# A Comparative Evaluation of Iris and Ocular Recognition Methods on Challenging Ocular Images

Vishnu Naresh Boddeti
Carnegie Mellon University
Pittsburgh, PA 15213
naresh@cmu.edu

Jonathon M Smereka
Carnegie Mellon University
Pittsburgh, PA 15213
jsmereka@andrew.cmu.edu

B.V.K. Vijaya Kumar
Carnegie Mellon University
Pittsburgh, PA 15213
kumar@ece.cmu.edu

## Abstract

*Iris recognition is believed to offer excellent recognition rates for iris images acquired under controlled conditions. However, recognition rates degrade considerably when images exhibit impairments such as off-axis gaze, partial occlusions, specular reflections and out-of-focus and motion-induced blur. In this paper, we use the recently-available face and ocular challenge set (FOCS) to investigate the comparative recognition performance gains of using ocular images (i.e., iris regions as well as the surrounding peri-ocular regions) instead of just the iris regions. A new method for ocular recognition is presented and it is shown that use of ocular regions leads to better recognition rates than iris recognition on FOCS dataset. Another advantage of using ocular images for recognition is that it avoids the need for segmenting the iris images from their surrounding regions.*

## 1. Introduction

Among the many biometric modalities, iris is of significant interest because of the excellent recognition rates shown when iris images are acquired in controlled conditions, e.g., when subjects are able to place their eyes close to the camera. But in some situations, subjects may be at a distance (e.g., 5 to 10 m) from the camera, may not be looking directly at the camera and may be moving. In such cases, the resulting iris images exhibit impairments such as blur induced by motion and defocus, specular reflections, partial occlusions due to eyelids and eye lashes and non-frontal gaze. The performance of conventional iris recognition algorithms degrades in the presence of such impairments for two main reasons: segmenting the iris from its surrounding regions (pupil and sclera) becomes harder and the nonlinear deformations and occlusions between the probe and the gallery images make it difficult to match the two images even after segmentation.

One way to improve the recognition performance in the presence of above impairments is to use ocular images, i.e., to use iris regions as well as the surrounding regions such as eyebrows, skin texture, etc. Inclusion of the additional regions may help by providing more clues about an individual's identity and some regions may be in better focus than other regions providing more usable information. Another advantage of using ocular regions is that we may be able to avoid the step for segmenting the iris from its surrounding regions, which is often an error-prone and computationally demanding task. So, the goal of this paper is a comparative evaluation of iris recognition and ocular image recognition methods on a common database of challenging ocular images. We use the recently-available face and ocular challenge set (FOCS) [1] containing 9581 images from 136 subjects as the common dataset to evaluate the algorithms.

For both iris recognition and ocular recognition methods, it is important to center the probe and gallery images before they are matched. Toward this goal, we have developed a new correlation filter-based eye detector, which is applied to all images in the FOCS dataset to center them. For iris recognition, we adapt the Chan-Vese segmentation method to extract iris regions and derive a binary code for each iris image using an optimized Gabor wavelet set. The usual Hamming distance-based comparison is used to determine whether the probe and gallery images are from the same class or not. For ocular recognition, there is no need for segmentation and a Bayesian graphical model is designed to obtain match scores that take into account the occlusions and nonlinear deformations between the probe and the gallery images. First, the images are divided into a number of non-overlapping patches and patch-based correlation

This work is sponsored under IARPA BAA 09-02 through the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0013. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of IARPA, the Army Research Laboratory, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

outputs are used to estimate the similarity of the patches as well as the nonlinear deformation between the probe and the gallery ocular images. The information provided by patch-based correlation outputs is input to a Bayesian graphical model which is trained using loopy belief propagation algorithm. The trained Bayesian graphical model is used to obtain a similarity score that reflects the estimated levels of occlusion and nonlinear deformations. We show that the recognition rates for ocular recognition method on the FOCS data set are noticeably better than the recognition rates for iris recognition indicating that the use of periocular regions is beneficial for recognition, particularly for challenging images.

The rest of this paper is organized as follows. Section 2 provides a brief summary of the FOCS data and Section 3 introduces the new eye detector and reviews the pre-processing applied to the images. Iris segmentation, normalization and binary code extraction and matching are discussed in Section 4. The new ocular matching algorithm based on Bayesian graphical models is discussed in Section 5 and recognition results on FOCS dataset are compared in Section 6. Finally, Section 7 provides a summary.

