IMPACT OF POSE AND GLASSES ON FACE DETECTION USING THE RED EYE EFFECT

Yednekachew Asfaw, Bryan Chen, Andy Adler
School of Information Technology and Engineering, University of Ottawa, Ontario, Canada
yednek@sympatico.ca, Bryan_chen7@yahoo.com, aadler@uottawa.ca

Abstract

In current image-processing algorithms for face detection performance is not completely reliable, especially in situations with variable lighting, and with low-resolution images. One possible approach to implement face detection is the use of the "red-eye" effect: the reflection produced by human eyes when exposed to co-axial infrared (IR) light. We investigated the effectiveness of the red-eye technique for variability in: skin tone, eye color, pose, angle of IR illumination, scene illumination, and the effect of shine from glasses. Algorithms were developed to detect eye locations from a single IR image. Image processing steps involved: normalization, blurring, dynamic threshold calculation, and candidate eye position validation. Average eye position estimation accuracy approaches 80 to 85 percent.

Keywords: *Face detection, Red Eye Effect*

1. INTRODUCTION

The ability to detect faces in a scene is critical to modern surveillance applications [4]. One interesting approach to improving face detection accuracy uses the red-eye effect. Human eyes brightly reflect coaxial infrared (IR) illumination and numerous methods have been developed to exploit red-eye effect. Haro et al [2] used use the physical properties of pupils along with their dynamics and appearance to extract regions with eyes. One limitation of the previous work was that it did not investigate the performance of red-eye effect under different conditions. In our approach, we consider skin tone, eye color, pose, angle of IR illumination, scene illumination, and the effect of shine from glasses.

We use different image processing techniques such as normalization, blurring and localization to improve the likelihood of detecting the eyes within the image.

2. EXPERIMENTAL METHODS

2.1 Experimental Setup

The setup consists of a single black and white camera with zoom lens of 2.5-75 mm and a NTSC output to a frame grabber. The camera is sensitive to the wavelength of the infrared light sources. The PC is equipped with a frame grabber card to capture the recording from the camera. A standard 60 W bulb with variable illuminations is used to adjust the over all lighting of the room. Figure 1 shows the setup of the IR light source.

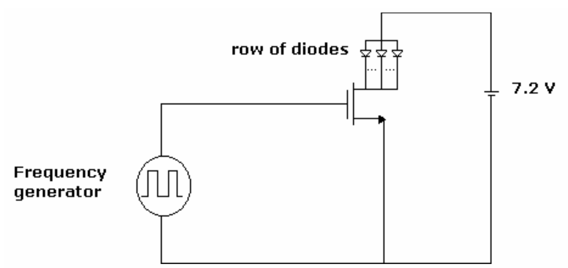

Figure 1: Setup for strobbing IR lighting

2.2 Preliminary analysis

Data was collected for all possible combinations of experimental variables shown in Table 1. All collected data were visually inspection to confirm its validity. We selected 2 frames from the 140 frame per sample, for a total of 48 images per candidate. The frames were picked out using a simple AVI editing program and saved as a JPG file to allow easer analysis in Matlab. Figure 2 shows a representative sample:

CCECE 2003-CCGEI 2003, Montréal, May/mai 2003 -000-
0-7803-7781-8/03/$17.00/© 2003 IEEE

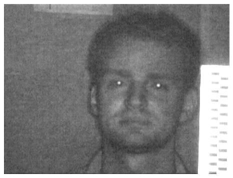

Figure 2: Sample with IR on and no room illumination

Table 1: pose, lighting, glasses, skin tone, eye color

Variable	Condition
Pose	0°,15°,30°,45°
Ambient light	Dark on axis IR, Dark off axis IR, Incandescent on axis IR.
Glasses	With and without Glasses
Skin tone	Five differing levels of skin tone
Eye color	Different types of eye colors

3. IMAGE ANALYSIS

Two different algorithms were developed to determine eye locations from a single frame IR illuminated image. This allowed us to determine the relative accuracies of different approaches to face detection. Both algorithms followed the same basic image processing techniques, but differed in implementation details. Each algorithm uses four stages to the image processing: Normalization, Blurring, Localization and validation.

Normalization: One of the most prominent sources of variability in facial appearance is lighting. Illumination corrections can be applied either by histogram equalization or contrast stretching [3].

Contrast stretching improves the contrast in an image by `stretching' the range of intensity values it contains to span a desired range of values. Histogram equalization uses a monotonic, non-linear mapping which re-assigns the intensity values of pixels in the input image such that the output image contains a uniform distribution of intensities [3].

