Statistical Pattern Recognition of the Iris

James Greco
David Kallenborn
Dr. Michael C. Nechyba

Abstract **– A method for applying pattern recognition techniques to recognize the identity of a person based on their iris is proposed. Hidden Markov Models are used to parametrically model the local frequencies of the iris. Also discussed is a transform of the iris image from two to one dimensional space and overcoming limited data with the generation of synthetic images.**

I. Problem

IRIS recognition is a problem that has received much media attention recently with the public's heightened awareness of public security. Iris recognition is the identification of a person's identity based on an image of their eye. This is done by analyzing the various patterns that makeup the iris.

These patterns are ideal for biometric identification because they are both hard to alter as well as exceptionally complex. It has been shown that iris patterns are stable from about one year of age until death, meaning that the patterns on the iris are relatively constant over a person's lifetime. [1] Another important feature of the iris is the great variability over many eyes. Each iris is unique (even the left and right eyes are different) and similar to fingerprints – Iris patterns are not shared between identical twins.

II. Previous Solutions

THERE have been numerous studies on the validity of using iris detection as a means of identifying a person's identity. The most successful and only complete solution is John G. Daugman's iris code method [1].

Daugman's algorithms localize the iris by using an integro-differential operator (Similar to the one implemented in *Sec. IV*) that identifies circular boundaries. This operator was chosen due to the circular nature of the eye, *Fig 1*. [6] Daugman used this operator to identify the pupil-iris boundary, the iris-sclera boundary, and the eyelid boundaries.
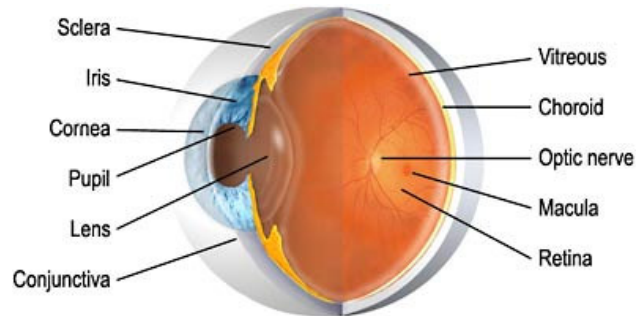

Fig. 1 – Anatomy of the Eye

After finding the iris in the image and the boundaries that enclose it, Daugman applied a Gabor wavelet transform to extract texture information from the iris. He then applied a threshold to the wavelet phase coefficients to generate a 512-byte "iris code". This code is compared to other iris codes in the database by using an XOR operation at a rate of over 100,000 comparisons a second. Daugman's company (Iridian Technologies) has a monopoly on the iris recognition field and claims to have never had a false match with the system on over a billion cross-comparisons. [2] Iris recognition is widely believed to be the best biometric recognition system because of this reliability.

Other solutions [3] include variations of this same technique and have been implemented with moderate success in non-commercial settings.

III. Overview of System

NONE of the previous solutions to the iris recognition problem have explicitly used pattern recognition.
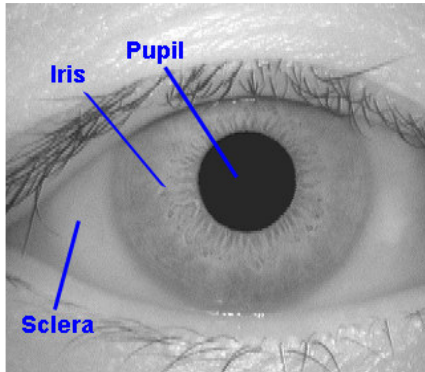

**Fig. 2 - External View of the Eye**

Recognizing a person's identity from an image of the eye has a few problems. First of all, the iris may be partially occluded by the eyelids and eyelashes, *Fig 2*. Thus, the iris may not be completely visible for classification. Also, individual images of the same eye will have slight variations in pupil dilation which causes a stretching of the iris and may not have the same orientation (i.e. rotation and translation of the iris).

The training and testing data was obtained from CASIA [7]. The eye database contained 108 different irises, photographed three times each.