## 2. FOCS Dataset

The ocular challenge data in FOCS data set consist of still images of a single eye region. These regions were extracted from near infrared (NIR) video sequences collected from the Iris on the Move system [9] from moving subjects in an unconstrained environment. The database has 9588 images of 136 subjects with between 2 and 236 images per subject. The database is characterized by impairments such as variations in illumination, out-of-focus blur, sensor noise, specular reflections, partially occluded iris and off-angle iris. Further the iris region is very small (about 50 pixels wide) within an image of resolution $750 \times 600$. Fig. 1 shows examples of some good and poor quality images in FOCS dataset.

## 3. Pre-processing

In this section, we describe the pre-processing applied to the images. Due to the drastic variations in the illumination observed in the database, we first perform illumination normalization via simple adaptive histogram equalization.

Unlike iris images where the eye covers most of the image, the eyes are only a small part of the ocular images. Estimating the center of the eye (center of the pupil to be precise) is very important for aligning the images to be compared. We use a template-based eye detector, more specifically we used a correlation filter-based eye detector. A correlation filter (CF) is a spatial-frequency array (equivalently, a template in the image domain) that is specifically designed from a set of training patterns that are representative of a particular class. This template is compared to a query image by computing the cross-correlation between the template and the query as function of relative shift. For computational efficiency the cross-correlation is computed in the frequency domain. Since the images and their 2D Fourier Transforms (FTs) are discrete-indexed, FT here refers to the 2D discrete Fourier transform (DFT) which is implemented using efficient fast Fourier transform (FFT) algorithm. The correlation filters are usually designed to give a sharp peak at the center of the correlation output plane $c(x, y)$ for a centered authentic query pattern and no such peak for an impostor.

While many different forms of correlation filters exist [15], we use unconstrained correlation filters for eye center detection because they are the simplest to design and they yield performance comparable to the other CF types. The general optimization formulation for unconstrained CF designs is,

$$\max_{\mathbf{f}} \quad \frac{\mathbf{f}^{\dagger}\mathbf{m}\mathbf{m}^{\dagger}\mathbf{f}}{\mathbf{f}^{\dagger}\mathbf{T}\mathbf{f}} \tag{1}$$

where $\dagger$ denotes conjugate transpose, $\mathbf{m} \in \mathbb{R}^N$ is the average of all vectorized 2D FTs of the training images, $\mathbf{f} \in \mathbb{R}^N$ is the vectorized 2D correlation filter and $\mathbf{T} \in \mathbb{R}^{N \mathrm{x} N}$ is a diagonal matrix. The numerator of the objective function is essentially the square of the mean peak value. Different forms of $\mathbf{T}$ lead to different CFs, e.g., using an identity matrix for $\mathbf{T}$ leads to a denominator that is the output noise variance when the input noise is white. The above optimization problem has a closed form solution for the filter,

$$\mathbf{f} = \mathbf{T}^{-\mathbf{1}}\mathbf{m} \tag{2}$$

We consider one particular form of $\mathbf{T}$ namely $\mathbf{T} = \frac{1}{N_1}\sum_{i=1}^{N_1} \mathbf{X_i}^{\dagger}\mathbf{X_i}$ where $\mathbf{X_i}$ is an $N \times N$ diagonal matrix whose entries are the vector versions of the 2D FT of the $i$-th training image. The resulting filter is called the Unconstrained Minimum Average Correlation Energy (UMACE) filter. This choice of $\mathbf{T}$ leads to sharp correlation peaks [8]. We have trained a UMACE filter for eye center detection in ocular images using about 1000 manually labeled images. For about 95% of the 9588 images in the FOCS database, the eye center detector found a point inside the pupil (Fig. 2 shows examples of some eye center detections on the FOCS data). While we are not reporting those details here due to lack of space, this eye detector out-performed state-of-the-art eye detector [2] on FERET [10] data set where the ground truth about eye centers was available.

## 4. Iris Recognition

Binary encoding [4][5] is one of the most popular and best performing algorithm for iris recognition under controlled conditions. However, recognition rates degrade considerably when images exhibit the kind of impairments seen

Figure 1. The top row shows images with good quality ocular regions and the bottom row shows some poor quality images (poor illumination, out-of-focus, closed eyes).

