Using Artificial Neural Networks and Feature Saliency to Identify Iris Measurements that Contain the Most Discriminatory Information for Iris Segmentation

Randy P. Broussard and Robert W. Ives

Abstract—One of the basic challenges to robust iris recognition is iris segmentation. To represent the iris, some researchers fit circles, ellipses or active contours to the boundary pixels of the segmented iris. In order to get an accurate fit, the iris boundary must first be accurately identified. Some segmentation methods operate on a pre-processed gray-scaled image, while others use a thresholded binary edge image. The Hough transform is a popular method used to search for circular or elliptical patterns within the image. Many irises are slightly elliptical, and suffer from eyelid/eyelash occlusion, specular reflections and often the pupil and iris centers are not co-located. Each of these issues can cause a segmentation error. This research uses of a feature saliency algorithm to identify which measurements, used in common iris segmentation methods, jointly contain the most discriminatory information for identify the iris boundary. Once this feature set is identified, an artificial neural network is used to near-optimally combine the segmentation measurements to better localize and identify boundary pixels of the iris. In this approach, no assumption of circularity is assumed when identifying the iris boundary. 322 measurements were tested and eight were found to contain discriminatory information that can assist in identifying the iris boundary. For occluded images, the iris masks created by the neural network were consistently more accurate than the truth mask created using the circular iris boundary assumption.

I. INTRODUCTION

To represent the iris, some researchers fit circles, ellipses or active contours to the boundary pixels of the segmented iris. The methods used to identify the iris boundary vary greatly from researcher to researcher. Some methods operate on a pre-processed gray-scaled image, while others use a binary edge mask to determine the boundary. Many use a Hough transform to search for circular or elliptical patterns within the image. Many irises are slightly elliptical, and the pupil and iris centers are typically not co-located. Occlusion by the eyelid and specular reflections can cause the visible boundary to have an irregular shape. Each of these issues can cause a segmentation system, based on a circular assumption, to produce inaccurate results. More advanced boundary representations such as ellipses and active contours can overcome some of these issues, but the iris boundary must be well identified if these methods are to operate accurately. In this research, we combine many of the standard methods used to localize and identify the iris boundary, using an artificial neural network, with the goal of determining which measurements and which iris regions jointly contain to most discriminatory information for iris segmentation. Many existing segmentation methods use a circular Hough transform to localize the iris boundary. This imposes an implicit circular assumption on the resulting boundary. In contrast, the neural network configuration used here identifies iris pixels on a pixel-by-pixel basis using measurements from the region of pixels surrounding the pixel. Since the majority of input features to the neural network do not contain a circular boundary assumption, the resulting boundary is free to take non-circular patterns. This research identifies those measurements that when combined can produce this pixel-by-pixel boundary. The goal of this research is to identify the measurements that discriminate between iris and non-iris pixels, not to create a better boundary.

II. BACKGROUND

Much research has been produced on iris segmentation in recent years. It has been demonstrated that very high identification accuracy rates can be achieved on good images when accurate iris segmentation is present [1]. Non-ideal iris databases often pose a segmentation challenge and the resulting identification accuracies achieved are often orders of magnitudes less than that achieved on pristine images [2-6]. It is widely believed that segmentation research will be a major force that improves iris identification accuracy and usability.

The majority of commercial iris identification systems in use today are based on a methodology was published by Daugman in 1993 [7], and is clearly the most cited method for iris segmentation. This method has been the basis for much of the segmentation research performed in recent years. The method operates in the image domain using integrodifferential operators to identify the maximum partial derivative with respect to increasing radius from a particular center point in the image. The center point is varied and the radius and center point that contains the maximum gradient are selected as the iris boundary.

Many variations to the Daugman method have been
III. METHODOLOGY

An artificial neural network is used to statistically classify each pixel of a polar iris image as either an iris pixel or other. The primary objective is to identify the features that statistically locate the outer iris boundary without using the assumption that the boundary is strictly circular in shape. Common measurements widely used for iris segmentation in the biometrics community are used as input features for the neural network. The neural network is used as a multidimensional statistical classifier to near-optimally combine multiple segmentation measurements into a single statistically based iris boundary estimation.

