Facial Recognition Using Color Discrimination and Hough Transform

Connor McCullough, University of Miami

Fig. 1. Unprocessed photo used for Facial Recognition

***Abstract*—Practical application of facial recognition must often be done in real time with low latency. This paper proposes use of the Hough Transform for facial recognition using pre-processing methods, parameter bounds, and dimension reduction so that the Hough Transform algorithm can perform at optimal speeds. Pre-processing includes splitting each detected object into its own bounded image before processing, as well as setting ellipse proportions to a constant to take advantage of universal proportions of faces.**

Keywords—Hough Transform, color discrimination, binary, facial recognition, ellipse, threshold.

I. Introduction

Facial Recognition is an Image Processing technique used in a wide variety of applications today. The process is multi-step and based on both color and shape content of the image. First, color discrimination is used to determine which pixels in the image match fall within the specified range for skin colored pixels based on random sampling. Next, noise removal is used to eliminate false positives and minimize the number of objects processed by the Hough Transform. Edge detection is then performed so only the edges of the faces are used in the Hough Transform. Lastly, the Hough Transform is performed, which converts the image to parameter space and finds ellipses that most closely match the shape of the faces. Because each parameter used in the Hough Transform increases the processing speed of the algorithm exponentially, steps must be taken to ensure that the parameters are properly bounded and that the image is properly pre-processed so that the Hough Transform operates in an efficient manner

II. Color Discrimination and Noise Removal

A. Color Discrimination

The use of color discrimination and thresholding can narrow down the number of objects processed by the Hough Transform to a reasonable number. The algorithm was implemented from scratch without MATLAB's computer vision toolbox by sampling of facial pixels. Instead of extensive data mining on the image to determine the true mean and standard deviation, 4 random sample points were acquired from each face in the picture to get a large enough sample size for color discrimination. The mean and standard deviations of each of these points were found and if the image value fell within the standard deviation from the mean, the point was set to 1 and all other points were set to 0. An extra thresholding parameter was added, allowing the color value to fall within a certain factor of the standard deviation. The optimal threshold in this case was found to be 1.2. Color thresholding was performed for the three RGB channels, creating three binary images based on if the values for each channel fall within the threshold for skin color. The three channels were then concatenated using a logic AND function, so that the thresholding had to satisfy all three channels to return ‘1’.

Fig. 2. Values for colors that fall within the accepted range for red (top), green(middle) and blue(bottom).

B. *Noise Removal*

Although the image has been converted to binary, it is not yet ready to be processed for the Hough Transform. Because the Hough Transform’s processing time increases exponentially with each added parameter, steps must be taken to reduce the processing time as much as possible. The Hough Transform is used to find the closes matches in an image for a given shape, based on the parameters of the shape. However, pre-processing can be done to eliminate objects which are obviously not the correct shape. Noise removal was applied liberally to reduce all unnecessary objects which may slow the algorithm down to the point of freezing MATLAB. Because of the voting system in the Hough Transform, it is ok if some features of the face get eliminated, so if there is any elliptical portion of a face present, it will be recognized by the algorithm.