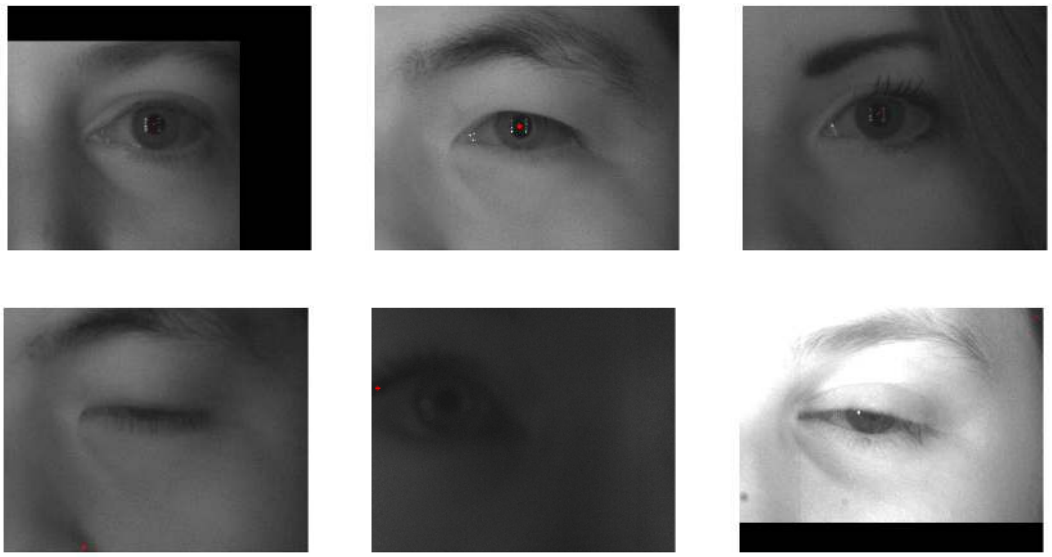


Figure 2. The red dot shows the eye center detected by UMACE. The top row shows example of successful eye center detections while the bottom row shows some failure cases.

in the FOCS database. In this section, we describe the binary code-based iris matching on the FOCS data set.

### 4.1. Iris Segmentation

The performance of iris recognition systems is greatly dependent on the ability to isolate the iris from the other parts of the eye such as eyelids and eyelashes. Commonly used iris segmentation techniques [6] use some variant of edge detection methods and since the blurring due to motion and defocus smudges the edge information to the point of there not being a discernible edge, iris segmentation becomes a challenging task. To alleviate this problem, we use a region-based active contour segmentation based on the seminal work of Chan and Vese [3]. Our technique segments the iris based on the distribution of pixel intensities or features extracted from the eye image from a region both inside and outside the contour rather than looking for sharp edges, making it more robust to blurring and illumination variations than an edge-based active contour. Our segmentation involves two steps, pupil segmentation followed by iris segmentation. We use the output of the eye center detector to initialize a contour for pupil segmentation. Once

the pupil is segmented, we initialize a contour just outside the pupil, to segment the iris.

### 4.2. Normalization

Once the iris boundaries have been found, we map the iris pattern into the polar domain as is popularly done. This has two effects.

1. Normalizes different irises to the same size thus allowing for better matching.

2. Any rotation of the iris manifests as a cyclic shift (along the angular-axis) in the polar domain which can be handled easily via circular shifts of the binary iris code.

### 4.3. Feature Extraction and Matching

Gabor filters with carefully selected parameters have been shown to be the most discriminative bandpass filters for iris image feature extraction among a variety of wavelet candidates [14]. A Gabor filter is a modulated Gaussian envelope and is given by

$$g(\rho,\phi) = \exp\left[-\frac{1}{2}\left(\frac{\rho^2}{\sigma_\rho^2}+\frac{\phi^2}{\sigma_\phi^2}\right) - j\rho(\omega\sin\theta) - j\phi(\omega\cos\theta)\right] \tag{3}$$

in the polar domain ($\rho$, $\phi$) where the filter is applied to the iris pattern. Here $\theta$ denotes the wavelet orientation, $\sigma_\rho$ and $\sigma_\phi$ denote the wavelet widths in radial and angular directions, respectively and $\omega$ denotes the modulation frequency of the wavelet. By varying these parameters, the filters can be tuned to extract features at different scales, rotations, frequencies and translations. We use a set of these differently localized Gabor filters as our feature extraction filter bank. The filter bank used in our experiments has 2 scales and 4 orientations for a total of 8 channels and features are extracted at every point of the unwrapped iris. More details on the parameters of the Gabor filters used and how they were chosen can be found in [14].