To determine which segmentation measurements to use as inputs, many measurements were generated and a feed-forward feature saliency technique (described in section VI) is performed to determine which combination of features contain the most discriminatory information. Gradient measurements, brightness measurements, local image statistics (mean, standard deviation, skewness, and kurtosis) and binary edge maps were evaluated for saliency. When computing each measurement, multiple neighborhood sizes and multiple threshold levels were applied when an empirical value was required in the process. The saliency process evaluated each feature and identified a subset of features that jointly contains the most discriminatory information for the problem at hand.

Ground truth for the training and test images were visually created by hand. The feature selection process and the neural network training optimize the feature set and neural network weights to approximate the ground truth. Care was taken to not over-train the neural network so that the resulting iris boundary estimation could slightly deviate from the ground truth boundary to better follow the natural curve of the iris boundary. The premise is that a pixel that statistically belongs to the iris or non-iris class should be classified as such even if its location slightly deviates from the highly structured ground truth boundary. This premise allows the neural network to form oddly shaped boundaries even when circular or elliptical ground truth is used.

IV. IMAGE DATABASE AND TRUTH DATA

Since this is research into the discriminatory information of segmentation measurements, the training set should be large and varied. The test set can be small since its purpose is to verify that information gained during training generalizes to other images. Test and training set size are always an issue when using supervised networks, because ground truth data must be generated (usually by hand) for each image. For this research, a 500 image subset of the University of Bath iris database [19] was used to train the neural network. The original database contains 2000 grayscale near infrared images. The training set consisted of 2 randomly selected images of each eye from 50 individuals. Testing was performed on a different subset of the database. Ground truth was generated using an in-house algorithm, which used local statistics and a circular iris boundary assumption [18]. The ground truth was then visually adjusted by hand to an edge detected version of the original image. Eyelid detection was also used in producing the ground truth.

V. ARTIFICIAL NEURAL NETWORK

A multi-layer perceptron (MLP) feed-forward artificial neural network was used for classification of each pixel. Most current segmentation algorithms contain one or more empirically determined internal thresholds to create an edge map or decision criteria. The neural network offers the advantage that all decision boundaries (thresholds) are multidimensional, near-optimal, and statistically based on the properties of the training data. An artificial neural network can also offer a performance advantage over other statistical classifier for those problems where the feature distributions do not match a simple Bayesian Probability Density Function (PDF) assumption [20] or when the PDFs of the measurements are unknown. The error back-propagation training algorithm was used to train the MLP. This training algorithm traverses the gradients of an error surface to minimize classification error [21]. Using error back-propagation training, the neural network adjusts mathematical hyper-plane boundaries to form a near-optimal discriminate between statistical class distributions in a multidimensional space. In this process, no PDF assumption is made. The neural network learns the statistical distributions of each class during training. A neural network’s classification accuracy approaches the accuracy of a statistical Bayes optimal solution [20] and therefore can near-optimally weight segmentation measurements, based on the discriminatory content of the individual measurements,
to form a near-optimal combination of those measurements. Because of this weighting, iris measurements that statistically have greater discriminatory information content will contribute more to the final boundary decision than measurements that statistically have lower discriminatory information content.

The polar images used to train the neural network were purposely unwrapped from the pupil center and not the iris center. This caused the boundary to take on a snake-like shape in the polar image. This forced the neural network to locate the boundary without making a circular assumption. The salient subset of segmentation measurements, found in the feature saliency process (section VII), were used as the input to the MLP for both training and testing. The number of nodes within the hidden layer of the neural network is kept low to deter memorization and to decrease computational run time. The number of hidden nodes was varied for thoroughness, but ten hidden nodes provided acceptable accuracy.

VI. FEATURE SET

The features used as input to the neural network consist of several of the segmentation measurements commonly used in iris research and fielded systems. The feature set was composed of both binary (edge detected) and real valued measurements. Image gray scale value, mean, standard deviation, skewness, kurtosis, horizontal gradient and vertical gradient measurements composed the core of the real valued features. Fig. 1 shows a list of these measurements and the neighborhood sizes used to compute them. These measurements were taken from the original pixel row (radius from pupil center) and various processed forms of that image. The processed forms of the polar image included Gaussian smoothing, adaptive histogram equalization and local kurtosis measurements. Fig. 2 shows examples of the original and processed polar images. The binary features included regional pixel sums taken from a thresholded integrodifferential processed image, a Canny edge detected image and a thresholded local kurtosis processed image.