Summary of System:
(1) Pre-processing is done on the image to extract an Iris Map which contains a normalized image of the data in the annular region defined by the boundaries of the iris.
(2) Synthetic data is generated from the Iris Map. The synthetic data generated for this project would correspond to multiple sequential images taken in a system with a video camera.
(3) Data is collected for feature extraction by sampling the iris in 50-pixel increments at concentric circles. (In the Iris Map, this corresponds to sampling rows spaced at constant intervals) A windowed Fourier Transform is then performed on the resulting data. The power vectors are then segmented into eight observables using Vector Quantization, to generate an observation sequence.
(4) In training, the sequences are used in developing the Hidden Markov Model for that iris. In testing, the sequences are evaluated for every HMM in the database.
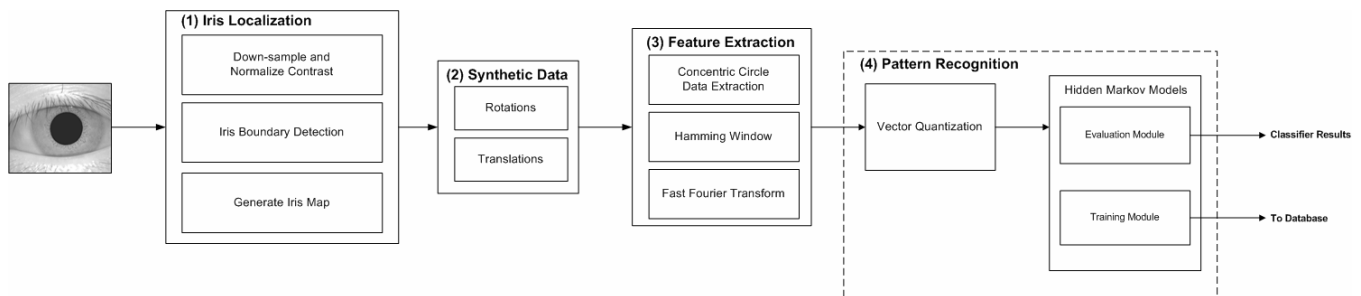
Overall Structure of the Iris Recognition System

IV. Iris Localization

PRE-PROCESSING begins by down-sampling the image by six and normalizing to a standard brightness. Two thresholds are taken to bring out the transition from pupil to iris and iris to sclera. The Sobel gradient is then applied to the thrsholded images to detect the high frequency transitions found along the boundaries of the iris. A discrete approximation of an integro-differential operator (1) is applied to the transformed image.

$$\max_{(r,x_0,y_0)} \left| \frac{\partial}{\partial r} \oint_{(r,x_0,y_0)} \frac{I(x,y)}{2\pi r} ds \right| \quad (1)$$

The operator is a voting algorithm similar to the evaluation of the Hough Transform that selects the best fitting inner and outer iris boundaries. The algorithm is robust enough to work on irises heavily occluded by the eyelids as demonstrated in *Fig. 3*.

With the boundaries detected, the iris data is radially mapped to a standard 450x50 image. Essentially, the iris is "unwrapped" clockwise to a polar coordinate system (r, θ). Data is compressed or stretched when creating the map to fit it to a normalized size. 450x50 was chosen experimentally to minimize the data distortion.

The down-sampling of the image will cause the boundaries of the iris to be off by as much as 6 pixels. With un-optimized code, the entire iris localization process to generate Iris Maps takes a 1/2 second on a 900 MHz Pentium 3. This compromise between resolution and time was found to be the best for our system.