The phases of the complex Gabor wavelet projections are quantized to 2 bits by mapping the phase to one of the four quadrants in the complex plane. All the bits obtained this way constitute a binary code. A pair of iris images are compared by matching their respective binary codes. The matching is done by computing the Normalized Hamming Distance between the two binary codes. There are also corresponding masks to identify which bits in the binary code to use for matching. The mask bits are set to either 1 or 0 depending on whether the corresponding code bits are used or not used (e.g., due to eyelid occlusions) for matching. When matching two binary codes A and B with respective masks $m_A$ and $m_B$ the dissimilarity $d$ is defined as,

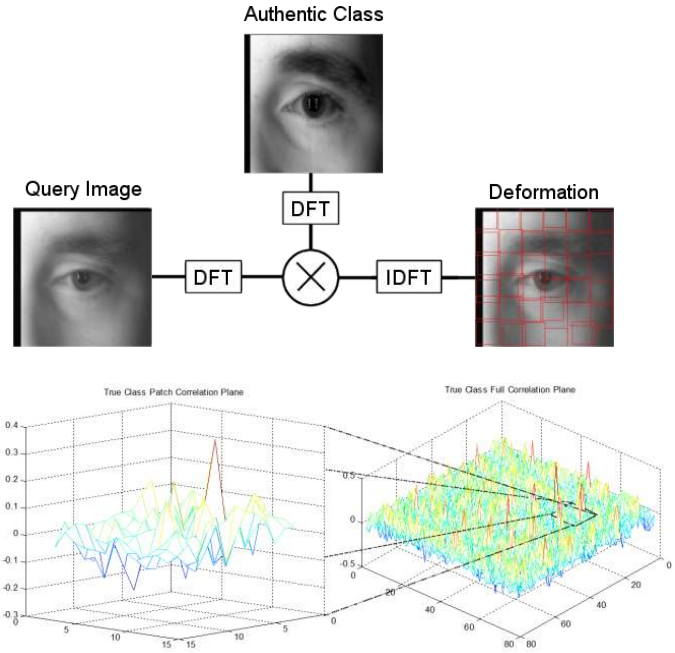



Figure 3. An example of a true class deformation when correlating an authentic class query image. The red boxes are centered on the highest peak in each region to display the shifts that occur, while the corresponding correlation plane and zoomed patch are displayed below.

$$d = \frac{||(A \oplus B) \cap m_A \cap m_B||}{||m_A \cap m_B||} \tag{4}$$

where $\oplus$ denotes an XOR operation and $||\ ||$ denotes the weight (i.e.,the number of nonzero elements) of the binary code. Any rotation of the eye is compensated for by matching the binary codes at different circular shifts along the angular axis and taking the minimum Normalized Hamming Distance value.

## 5. Probabilistic Matching

We adapt the technique originally proposed by Thornton et al. [13] for iris recognition, to ocular recognition. The overall process produces a similarity score between a template and query image that takes into account the relative deformation between the two. The input image is first segmented into non-overlapping patches which are cross correlated with the template using the fusion Optimal Trade-off Synthetic Discriminant Function (OTSDF) correlation filter, which is based on the standard OTSDF filter proposed by Refregier [11]. When comparing a patch from the template with a patch from the query image using the fusion OTSDF the result will produce either a sharp correlation peak at the center of the location of the best match or noise with peaks close to zero. Deformation occurs when the correlation peak is shifted from the center of the image region

as seen in Figure 3. To effectively learn and distinguish a true deformation from just random 'movements', approximation is performed by building a Bayesian graph through maximum a posterior probability (MAP) estimation.

Adapting this technique to ocular images presents its own challenges as the original framework incorporates occlusion along with deformation as hidden states to the observed variable. In iris images, occlusion comes from eyelids, eyelashes, sclera, or parts of the pupil from poor segmentation, while occlusion in ocular images is much more subjective and undefined. For now we only address occlusion as the lost portion of the image from centering based on the output of the eye detector.

### 5.1. Fusion OTSDF Correlation Filter

When segmenting the ocular region into non-overlapping patches, it is difficult to design a robust CF to distinguish, for example, part of an eyebrow of one user from part of an eyebrow of another user. By designing several CFs per region there is an increased chance of obtaining higher similarity values for authentic class images. The fusion OTSDF filter is a powerful multi-channel CF that uses many degrees of freedom to jointly satisfy design criteria leading to robust discrimination for detecting similarities between members of the same pattern class. In contrast to the individual OTSDF filter design, the fusion CF design takes advantage of the joint properties of different feature channels to produce the optimal output plane. Each channel produces a similarity metric based on a relative transformation of the observed pattern and the inner sum represents a spatial cross-correlation between the channels giving an increased chance that the similarity metric will produce high peaks for true class images.