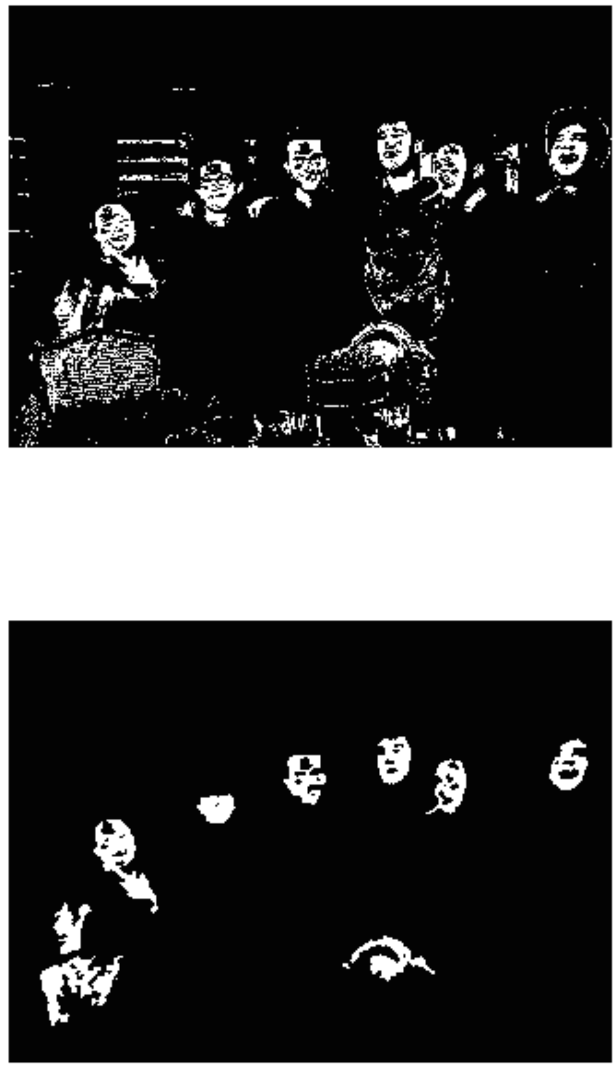

Fig. 3. Composite binary image before noise removal (top) and after (bottom).

The technique for noise removal was similar to that of noise removal for character recognition, with a few parameters tweaked for optimization for this use. There was use of morphological functions, as well as a brute force method of removing any object that does not have a sufficiently large area. This eliminated small patches which tested positive for skin color. Use of bridging and the “majority” function, which fills in holes if most of their neighbors are ‘1’s, connected parts of the face which were detected to be separate objects, so that all faces were found to be objects. At the end of the noise removal, there were only 14 objects left for the

Hough Transform to process, as opposed to the 1000 objects that were present before noise removal. The noise removal also successfully preserved all 6 faces in the image enough so that the Hough Transform can still be used. While there are still false positive objects, these can be removed using the Hough Transform when they are found not to have a matching spherical shape.

Fig. 4. Image after edge detection, ready for processing by Hough Transform

The final step before processing is to use edge detection on the images, so that even fewer points are being processed in the algorithm and matching is optimized even further. Each object is also processed as its own image, as reducing the total size of the image being processed also allows the algorithm to function much more rapidly.

III. Hough Transform

Once the image has been reduced to its most simple form, the Hough Transform can be implemented. The Hough Transform converts an image to parameter space, creating a function of all the parameters and how well the given parameters match the shapes in the image. Basic shapes that the Hough Transform is used on include circles and lines, which are functions of only one parameter each. The processing time for n parameters is x^n, so steps must be taken to reduce the dimensions as much as possible. Luckily, because the Hough Transform using a "voting" system to accumulate each matching point in the image with the given shape, the shapes do not have to be perfect in order to obtain matches and the parameters can have many constraints in order to keep the processing time manageable. After all the potential shapes and their respective votes have been calculated, the shapes with the most votes are plotted over the original picture, using a threshold to choose numbers of votes relative to the maximum number of votes. First, a basic 1-dimensional Hough Transform using a sphere was used. By limiting the threshold to only the top 95% of votes, the function plots one sphere for the majority of the faces, based on the following equations:

$$\frac{x^2}{R^2} + \frac{y^2}{R^2} = 1 \quad (1)$$

$$R = \sqrt{x^2 + y^2} \quad (2)$$

Equation (1) is the general equation for a circle, while (2) is the parametric expression. (2) is used for the Hough Transform as functions must be expressed explicitly in MATLAB. However, as expected, the spheres do not accurately match the shape of the human face.

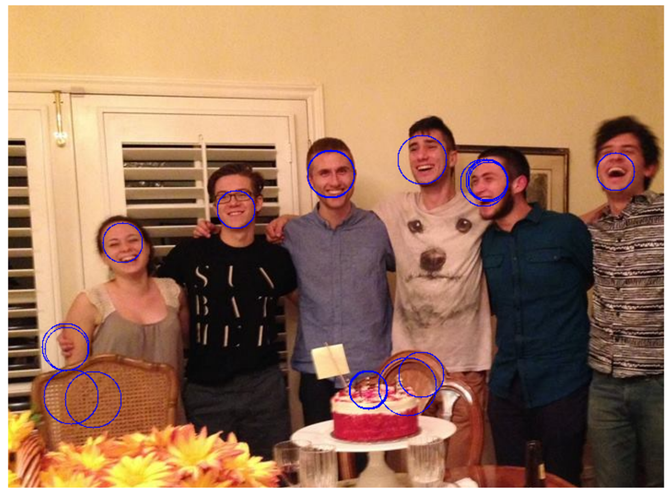

Fig. 4. Image with Hough Transform facial recognition using spheres

In order to get a more accurate outline of the human face, the ellipse equation must be used in the Hough Transform:

$$y = \frac{x^2}{a^2} + \frac{y^2}{b^2} \quad (3)$$

$$y = \frac{x^2}{R^2} + \frac{y^2}{R^2 k^2} \quad (4)$$

$$R = \sqrt{x^2 + y^2/k^2} \quad (5)$$

Equation (3) is the general equation for an ellipse, which uses to different values for radiuses for each axis. By manipulating this equation, the Hough Transform can be done on an ellipse with as little added processing overhead as possible. By expressing b as a ratio of R, only one term is added to the equation, which can still be expressed in terms of R. Setting a range for k will give a variety of widths of ellipses. However, the added parameter will increase the overall algorithm speed exponentially. By examining the human faces by inspection, it can be seen that most of the differences are in overall size, but not proportion. In fact, many sources note that the human face exhibits proportionality based on the golden ratio, approximately 1.618. Therefore, setting k = 1.618 should increase the accuracy of facial detection, while not increasing processing size at all.

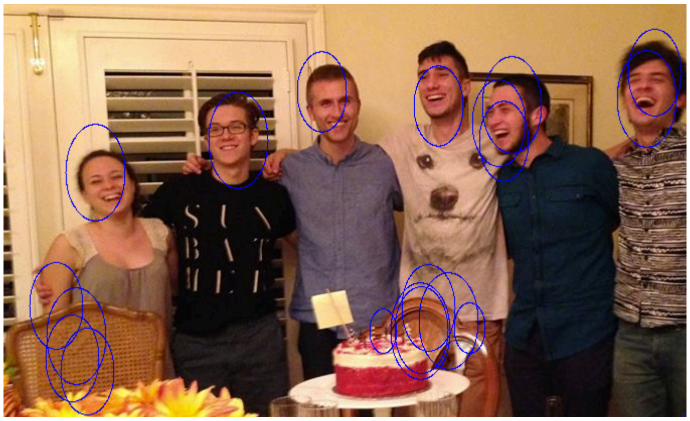

Fig. 5. Image with Hough Transform facial recognition using ellipses.

IV. Limitations

While the described method gives a quick and effective method for facial detection, there are several improvements that could be made to the algorithm. As can be seen, there are several false positives generated by the chairs. These objects proved especially difficult because not only did they fall within the acceptable color range for skin tone, but were a mesh grid, allowing colors not in the range to be inside of the overall shape of the chair. This caused the edge detection to have several holes and not accurately match the shape of the chairs. Because of the amount of edges generated by the mesh, a variety of ellipses were able to pass through the chair with a high enough number of votes to show up on the final image. This could be avoided by more thorough removal of holes in the shapes in order to prevent the grid to show up on the final edge detection. The middle chair proved especially difficult as well, because it both had the appropriate colors to fall into the range of colors, and also had an elliptical curvature similar to that of the human heads. In this case, more extensive detection such as nose, eye, and mouth detection may be necessary in order to eliminate these false positives.

Another shortcoming is that while the ellipse with ratio of 1.618 works well for faces which are facing forward and are straight, it loses accuracy on the faces which are faces the side, or are tilted. The tilting could be solved by incorporating an angle parameter into the equation for the ellipse. The heads facing other directions could be solved by allowing the ratio between a and b in the ellipse equation to be a parameter and not a constant. However, both of these changes together would cause the Hough Transform to become 3-dimensional. The current configuration is the optimal settings for a 1-dimensional Hough Transform.

One final limitation is that the necks of subjects are often included in the Hough Transform ellipse, which offsets the overall shape of the ellipse. In order to solve this, more thorough analysis of the color and shape differences between the neck and head would need to be performed in order to fine tune the algorithm.

Conclusion

The method decribed gives a simple and computationally efficient method of facial recognition, most suitable for detecting faces directly facing a camera. The pre-processing eliminates a significant amount of the proccessing time by reducing the number of shapes to be processed by the Hough Transform, and the bounding and fixing of parameters reduces processing time by an exponential factor. The trade off of processing time versus accuracy of faces under non-ideal conditions (facing away from camera or tilted) can be clearly seen. However, the increased accuracy gained from support of multiple ellipse ratios and tilted images comes at a high cost as the processing speed of the Hough Transform increases exponentially for every parameter used. For basic detection, an ellipse ratio equal to that of the golden ratio seems to have a close match which the majority of faces.

References

[1] http://www.goldennumber.net/face/