V. Feature Extraction

PREVIOUS solutions to the iris recognition problem use wavelets as a means of extracting texture information from the iris. Our system uses a Fast Fourier Transform (FFT) as a substitute; a wavelet system is beyond the scope
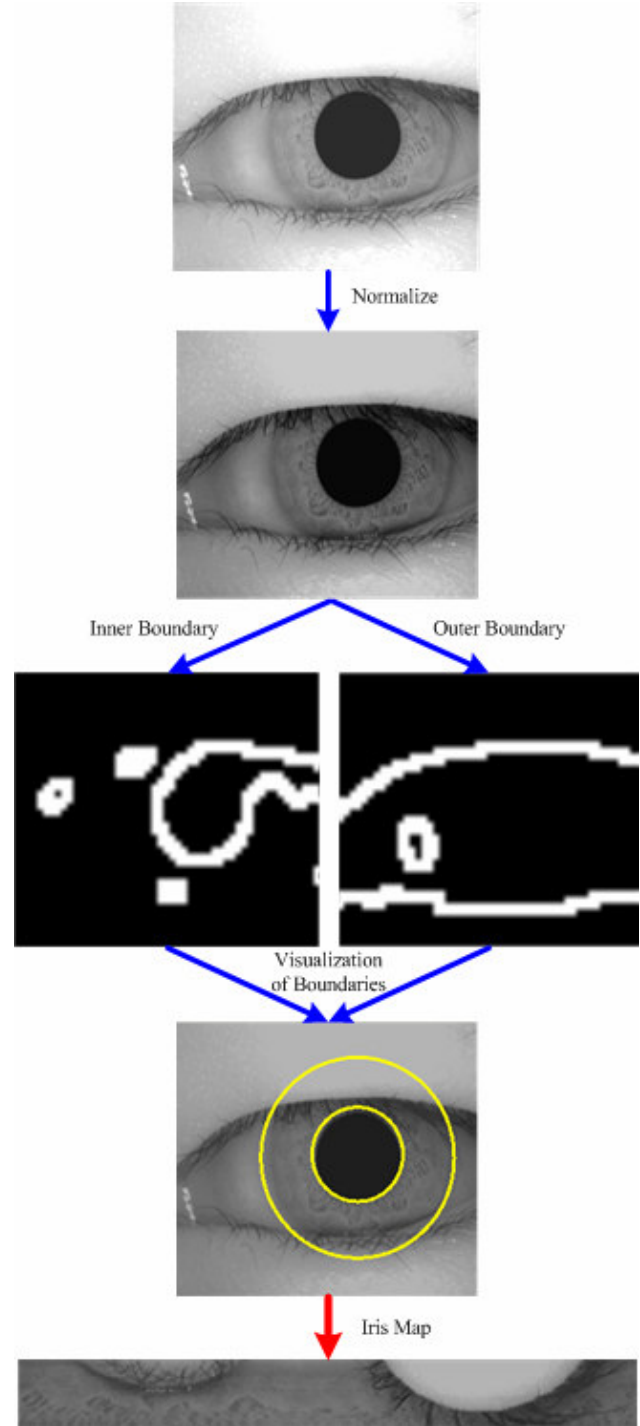

Fig. 3 – Iris Localization System

of this paper; however, it would improve classification performance.

In testing, mid-to-high frequencies were found to be the most important features in classification. Upon first glance of an iris, the mid-to-higher frequency content appears to be

concentrated near the boundary of the pupil. The eye from *Fig. 4a* was sampled in intervals of 50 along a log spiral. An FFT of the individual samples was then taken and the mid-to-high spectral bands were averaged to produce the results shown in *Fig 4b*.
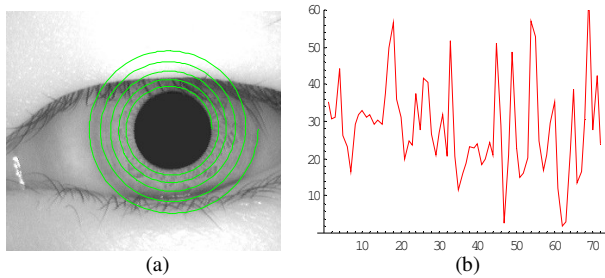

(a)   (b)

Fig. 4 – Mid-to-High Spectral Content

The spikes on the outer bands are largely due to the interfering eyelids. With this in mind, logarithmic feature extraction techniques were used in early tests of the system.

As shown in *Fig. 4a*, data is sampled from the iris by starting at the pupil boundary and expanding outwards in a log spiral to the sclera boundary. The log spiral is defined by in Cartesian form as:

$$x = r_{pupil}e^{a\theta}\cos\theta \qquad 0 \le \theta \le 2\pi * NumSpirals$$
$$y = r_{pupil}e^{a\theta}\sin\theta \qquad a = \frac{\log(r_{sclera}/r_{pupil})}{2\pi * NumSpirals} \quad (2)$$

The log spiral has four advantages.
  (1) Reduces the dataset from two to one dimensional space.
  (2) Eliminates discontinuities in data caused by sampling with circles.
  (3) Concentrates on the higher frequency content of the inner iris.
  (4) Reduces the likelihood of invalid eyelid and eyelash data.