Setting up the fusion OTSDF filter is much like setting up an individual OTSDF filter, with some minor adjustments to account for the additional channels. Ultimately the goal of a CF is to obtain a pre-specified result when correlating an image with a pre-designed filter (note that unless specified otherwise, all operations are in the Fourier domain):

$$\mathrm{x_i}^{+}\mathrm{h} = u_i \tag{5}$$

where $h$ is the filter, $x_i^+$ is the conjugate transpose of the FT of the image $i$, and $u_i$ is the pre-specified filter response to the vectorized 2D Fourier transform of image $i$ (usually $u_i = 1\ \forall i \in$ true-class and $u_i = 0\ \forall i \in$ false-class images). By lexicographically arranging $N$ training images of $k$ feature channels into column vectors and concatenating them into a large training matrix X, the entire system of linear constraints becomes:

$$x^{(i)} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} h = \begin{bmatrix} h1 \\ \vdots \\ h_k \end{bmatrix} \mathrm{X} = \begin{bmatrix} \mathrm{x_1}^{(1)} & \cdots & \mathrm{x_1}^{(N)} \\ \vdots & \ddots & \vdots \\ \mathrm{x_k}^{(1)} & \cdots & \mathrm{x_k}^{(N)} \end{bmatrix} \tag{6}$$

$$\mathrm{X}^{+}\mathrm{h} = u \tag{7}$$

To to produce sharp peaks in the correlation plane and to give the filter a robustness to additive white noise, the CF needs to be designed to minimize average correlation energy (ACE) and output noise variance (ONV). It has been shown [7] that the ACE and ONV criteria respectively, must be minimized subject to the peak constraints resulting in the following form:

$$L = h^{+}Qh \tag{8}$$

in which $Q$ represents the energy between the feature channels when minimizing ACE ($L = E_{avg}$, where $E_{avg}$ is the average energy in the correlation plane), or the covariance of the noise ($n$) when minimizing ONV ($L = var\left(h^{+}n\right)$). Refregier [11] showed that when minimizing for two quadratic criteria ($h^{+}Dh$ and $h^{+}Ph$), the set of optimal solutions may be derived from minimizing a single weighted combination of the two criteria:

$$T = \tau P + (1 - \tau)\, D \tag{9}$$

$$D = \begin{bmatrix} \frac{1}{N}\sum_i X_1^{(i)*} X_1^{(i)} & \cdots & \frac{1}{N}\sum_i X_1^{(i)*} X_k^{(i)} \\ \vdots & \ddots & \vdots \\ \frac{1}{N}\sum_i X_k^{(i)*} X_1^{(i)} & \cdots & \frac{1}{N}\sum_i X_k^{(i)*} X_k^{(i)} \end{bmatrix} \tag{10}$$

where $T$ is a weighted sum of the cross-power spectrum matrix (energy between the feature channels), $D$, and the covariance matrix for the noise, $P$. The parameter, $\tau$, offers a trade-off between ACE or ONV. In contrast to traditional CF designs, fusion CF designs jointly optimize the performance of multiple channels.

After minimizing $h^{+}Th$ subject to $\mathrm{X}^{+}\mathrm{h} = u$, a closed form solution for the $k$-channel CF can be found as (a general form of the solution is derived in [7]):

$$h = T^{-1}X\left(X^{+}T^{-1}X\right)^{-1}u \tag{11}$$

To apply the filter $h$ to a image, it is necessary to reshape the vector to a 2-D image equal in size to the training image, i.e., $h \rightarrow h\,(m, n)$. The test image can be much larger than the training image and may contain multiple targets of interest. When correlating $h\,(m, n)$ with a query image $y\,(m, n)$ the result is the correlation plane $g\,(m, n)$:

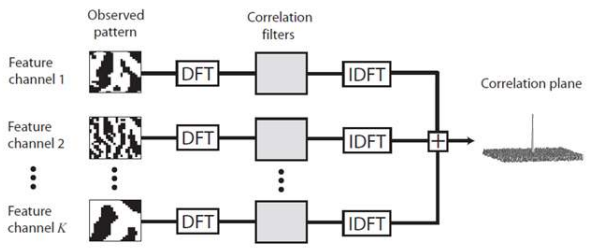



Figure 4. Fusion OTSDF correlation filter

$$g(m,n) = \sum_k \sum_l y(m+k, n+l)h(k,l) \quad (12)$$
$$= y(m,n) \otimes h(m,n) \quad (13)$$

where $\otimes$ is the correlation operator.