Blurring: The normalization process has the effect of enhancing background noise. Blurring is used to reduce the effect of background noise. One of the popular methods used for blurring is Gaussian blurring [3]. A Gaussian blur effect takes each pixel in an image and mixes it with adjacent pixels with Gaussian probability. The mean filter is another blurring technique that computes the value of an output pixel by simply averaging the values of its neighboring pixels [3].

Localization and validation: Pupil locations are identified by thresholding of the difference of the dark from the bright pupil images. Once the pupils are detected, it is possible to use holistic approach, which involves template matching, to validate the eyes using global representations [4]. Another approach used was to validate candidate pupil positions using their symmetrical characteristics [4].

3.1 Algorithm one

There are five major steps in this algorithm. The first step is normalization using contrast stretching. This method will stretch the range of the intensity. The bottom one percent (1%) and the top one percent (1%) pixel intensity value of the image are used as the adjustment limits. This provides a dynamic intensity adjustment on the image.

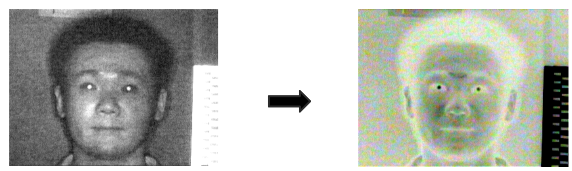

Figure 3: Contrast Stretching and Color inverting

The second step is to use Gaussian blurring method to eliminate the noises with a ten-by-ten Gaussian Filter. After the blurring step, the image is converted to black-and-white image.

The average pixel value of the frame is passed to a polynomial to generate a specific threshold value for black-and-white filtering. The polynomial is created by taking the upper and lower filter value with the average pixel value on a set of testing frames. This allows each image to have its own filter value. The pixels with intensity higher than the threshold value become white, and vice versa.

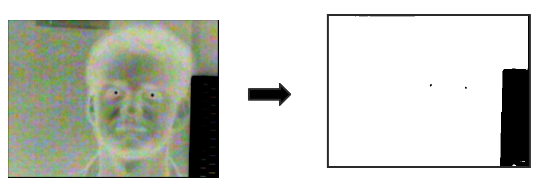

Figure 4: black and white filtering

The fourth step is to localize the eye candidates using template matching. A circular convolution is performed between the image and a pupil template. If a match region is found, the region's color is turned into black. After the template matching, each white region is labeled a unique identification.

The last step is to examine the white regions in the image. The algorithm will check if there are more than two regions to compare with. If the region number is less than two, the black and white filter value will be increased to produce more "candidates" due to low red-eye intensity; otherwise the smallest two regions are verified to meet the following criterions. The areas are less than 0.5 percent of the testing image. The angle between the two region is lass than twenty degrees (20) The distance between them needs to be within nine percent (9%) and twenty percent (20%) of the testing image's width.

3.2 Algorithm Two

This algorithm is divided up into four stages: Normalization, Dynamic Thresholding, Mean Filtering, and Validation.

To create a normalized intensity image, first a uniform background image is created from the original using the morphological opening operator. This function removes small objects from an image while preserving the shape and size of larger objects in the image. The normalized image is produced by subtracting the background from the original.

Dynamic thresholding is used to identify candidate pupil positions. The threshold value is calculated from the mean intensity and the intensity standard deviation of the image. Using the calculated threshold value the normalized image is converted to black and white.

The mean filtering stage is used to eliminate the noise in the background after thresholding. These contributions are then eliminated using mean filtering with a three-by-three kernel.

This process to reduce background noise causes the size of the pupils is becomes smaller. To compensate for this effect we perform dilation of the image to return pupils to their original size. Subsequently, all candidate points are labeled and go through the validation process.

There are three criteria's for validation: distance between candidate regions, angle between candidate regions and the location of the candidate region in relation to the image. The angle between the two candidates has to be less than 20° to the horizontal. Also, the distance between the two regions has to be from forty (40) to fifty (50) pixels. The candidate region has to be inside the active area within ten (10) pixels from the top and bottom of the image.

All regions go through validation process pair by pair until a match occurs. The regions' are first sorted by size in an increasing order, and then the two smallest regions are selected for validation.