While in theory logarithmic sampling methods should have performed better than concentric methods, in practice the corruption of data from the eyelids and eyelashes could be taken advantage of in the model. The high frequency transitions from iris to eyelid/eyelash made the outer portion of the iris just as important as the inner iris as long as the test image had a similar level of iris visibility as the training image. In a more expressive system, the eyelids and eyelashes would have been removed from the Iris Map and not accounted for in the model.
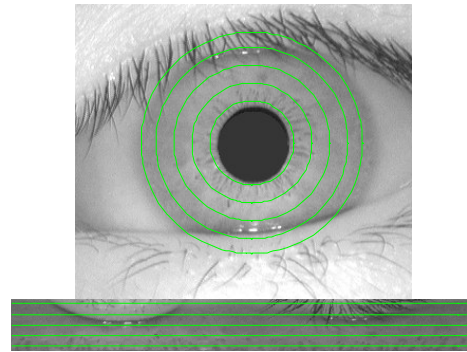

Fig. 5 – Data Sampling at Concentric Circles

Data is sampled in the final system by taking the brightness levels in 50 pixel intervals along the concentric circles. (With 50% overlap). The brightness vector is then multiplied by a 50-length Hamming Window to prevent aliasing.

$$H[n] = 0.54 - 0.46\cos\left(\frac{2\pi(n-1)}{49}\right), n \in \{1,2,..,50\} \quad (3)$$
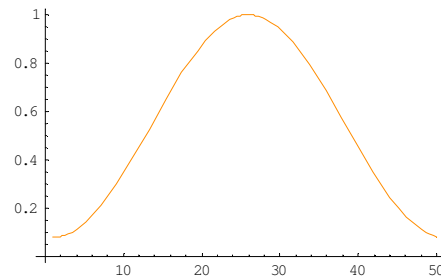

Fig. 6 – The Hamming Window

After applying the Hamming Window, a Fast Fourier Transform is evaluated on the resulting data. Only the magnitudes of the twenty-five coefficients, corresponding to the positive frequencies, are used since phase information was found to give poor classification results.

VI. Data Synthesis

LIMITED training data (2 images per eye) severely limited the performance of the iris recognition system. If the test image was tilted slightly or the iris localization program slightly misclassified the pupil boundary, there was almost no chance of correct classification. More iris data needed to be generated in order to allow for a more robust model. Due to the cost of capturing new images, generating synthetic Iris Maps was the only reasonable option.

The Iris Map is very conducive towards this modification. Tilting the eye clockwise is equivalent to shifting the Iris Map to the right and tilting counter-clockwise is equivalent to shifting the iris map to the left. Starting measurements closer to or farther from the pupil boundary is equivalent to shifting the Iris Map up and down respectively. Twenty-four synthetic images are generated uniformly over a -10° to +10° iris tilt (-12 to +12 pixels) and a -4 to +4 radial pixel shift.

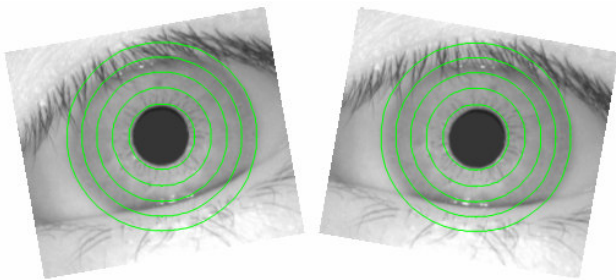**Fig. 7 – Rotations of ±10°**

The synthetic and real images are then Vector Quantized to generate multiple observation sequences for the training and testing of the model.

VII. Modeling

ALL training images were used to generate the VQ Codebook. The power vectors from the optimal 8-prototype vector codebook are displayed in *Fig 8*.