### 5.2. MAP Estimation

The overall goal of this process is to authenticate a true class image $I$ by a template $T$ and reject a false class image. CFs can provide a reliable approach to obtaining a similarity measure between a template and query image. However, in the presence of deformations, the matching performance of CFs deteriorates. After independently correlating non-overlapping regions of each image, a good measure of similarity needs to account for any deformations present. One method of doing this, is to use MAP estimation to find the most likely parameter vector $\mathbf{d}$, which describes the deformations, for some (possibly nonlinear) image transformation between $I$ and $T$, assuming $I$ is a true class image. Maximizing the posterior probability distribution on the latent deformation variables results in the following:

$$\hat{d} = \arg\max_{\mathbf{d}} p(\mathbf{d}|T, I, \text{true match}) \quad (14)$$
$$= \arg\max_{\mathbf{d}} p(I|T, \mathbf{d})\, p(\mathbf{d}|T) \quad (15)$$
$$= \arg\max_{\mathbf{d}} p(I|T, \mathbf{d})\, p(\mathbf{d}) \quad (16)$$

Ignoring the normalization constant in Bayes rule and assuming that the deformation $\mathbf{d}$ and template $T$ are statistically independent shows that the image undergoing a specific deformation can be matched to the template. As it is not necessary to learn all possible deformations, the model can be restricted to the values described by the parameter vector $\mathbf{d}$. This restricts the prior distribution to these specific parameters which is defined on a space with low dimensionality and is modeled as a multivariate normal distribution. Specifically, $\mathbf{d}$ is defined so that no deformation occurs at the zero vector, which is assumed to be the mean of the distribution, leaving only the covariance to be estimated.

To determine the deformation parametrization, a coarse vector field model is used, in which the input image $I$ is divided into a set of small regions with corresponding translation vectors $\{(\Delta x_i, \Delta y_i)\}$ and the deformation parameter vector $\mathbf{d} = (\triangle x_1, \Delta y_1, \cdots, \triangle x_N, \Delta y_N)^t$. We can think of the prior term being specific to the pattern-type, and thus $\mathbf{\Sigma_d^{-1}}$ can be estimated directly from matching pairs of the pattern-type. Testing has shown that there are significant correlations between neighboring regions of the vector field and that there is little to no correlation between non-neighboring regions.[13]

Since the generative probability is defined over a large dimensionality (number of pixels in $I$), estimation can become a daunting task. Thus, the fusion OTSDF output $S(I, T;\mathbf{d})$ is used as a similarity measure between the image $I$ and the template $T$ at relative deformation $\mathbf{d}$, setting $p(I|T, \mathbf{d}) = p(S(I, T;\mathbf{d}))$. Rewriting the expression shows this as a minimization problem where the final match score can be computed by taking the similarity function values at the specific estimated deformation.

$$\hat{d} = \arg\max_{\mathbf{d}} \left\{ p(S(I, T;\mathbf{d})) \cdot \exp\left(-\frac{1}{2}\mathbf{d^t \Sigma_d^{-1} d}\right)\right\} \quad (17)$$
$$= \arg\min_{\mathbf{d}} \left\{ -\ln\left(p(S(I, T;\mathbf{d}))\right) + \left(\frac{1}{2}\mathbf{d^t \Sigma_d^{-1} d}\right)\right\} \quad (18)$$

Essentially a Bayesian graphical model is formed as Markov random field (MRF) for the deformation variables as they have spatial significance to neighboring regions. Creating a 2D lattice MRF where each deformation variable represents an image region, a potential function $\Psi_{\mathbf{d}}(\mathbf{d}_i, \mathbf{d}_j)$ can be formed to give some measure of likelihood over the joint values of two separate regions. Being non-negative and approximately Gaussian centered at zero local deformation, we want $\Psi_{\mathbf{d}}(\mathbf{d}_i, \mathbf{d}_j)$ to emphasize vector pairs that are more probable.

$$\Psi_{\mathbf{d}}(\mathbf{d}_i, \mathbf{d}_j) = \exp\left[-\frac{1}{2}\left(\alpha\|\mathbf{d}_i\|^2 + \alpha\|\mathbf{d}_j\|^2 + \beta\|\mathbf{d}_i - \mathbf{d}_j\|^2\right)\right] \quad (19)$$

The $\alpha$ parameter specifies the penalty on the absolute magnitudes of the deformation vectors and the $\beta$ parameter specifies the penalty on the relative difference between the magnitudes of the deformation vectors. Learning $\alpha$ and $\beta$ can be done by using generalized expectation-maximization (EM) algorithm as directly learning the true deformation fields is too difficult, a gradient ascent approach can converge on a solution iteratively. Heuristics show the tuning parameters for the prior converge to approximately $(\alpha, \beta) \approx (0.05, 0.09)$ regardless of initialization [12]. Thus the prior can be defined as:

$$p(\mathbf{d}) = \frac{1}{Z_{\mathbf{d}}} \prod_{(i,j)\in\varepsilon} \Psi_{\mathbf{d}}(\mathbf{d}_i, \mathbf{d}_j) \quad (20)$$
$$= \frac{1}{Z_{\mathbf{d}}} \exp\left(-\frac{1}{2}\mathbf{d^t \Sigma_d^{-1} d}\right) \quad (21)$$

where $\varepsilon$ contains all index pairs of nodes which are connected by edges in the MRF, and $Z_{\mathbf{d}}$ is the normalization constant, making the prior a proper distribution.

