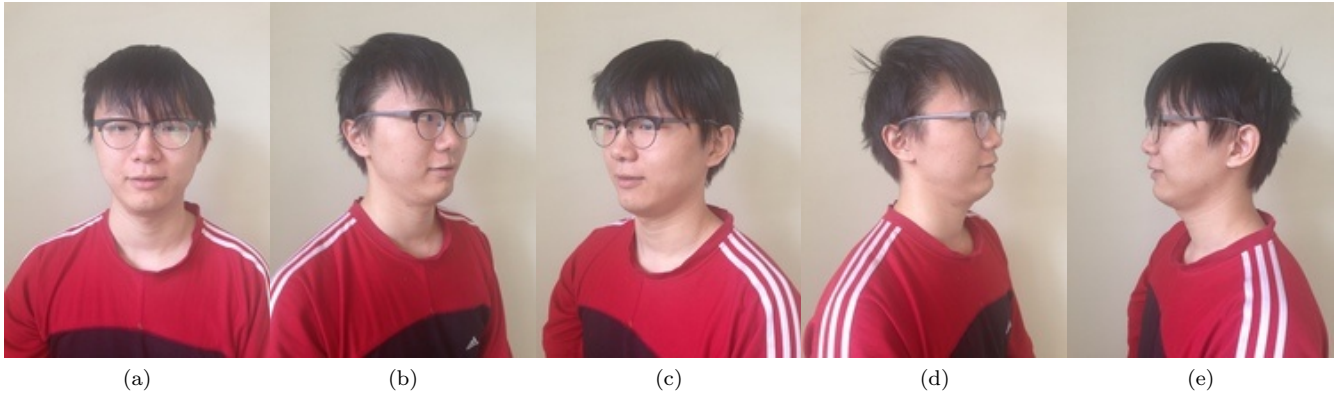


Figure 1: Training images at the close, medium and far scales.

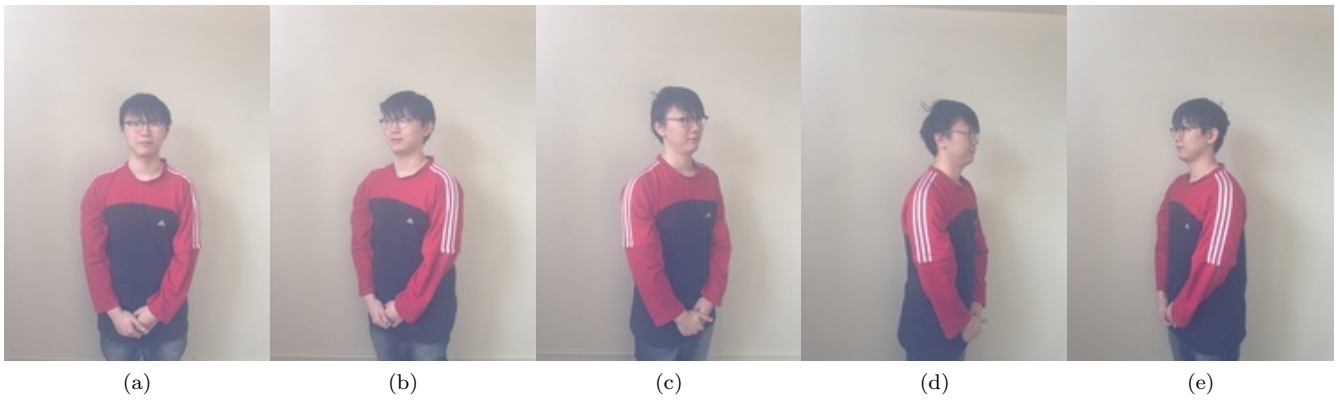


Figure 2: Training images at the medium scale.