HMMs are used to parametrically model a single iris. They were chosen primarily for their sequential statistical dependence. This dependence allows us to model position and local frequencies of the iris. As can be visually seen
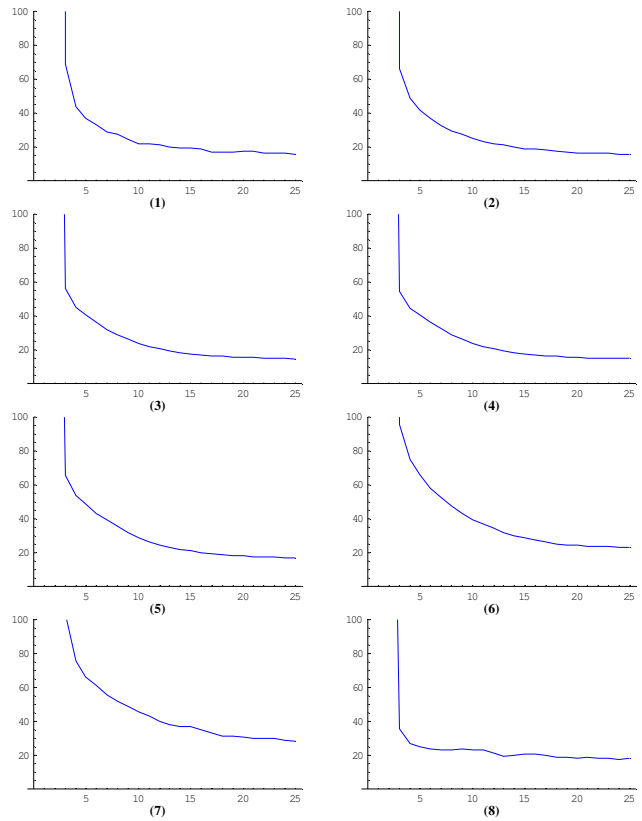
Fig. 8 – VQ Codebook

from the codebook, there is not a great amount of variability in iris frequency that can be detected by the FFT. (Wavelets do a much better job)

Our testing of the iris data was limited to ten irises at one time because of memory limitations when generating and testing a large number of models off of 50 images per eye. Ten test sets, containing ten eyes each, were created from the first 100 images in the database.

Through testing, it was found that 4 states and 8 observables produced the best results. *Table 1* summarizes the results for varying the number of states in the HMM over three sets.

	3 State	4 State	5 State	6 State
Set 0	10%	0%	0%	30%
Set 1	10%	0%	20%	40%
Set 2	20%	10%	10%	30%

Table 1 – Classification Error: Varying State Lengths

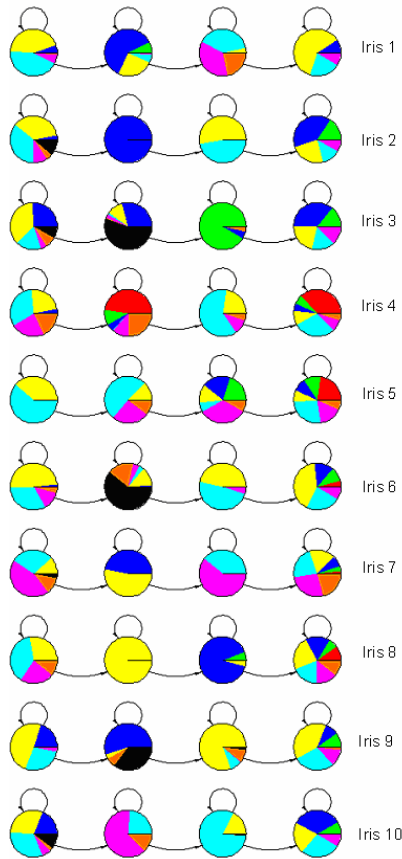The 4 state HMMs for the irises in Set 0 are displayed in *Fig 9*.

Fig 9 – Visualization of HMMs for Iris Set 0

Due to limitations in the iris database, each model is trained off of 2 real images and 48 synthetic images. To stress the significance of generating synthetic data, *Table 2* summarizes results of differing amounts of training data over 3 sets of irises. More real images would have given us the freedom to eliminate the need for some of the synthetic images.

