REALTIME HAND GESTURE RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS BASED ON SKIN COLOUR SEGMENTATION

December 20, 2017

Elizaveta Kharlova

Student ID: 1403693

Nazmus Sakib

Student ID: 1546701

Contents

INTRODUCTION

Gesture recognition is one of the current problems in HCI (Human Computer Interface). Principal Component Analysis (PCA) is a computationally friendly and robust way to perform image recognition. This method has been used to perform face recognition in [3]. This can be applied to gesture recognition in real time. The main challenges of this approach are reducing of variance between the images in order to improve recognition, decreasing computational complexity, and selecting the training set.

PROCEDURE

The process consisted of two main parts: training and recognition. In the training part, images were processed with segmentation and PCA in order to form a basis set. In the recognition part, each new frame was processed and projected onto the basis in order to produce a guess. This procedure is similar to [1].

Nazmus Sakib was responsible for the segmentation. Elizaveta Kharlova was responsible for the PCA.

Segmentation

Segmentation is a very important part before the recognition. The accuracy of the PCA recognizer largely depends on proper segmentation. In our system we tried to get the grayscale image of hand subtracting the background.
The input video was captured in RGB. Then it was downsized with nearest neighbor interpolation. We chose nearest neighbor because it is fast despite the reduction of quality of the image. Each of the RGB components were scaled like the grayscale in order to compensate the illumination. This uniform scaling helps thresholding the skin color and subtracting the background. The thresholding was done on YCbCr image to form a new binary image. The binary images can show head, hands and any regions that look same as skin color.Then we use a binary label set algorithm to get regions.So the regions of interest were bounded with bounding box procedures and sorted with respect to area Figure (1). In our system we take the box with the largest area where we assume that the human hand will be closer to the camera than his head.
 Next we cropped the image based on the boundary. Since the content in the box depends on skin color thresholding, there are some gaps in between the regions. So we used flood
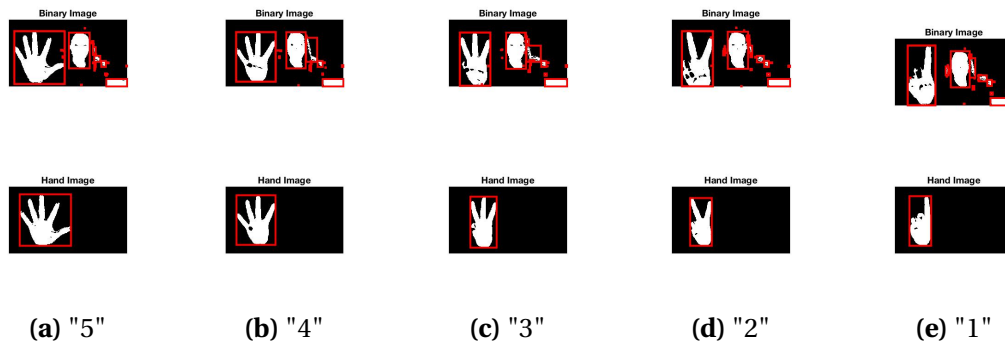
(a) "5"   **(b)** "4"   **(c)** "3"   **(d)** "2"   **(e)** "1"

Figure 1: Hand Image Segmentation

filling algorithm to fill the holes inside the hand image. Now we took the grayscale pixel corresponding to the binary image. This process is a repetitive over each frame of the video sequence. The images were passed into PCA process for recognition. The following Figure(2) shows the activity diagram of the whole procedure.
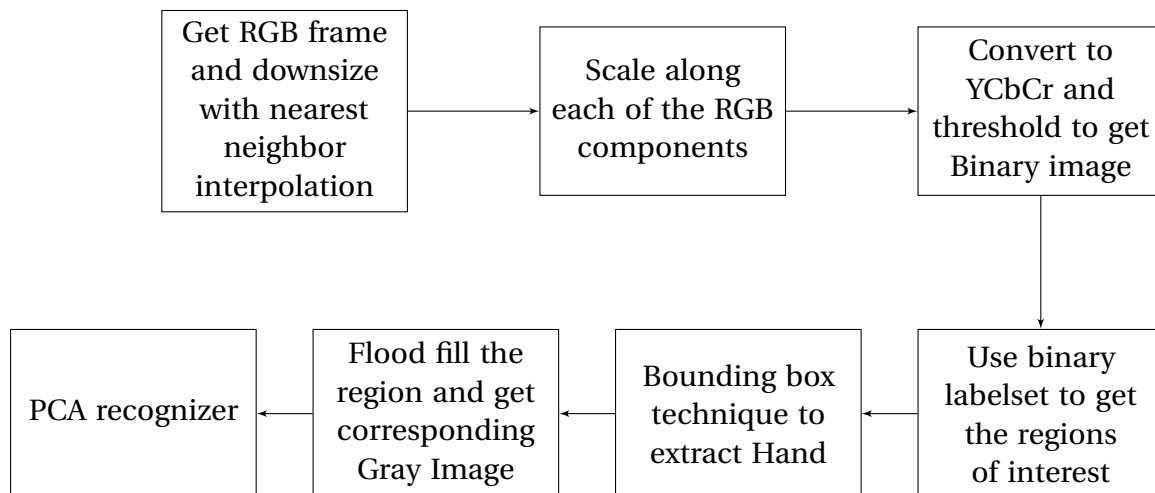