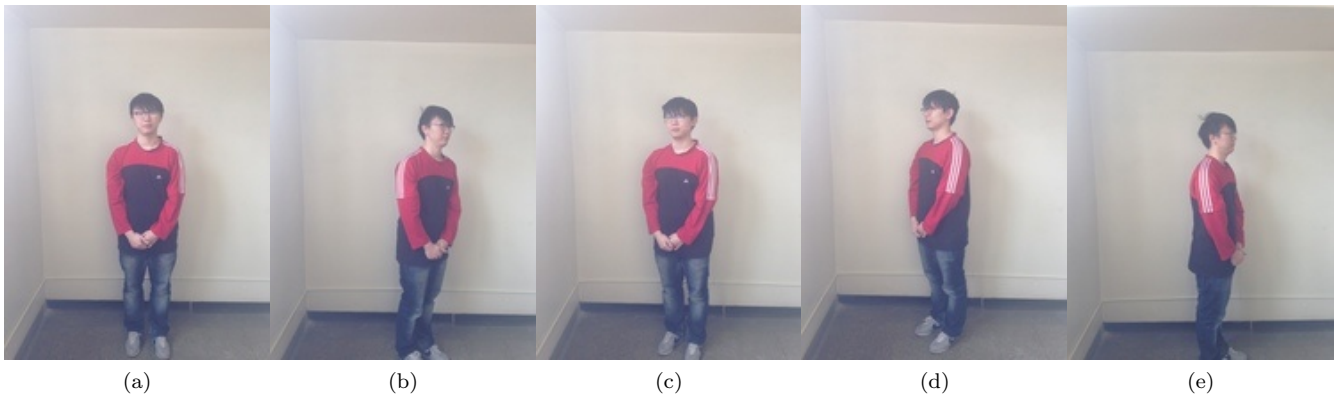


Figure 3: Training images at the far scale.



Figure 4: Testing images of the subject with their head facing upwards.

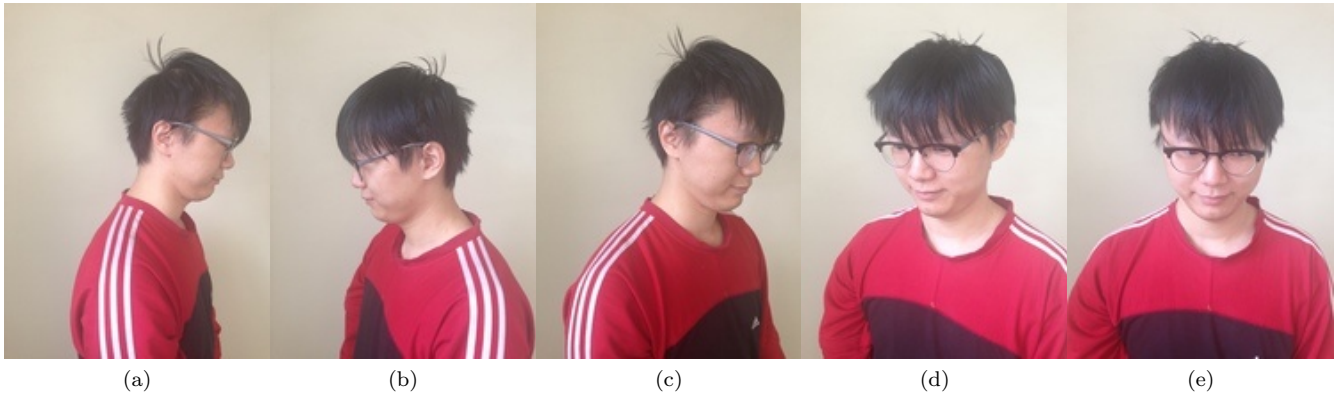


Figure 5: Testing images of the subject with their head facing downwards.