	1 real, 0 synth.	2 real, 0 synth.	1 real, 24 synth.	2 real, 48 synth.
Set 0	70%	70%	20%	0%
Set 1	90%	80%	50%	0%
Set 2	80%	80%	30%	10%

Table 2 – Classification Error: Varying Number of Real and Synthetic Images

VIII. Classification

CLASSIFICATION of a single iris is performed by finding the maximum of the median of P(O| λ) for the set of all synthetic and real images (*n*) over all number of models in the database (*m*).

$$Max[Median[\bar{P}(O_i \mid \lambda_j)]] \quad (4)$$
$$i \in \{1,2,..,n\}, j \in \{1,2,..,m\}$$

In our system, n = 25 and m = 10. The classification results for all ten image sets with 4-state models are listed in *Table 3*.

	% Error	# Misclassified
Set 0	0%	0
Set 1	0%	0
Set 2	10%	1
Set 3	10%	1
Set 4	0%	0
Set 5	0%	0
Set 6	10%	1
Set 7	0%	0
Set 8	0%	0
Set 9	10%	1

Table 3 – Classification Results

96 out of 100 irises were classified correctly.

IX. Analysis

The 96% classification can be somewhat misleading. The images were only classified against a database of 10 iris images. Realistic applications will, obviously have much larger databases, perhaps thousands of irises. However, this is beyond the scope of this project. In contrast, the images that were misclassified were all images that included almost completely closed eyes. In addition, these images were typical rotated beyond the 10º that is taken into account with the synthetic data.

For example, iris 91-3, shown in *Fig 10*, in the CAISA database had almost a 30º tilt. This,

in addition to miscalculated iris boundaries, led to false classification.
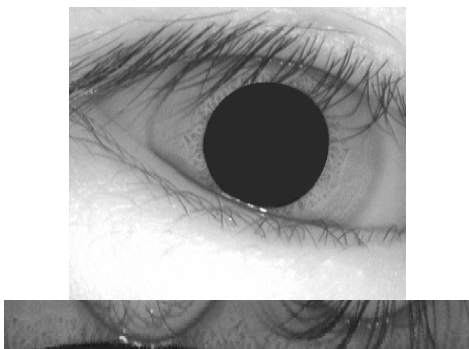

Fig 10 – Typical Misclassified Iris and Iris Map

Notice that the Iris Map contains much distortion in the lower portion. Also, the eyelids and eyelashes occlude or distort almost half of the image. The combination of these factors led to the eventual misclassification of the iris.

Although our approach is different from Daugman's non-parametric method, we achieve comparable results over small datasets. Our system remains untested over moderate to large databases, mainly due to time and computational restraints. One large difference between the two methods, is that Daugman is able to compute orders of magnitude faster than the HMM approach.

X. Conclusions

WHILE the results of the system were successful, there are significant caveats that must be overcome to be used in a real time and robust system.

(1) A smarter heuristic to detect the iris boundary to single pixel precision with little added overhead.
(2) The eyelids and eyelashes must be taken out of the Iris Map. While our dataset was relatively consistent, it would not be reasonable to rely on similar iris visibility. Given more time, this could have been implemented with edge-detection and edge-linking techniques.

(3) When implemented in a camera system, sequential images should be taken in addition to generating synthetic data.
(4) The Markov Modeling must be implemented in C++. Mathematica severely limited our ability to train and test with the large amount of data the models required.

Despite its limitations, this system can be used for small applications, such as personal computer security, where the number of users is relatively small. The system shows much potential for improvement.

Visual C++ source code for the Iris Localization and a Mathematica notebook for the modeling/testing is available at http://plaza.ufl.edu/jimgreco/iris/

XI. References

[1] J. G. Daugman, High Confidence Visual Recognition of Persons by a Test of Statistical Independence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1148-1161, 1993.

[2] J. G. Daugman, How Iris Recognition Works, *University of Cambridge*, 2001

[3] J. M. Ali and A. E. Hassanien, An Iris Recognition System to Enhance E-Security, *Advanced Modeling and Optimization*, vol. 5, pp. 93-104

[4] M. C. Nechyba, An Isolated-Word, Speaker-Dependent Speech Recognition System, *University of Florida*, 2002

[5] L. R. Rabiner, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proc. of the IEEE*, vol. 77, no. 2, pp. 257-286, 1989

[6] M. Erickson, Eye Anatomy, *St. Luke's Cataract and Laser Institute*, 2003

[7] Iris Image Database, *Chinese Academy of Sciences Institute of Automation (CASIA)*, http://www.sinobiometrics.com/resources/