Since many of the features used are taken from a region of pixels, multiple square and rectangle region sizes were used for each measurement and all sizes were included in the feature saliency process. Applying the measurements listed in Fig. 1 to the original and processed images for various region sizes produced 124 real valued features. Summing the binary pixels from various regions within the edge detected images produced an addition 198 features. All measurements were taken for each pixel surrounding the suspected outer boundary location. All measurements were taken on images converted from rectangular to polar coordinates referenced to the center of the pupil. Row summation in the polar image was used to approximate the functionality of the circular Hough transform. The incorporation of this large number of features is intended to represent each major aspect of the common iris segmentation methodologies. The feature saliency process could select from either gray-scaled or edge detected measurements taken from either individual or regions of pixels to determine which set of measurements jointly contain the most discriminatory information. Afterwards, the neural network could weight and combine these measurements to best approximate the ground truth of the iris boundary.

VII. FEATURE SALLICENCY

Feature saliency was performed to determine a subset of features that jointly contain the greatest discriminatory information that will distinguish the iris from the rest of the eye. To find this feature subset, a feed-forward, MLP classifier based feature saliency technique was used. The feed-forward technique is a sub-optimal approach, but is computationally practical and delivers excellent results in practice. To begin, the feed-forward technique evaluates the discriminatory power of each feature individually and keeps the single feature that demonstrates the greatest discriminatory power. Next, each feature is individually used in conjunction with the first feature and the pair that contains the greatest discriminatory power is selected. This process, of adding individual features to the group of
previously selected features, is continued until no discriminatory power is gained by adding an additional feature to the previously selected features. Because this is a local search methodology, the results are considered sub-optimal. The resulting feature subset may be optimal, but is not guaranteed to be. For this research problem, using this saliency technique proved computationally expedient and provided acceptable results.

VIII. NEURAL NETWORK AND POST PROCESSING

The MLP was trained on the 500 120x180 pixel image training set which consisted of a total of 10.8 million pixels. The 8 features, selected in the feature saliency process (Fig 4), were used as input to the MLP. To remove any effects caused by the random neural network internal weight settings at the start of training, the neural network was trained 10 times and the weights from the best training run were retained for use in the testing phase.

Fig. 3. Overview of using an Artificial Neural Network to segment the iris.

Fig. 3 demonstrates how the neural network was configured to segment the iris. For each pixel in the iris image, the selected features were computed and presented to the neural network for classification.

IX. RESULTS

Initial results indicated that the local brightness-based features did not hold the discriminatory information necessary to properly segment the images within the test set. Pixel location features, such as pixel distance and angle from the pupil center, were added. The addition of these features allowed the neural network to statistically determine the most likely locations of the iris boundaries and focus its decision making on those pixels.

The test set contained 200 images composed of left and right eyes of 20 individuals. No training images were contained in the test set. The number of nodes within the hidden layer of the neural network was kept low to deter memorization and to decrease computational run time. Ten hidden nodes were sufficient to achieve visually acceptable classification accuracy.

Both real-valued and binary (thresholded) features were tested separately and together. The binary features consisted of 105 various sized (NxM) regional sums taken from a Canny edge detected version of the iris image. Added to these features were 90 measurement from a circular Hough transform applied to one eighth, one forth, one half and full circle regions of the Canny edge detected image. The artificial neural network achieved nearly identical numerical classification accuracies (on the test set when compared to the truth data) when using either the real-valued or binary features. The actual accuracy numbers are not a particularly meaningful metric when evaluating segmentation quality since they are relative to the number of iris and non-iris pixels in the truth image. However, they are useful in making relative comparisons between competing methods. Visually the binary features produced smoother more consistent boundary. Visual inspection of the images also showed it was more desirable for the neural network to follow the natural boundary of the iris than the circular ground truth boundary, and thus a small amount of inaccuracy was desirable. Fig. 4 shows the eight most salient “real-valued” features found in the feature saliency process.

Fig. 5 shows the most salient features, from the combined (real-valued and binary) feature set which totaled 322 measurements, found in the feature saliency process. The saliency process found that, in general, the binary features contain more discriminatory information than the real-valued features. This is likely due to the additional processing that is used to determine and apply a robust threshold.