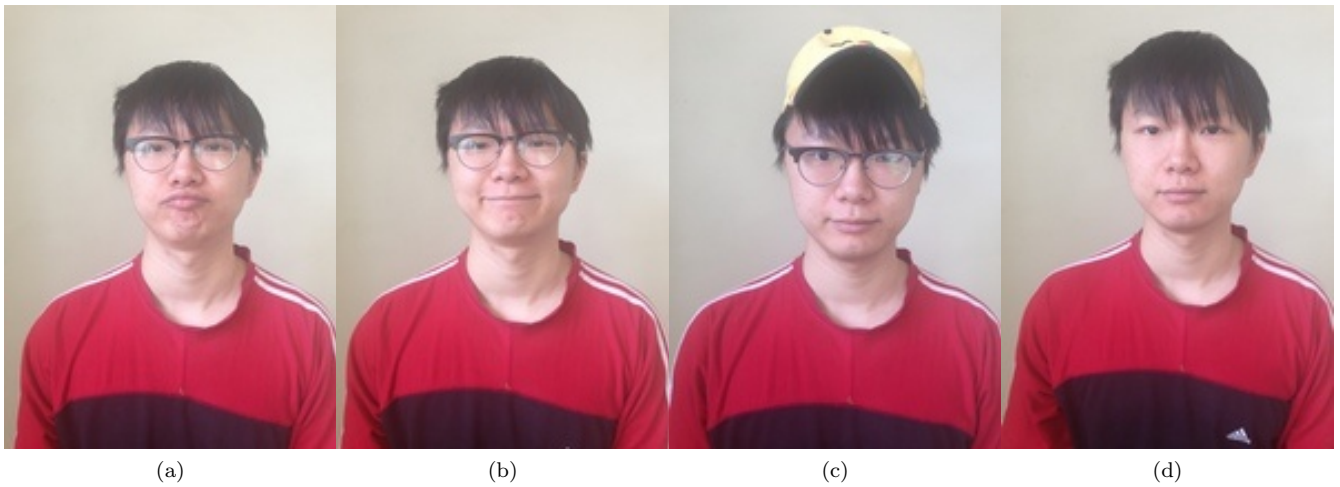


Figure 6: Testing images of the subject with accessories and different facial expressions.

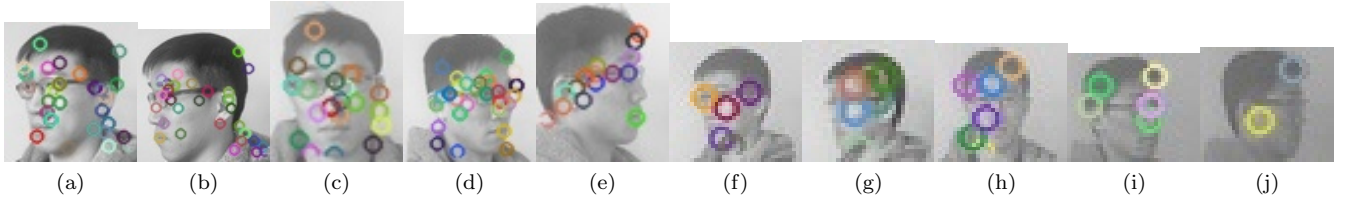


Figure 7: SIFT Features for 10 selected faces.

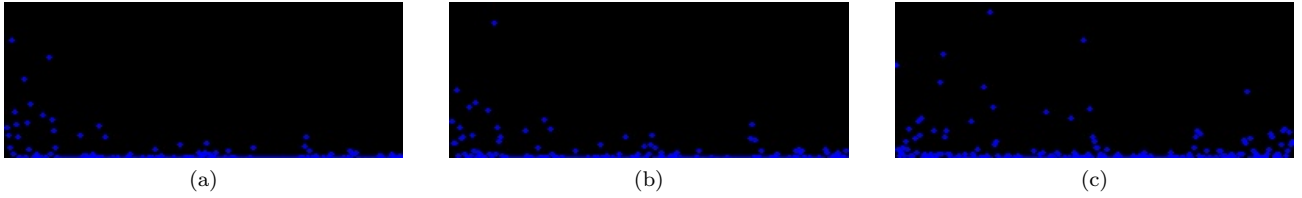


Figure 8: Three examples of region histograms for a given training image. The horizontal axis represents pixel intensity (from 0 - 255) while the vertical axis represents the frequency of occurrence of each pixel intensity.

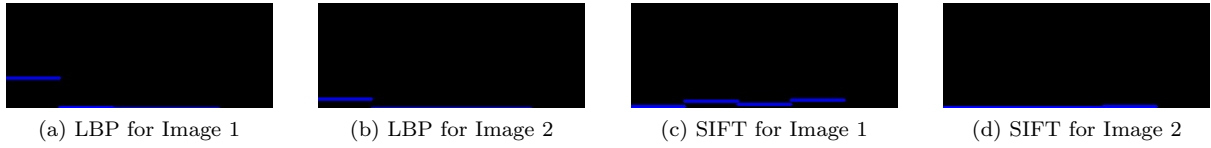


Figure 9: Histograms representing the codewords for SIFT and LBP for two select images. The horizontal axis represents the number of code words and the vertical axis represents the number occurrences of a given code word.