### 5.3. Score Calculation

The developed model is designed to assign a higher match score to discernible correlation peaks in similar patterns, and give a lower match score to uncorrelated query images exhibiting seemingly random 'movements'. However, the objective involves estimating the posterior distribution of deformation given the observed image, which can become computationally expensive given the number of values the deformation vector can take. Thus, a variant of Pearl's message passing algorithm is implemented to estimate the marginal posterior distributions at each patch, or node in the Bayesian model. Assuming a sparse, acyclic graphical model, loopy belief propagation (LBP) is an iterative solution to estimating the marginal posterior distribution at each node.

LBP operates on a set of 'beliefs' about the deformation in the model based on'messages' passed between nodes on which there is a direct statistical dependence. These beliefs are a product of the incoming messages from neighboring nodes giving the current estimate of the marginal posteriors at each node by maximizing the likelihood of the deformation values given the observed evidence from the correlation plane at a specific local region. The algorithm continues until the beliefs converge or a maximum number of iterations occur (currently set at four). The resulting beliefs are used to estimate the expected match score.

$$\bar{M} \quad = \quad \sum_{\mathbf{d}_k} S_k\left(\mathbf{d}_k\right) P\left(\mathbf{d}_k \mid O\right) \tag{22}$$

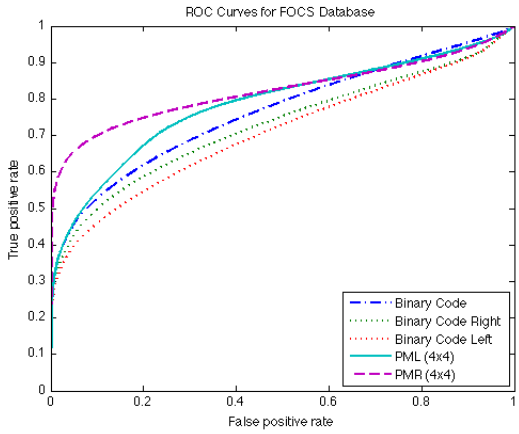
The final score is the summation of the similarity measures, $S_k\left(\mathbf{d}_k\right)$, from correlation multiplied by the marginal posterior distribution of deformation given the observed image, $P\left(\mathbf{d}_k \mid O\right)$, for each image region.

## 6. Score Fusion

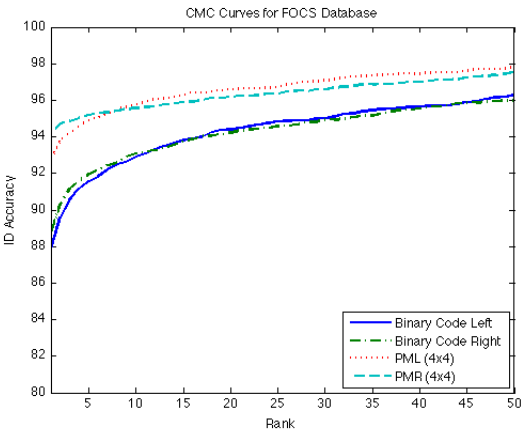
Due to the challenging nature of the database under consideration, we consider a score level fusion of the two techniques under comparison namely, binary iris code and probabilistic matching. In this paper we consider a simple weighted linear combination of the match scores, with the weight being the same for all the match pairs.

## 7. Experimental Results

We tested the two methods, binary iris code and the probabilistic matching (PM) on the FOCS dataset that was described earlier, first pre-processing the images by centering them with the eye detector and applying histogram equalization.




(a) ROC Curves (best viewed in color)




(b) CMC Curves (best viewed in color)

Figure 5. (a) ROC curves for the different experiments (PML refers to probabilistic matching using left graph and PMR refers to probabilistic matching using right graph) and (b) Cumulative Match Characteristic curves.

For the binary code, we computed the full similarity matrix for all the image pairs in the FOCS database achieving an Equal Error Rate (EER) of 30.8%. When we limited the comparison to left iris images versus left iris images we observed an EER of 35.2% and a Rank-1 ID accuracy of 87.9% and for right versus right we observed an EER of 33.1% and a Rank-1 ID accuracy of 88.7%. The full FAR-FRR trade-off in the form of an ROC curve is shown in Fig. 5 along with the Cumulative Match Characteristic (CMC) curve.