Radius from pupil center
Presence of eyelid at this pixel
5x1 mean (of Hist EQ image)
1x15 horizontal gradient (of Hist EQ image)
1x15 horizontal gradient
15x15 standard deviation (of Hist EQ image)
5x5 vertical gradient
15x1 kurtosis (of Kurtosis image)

Fig 4. The eight most salient real-valued features found in the feature saliency process

11 pixel wide Hough transform of right half of Canny image
Presence of eyelid at this pixel
3 pixel wide Hough transform of left fourth of Canny image
9 row sum of Canny polar image corresponding to eighth located at 10 o’clock
7 row sum of Canny polar image corresponding to eighth located at 6 o’clock
Original image pixel value
1x5 kurtosis of Hist Equalized polar image
9 row sum of Canny polar image corresponding to eighth located at 8 o’clock

Fig 5. The eight most salient features found in the feature saliency process
In many images, the location of the iris boundaries was subjective due to the gradual color transition of the limbic boundary. As with any iris segmentation research, the subjective nature of the iris boundary can be called into question and the accuracy of the truth masks used to measure classification correctness is subjective. For many images, visual inspection often provided greater insight to correctness than numerical accuracy. In this research, no attempt was made to locate the pupil boundary.

Post processing was used to increase the quality of the neural network produced iris mask. A number of the incorrectly classified pixels were not contiguously connected to the largest grouping of pixels in the mask. The iris boundary, as found by the neural network, was median filtered to reduce a few spurious artifacts. Additional improvements could have been pursued by using basic morphological operations, such as open or close, but no attempt was performed.

Figures 7 and 8 show the neural network produced iris mask overlaid onto the original image. The white line denotes the neural network produced boundary and the green line represents the boundary overlaid onto the original image of the eye. The same eye was used to generate figures 6 and 7. Visible behind the neural network produced boundary is the circle representing the ground truth (without eyelid truth). The two crosses denote the location of the iris (green) and pupil (yellow) centers. Figure 7 contains upper and lower eyelid occlusion while figure 8 contains blurring, rasterization artifacts, eyelash occlusion. Both irises have slightly non-circular boundaries and non-concentric centers. Figure 8 is an image of an Asian eye, which is often more difficult to segment than eyes of other races.

Figure 6 shows the truth mask and neural network produced iris mask are presented for comparison. The images in figure 6 represent typical results from the test set. Space precludes including more of the test set images in this paper. The test set was limited to 200 images due to the time required to adjust the truth masks by hand. Sample images from the remainder of the database gave visually similar results.
Figure 9 shows an iris which contains a large amount of occlusion. The eyelid occludes a large portion of the upper iris and the remainder is occluded by eyelashes. This iris also contains a non-circular boundary and non-concentric centers.

X. CONCLUSIONS

Neural networks and feature saliency techniques can be used to identify those measurements that can best contribute to accurate iris segmentation. The feature saliency results give much insight into a near-optimal approach to localizing the iris outer boundary. The process occurs in an iterative fashion that roughly localizes, and then iteratively refines the boundary. The following are the author’s educated, subjective interpretations of the results. First a large area measurement (Hough transform using a half circle) is used to obtain a rough circular location of the boundary (even if iris center is different than pupil center). Half a circle is used because it will fit a non-circular boundary better than a full circle. Next, the eyelid occluded regions are removed from consideration. Third, small arcs (one forth and one eighth circles) are used to complete the opposite side of the boundary (forming a non-circular or circular boundary). Next, measurements for pixels or small region of pixels are used to further refine the boundary. Last, a small arc from the bottom of the iris boundary is used to remove any reflections that typically occur on the lower eyelid. Upon all visual inspections, the neural network improved segmentation accuracy over the circular assumption based segmentation algorithm. On occluded images, the iris masks created by the neural network were consistently more accurate than the iris mask created using the circular iris boundary assumption and eyelid detection. Visually, the neural network segmentation results approached the perceived accuracy of the manually created truth masks. On some images, the iris boundary was so vague that one or more of the core segmentation methods failed to identify the boundary. By combining measurements, the neural network was able to detect a boundary for all images within the training and test set.

REFERENCES