Figure 2: Segmentation Activity Digram

PCA

Once the segmented grayscale images were loaded into one matrix, they were resized to an arbitrary average size (nxm) of the received images (here 100x80). They then were passed onto the PCA processing. This procedure is similar to the one described in [2] and [3].
The first step was to rearrange each image such that it forms a column in the following way,

where k is the number of images

$$N_{nxmxk} = [Y^1_{nxm} Y^2_{nxm}...] => M_{(nxm)xk} = [I^1_{n*mx1} I^2_{n*mx1}...] \tag{1}$$

Then mean of the resulting matrix was subtracted based on (2).

$$M = [I_1 - I_{mean}[...]I_k - I_{mean}] \tag{2}$$

Then the covariance matrix was calculated.

$$C = M^T * M \tag{3}$$

This way the size of the covariance matrix is reduced to kxk, corresponding to the number of images unlike in [2] and [3], where the covariance matrix is the size of the image column. Decreasing the size of the covariance matrix reduces computational complexity.
The next step is to calculate the eigenvalues of the covariance matrix, and determine the basis by (4), where V is the matrix of all eigenvectors of M.

$$B = M * V \tag{4}$$

The weights are then calculated for the future comparison.

$$W = B^T * M \tag{5}$$

This concludes the calculations of the basis and weights for the training data set. We used five gestures corresponding to numbers 1 to 5 and ten training images for each gesture. Therefore, the training set size is 50 images, and the covariance matrix and the weight matrix have size 50x50. A larger data set leads to excessive variance in the errors, which makes the gestures indistinguishable from one another.
The second part of the process is to acquire a new image and determine to which gesture it corresponds. Firstly, the image is processed using (1) and (2). Then its projection on the existing basis set is calculated.

$$P = B^T * image \tag{6}$$

Once the projection is calculated, the Euclidean distance between the projection and the weight vectors is found. The error is found for each weight vector corresponding to a certain gesture. The errors are then summed for the vectors corresponding to the same gesture.
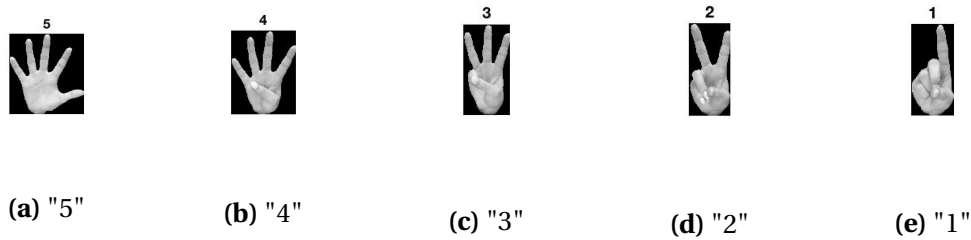
(a) "5" **(b)** "4" **(c)** "3" **(d)** "2" **(e)** "1"

Figure 3: Recognized gestures

The minimum error has maximum covariance with the basis, and therefore, the gesture is detected.

RESULTS

In Figure (3) you can see that the gestures have been correctly detected by the algorithm. The significant decrease in computational complexity is followed from the fact that the training basis set is calculated separately and before running the real time recognition algorithm. The procedure showed significant improvement upon the inclusion of the segmentation and background subtraction. Otherwise, the changes in the background and lighting conditions affected the results and caused large errors in the algorithm. PCA is a good choice with respect to the computation complexity. We assume it performs better than popularly used Hidden Markov Model (HMM) methods in [4]. Although the accuracy might be compromised.

Moreover, during the testing procedure it was determined that the gestures are best recognized when the training set is based on the same person's gestures. This depends on the fact that the PCA calculates the basis based on the pixels in the image, and therefore will have different results. However, it recognizes gestures, which are more distinct, from a different person. The fact that the images have all been resized and appropriately processed leaving only a gray-scale palm of the hand and subtracting the background helps detect the covariance.

Thus, a proposal of future work is to configure the algorithm by first asking the user to demonstrate the necessary gestures, and taking those frames as the training images. This will also lead to the algorithm being modular in a way that the user can input any gesture, which will later be recognized. Another way of improving the algorithm is to implement hand tracking such that the bounding box rotates together with the hand, which will significantly

decrease variance for the PCA.

CONCLUSION

PCA gesture detection provides a good opportunity for implementing a real time algorithm. The first part was the segmentation and background subtraction and the second part was the application of the PCA to the received image and the comparison to a previously trained basis set. The proposed algorithm has been implemented and tested. The results were presented. The recognition was satisfactory, but had larger variance for different users. The future improvements were proposed.

REFERENCES

[1] Dardas, Nasser H., Petriu, Emil M. "Hand Gesture Detection and Recognition Using Principal Component Analysis," in *Proc. IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, 2011.

[2] Ahuja, Mandeep Kaur *et al.* (July,2015). "Hand Gesture Recognition Using PCA." *International Journal of Computer Science, Engineering and Technology*. [Online]. 5(7), pp. 267-271.

[3] Kota, S.R., Raheja, J.L., Gupta, A. "Principal Component Analysis for Gesture Recognition using SystemC," in *Proc. International Conference on Advances in Recent Technologies in Communication and Computing*, 2009.

[4] Chang-Yi Kaoa and Chin-Shyurng Fahn. "A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion" in *Advanced in Control Engineering and Information Science*. Procedia Engineering 15 (2011),3739-3743.