For the probabilistic matching algorithm we normalized the size of the images to 128x128 pixels, and divide each image into 36 non-overlapping patches (6x6 configuration). Then the images were randomly separated such that half of

the database, equally distributed between left and right ocular images, were used for training and the remaining for testing in which no distinction between the left and right ocular regions is made during matching. A correlation filter is built for each image using the fusion OTSDF filter with $\tau = 10^{-5}$ and every test image compared against every filter to create a 4791x4791 score matrix. With the trained graph, as we'll call it, we were able to obtain EER of 40.11% and 53.2% Rank-1 ID accuracy.

Breaking down the problem we consider the scenario where a separate graph is built for the left and right ocular regions dividing the images into a 6x6 patch configuration. We use half of the left and right ocular images to learn the parameters of the two graphs respectively. Testing the respective graphs on all the left and right ocular images yielded an EER of 31.28% with a Rank-1 ID accuracy of 86.8% and an EER of 29.74% with a Rank-1 ID accuracy of 84.0% respectively. For this configuration, the score level fusion yields an EER of 29.72% and 28.07% for the left and right ocular images respectively, suggesting that score level fusion does help improve verification performance. These results also suggest that knowing whether a query image is of the left ocular region or right ocular region can benefit the accuracy of the system significantly.

Finally, we consider another scenario where the ocular images are divided into a 4x4 configuration, i.e., each patch is now of a larger size, with everything else being the same. The PM algorithm now yields an EER of 26.81% with a Rank-1 ID accuracy of 92.8% and an EER of 23.83% with a Rank-1 ID accuracy of 94.2% for the left and right ocular images respectively. However, the score level fusion for this configuration does not help improve the verification performance yielding an EER of 26.80% and 23.81% for the left and right ocular images respectively, suggesting that the simple score level fusion considered in this paper is not always helpful.

## 8. Summary

In this paper, we compared two biometric recognition methods on the recently made-available FOCS data set of challenging single eye images. In the first method based on iris recognition, we used an eye detector to center the images, applied adaptive histogram equalization to deal with illumination variations, segmented the iris using an adaptation of the Chan-Vese segmentation algorithm and used carefully selected Gabor wavelets to obtain a binary code to represent each image. In the second approach based on ocular recognition, same eye detector was used to center the images, no segmentation was needed and patch-based fusion correlation filter outputs were input to a Bayesian graphical model that was trained to produce match scores that take into account any nonlinear deformations between the two images. Equal error rates were 30.8% for binary code iris recognition and 26.81% for left ocular regions and 23.83% for right ocular regions, indicating that use of ocular regions may be beneficial when dealing with challenging ocular images. We also observed that the proposed ocular recognition method benefits from the knowledge of whether it is a left eye or right eye whereas the binary iris code recognition method appears to not benefit.

## References


[1] Face and ocular challenge series (focs) http://www.nist.gov/itl/iad/ig/focs.cfm. 1

[2] D. Bolme, B. Draper, and J. Beveridge. Average of synthetic exact filters. 2009. 2

[3] T. Chan and L. Vese. Active contours without edges. *Image Processing, IEEE Transactions on*, 10(2):266–277, 2001. 3

[4] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(11):1148–1161, 1993. 2

[5] J. Daugman. How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):21–30, 2004. 2

[6] J. Daugman. New methods in iris recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1167–1175, 2007. 3

[7] A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent. Minimum average correlation energy filters. *Applied Optics*, 26(17):3633–3640, 1987. 5

[8] A. Mahalanobis, B. V. K. Vijaya Kumar, S. Song, S. Sims, and J. Epperson. Unconstrained correlation filters. *Applied Optics*, 33(17):3751–3759, 1994. 2

[9] J. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. LoIacono, S. Mangru, M. Tinker, T. Zappia, and W. Zhao. Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE*, 94(11):1936–1947, 2006. 2

[10] P. Phillips, H. Wechsler, J. Huang, and P. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998. 2

[11] P. Réfrégier. Filter design for optical pattern recognition: multicriteria optimization approach. *Optics Letters*, 15(15):854–856, 1990. 4, 5

[12] J. Thornton. *Matching deformed and occluded iris patterns: a probabilistic model based on discriminative cues*. PhD thesis, Pittsburgh, PA, USA, 2007. AAI3262821. 6

[13] J. Thornton, M. Savvides, and B. V. K. Vijaya Kumar. A bayesian approach to deformed pattern matching of iris images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 596–606, 2007. 4, 6

[14] J. Thornton, M. Savvides, and B. V. K. Vijaya Kumar. An evaluation of iris pattern representations. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6. IEEE, 2007. 4

[15] B. V. K. Vijaya Kumar, A. Mahalanobis, and R. Juday. *Correlation pattern recognition*. Cambridge Univ Pr, 2005. 2