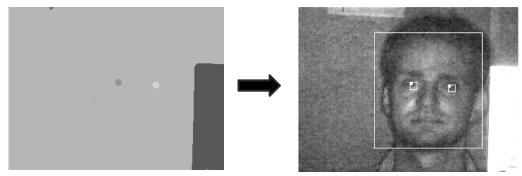

Figure 5: Successful face detection: notice sample with candidate at the top margin

4. RESULTS

These algorithms were used to identify eye locations in each experimental image. The identified positions were then visually inspected and classified as pass for results within ±5 pixels. The average rate of correct identification was 85% for algorithm 1 and 83% for algorithm 2.

By analyzing the algorithm performance as a function of the experimental variables, we identified the factors that affect the performance of the algorithms. Both algorithms were roughly equally affected by each variable. Table 2 shows the level at which each variable begins to significantly reduce the algorithm performance.

Table 2: Settings degrading algorithm performance

Variable	Setting
Pose	45°
Ambient light	Incandescent on axis IR
Glasses	With Glasses
Skin tone	Light skin color
Eye color	Brown

As the angle increases, the cornea of the eye starts to contribute reflections from of the infrared. Since the reflection from the cornea has similar coverage area as the red-eye effect on the pupils, the algorithms choose the reflection of the cornea instead of the pupils. The same result was observed with IR off-axis.

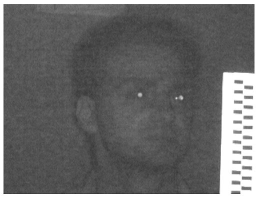

Figure 6: Cornea Effect: From cornea on left eye

The effectiveness of the red-eye effect is reduced by the glasses. Although red-eye effect is visible, the glasses result in dimmer reflection from pupils and other brighter regions on the face becomes “eye” candidates. In addition, reflections were found from eyeglass frames. If the reflection from frame satisfied the validation criterions, the locations of the reflections could be wrongly considered as the eye region.

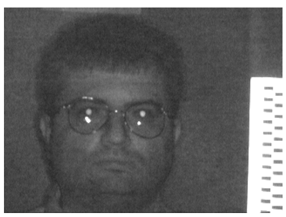

Figure 7: Glasses Effect: notice reflection from glasses

Furthermore, eyes with dark colors can diminish the red-eye reflection intensity level and decrease the rate of face detection.

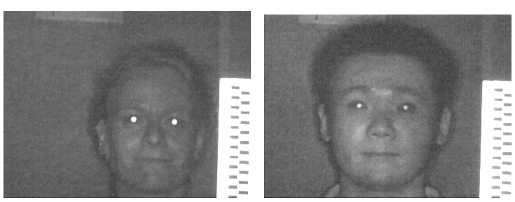

Figure 8: Eye color effect: left blue eye and right brown eye color

6. CONCLUSION

We have investigated the design of face detection algorithms using the red-eye effect. Images were taken of IR illuminated faces under a variety of different experimental conditions. Two different algorithms were developed to identify eye locations in the images. Results were analyzed in terms of algorithm accuracy as a function of each variable. Overall, this implementation provides an 80 to 85 percent success rate, however the performance varied significantly for different conditions.

When the result is categorized according to the different experimental variables, red-eye detection success rate increases as the skin color gets darker, as the eye color becomes lighter and overall lighting condition becomes darker.

The IR placement and glasses degraded the effectiveness of the red eye technique. When the IR is placed off axis the pupils do not shine as brightly as it would for on axis placement. The shielding and reflective from glasses, decreased the success of red-eye as face detection technique.

The performance of the two different algorithms was very similar, and tended be impacted similarly by the experimental variables. This suggests that face detection from a single frame IR image is intrinsically difficult, especially in a surveillance application, where pose, lighting and glasses shine are largely uncontrolled.

ACKNOWLEDGEMENT

We would like to thank Gary Van Beek, and Adrian Moica for experimental and logistical support.

REFERENCES

[1] Morimoto, C.H. , D. Koons, A. Amir, and M.Flickner. "*Pupil Detection and tracking using multiple light sources,*" Image and Vision Computing, vol 18,no. 4, pp 331-335, March 2000.

[2] Haro, Antonio, Myron Flickner, and Irfan Essa. "*Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance,*" IEEE CVPR 2000, pp 163-168, June 2000.

[3] Russ, John C., THE IMAGE PROCESSING HANDBOOK. Raleigh, North Carolina: CRC Press, 1999.

[4] T. Fromherz, P. Stucki, and M. Bichsel, “*A survey of face recognition*,” MML Tech. Rep. 97.01, Dept. Comp. Sci., Univ. of Zurich, 1997