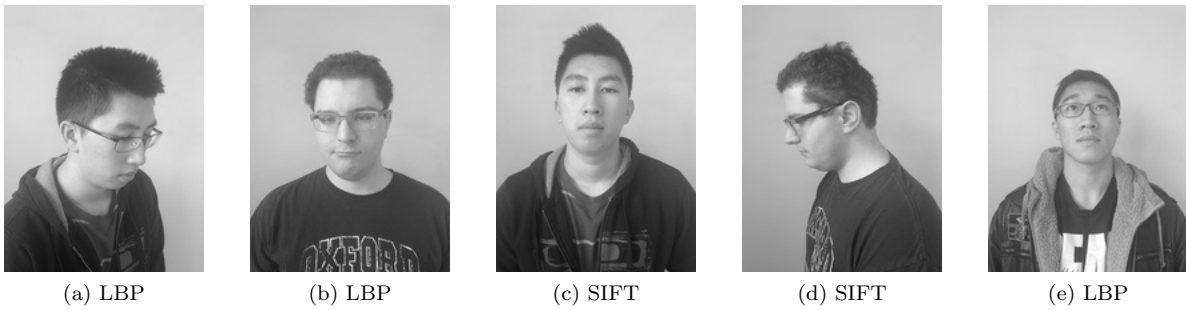


Figure 10: Five examples of classification using both LBP and SIFT. (a) is classified as Gavin when it is in fact Haisin. (b) is classified correctly as Michael. (c) is classified correctly as Haisin. (d) is classified correctly as Michael. (e) is classified as Michael but is in fact Eric.

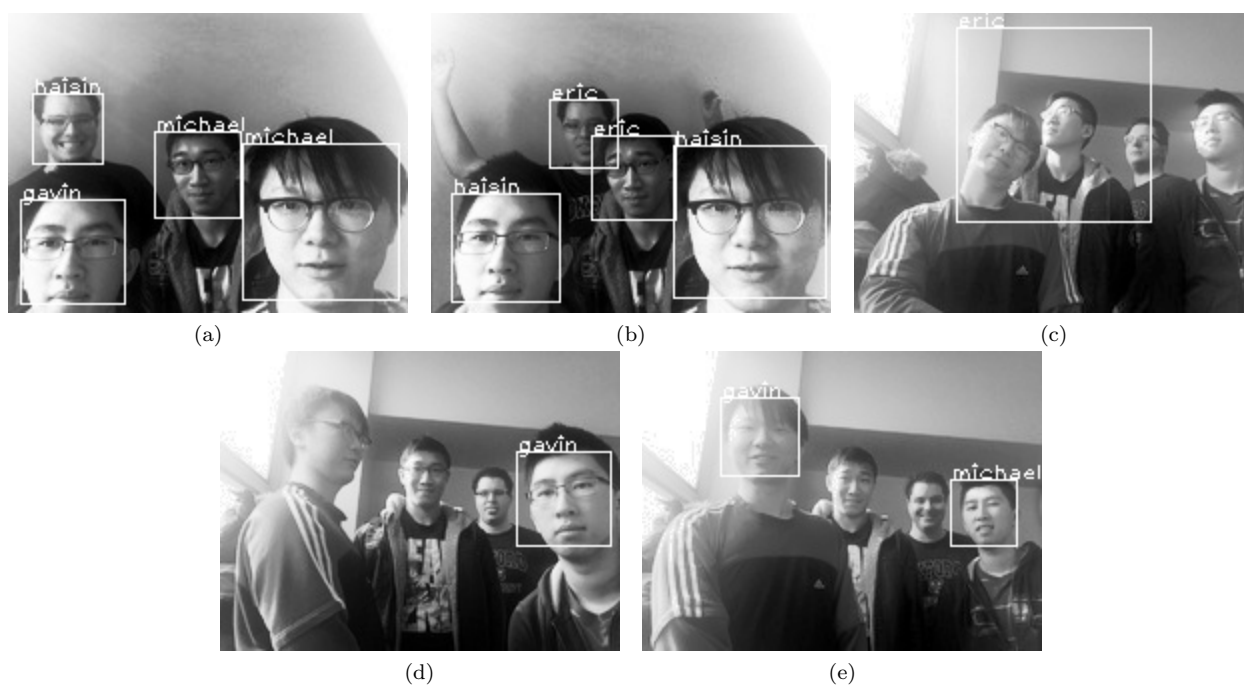


Figure 11: Five group photos with faces detected and tagged.