Chapter 17
Application of Correlation Filters for Iris Recognition

B.V.K. Kumar, Jason Thornton, Marios Savvides, Vishnu Naresh Boddeti, and Jonathon M. Smereka

Abstract Excellent recognition accuracies have been reported when using iris images, particularly when high-quality iris images can be acquired. The best-known strategy for matching iris images requires segmenting the iris from the background, converting the segmented iris image from Cartesian coordinates to polar coordinates, using Gabor wavelets to obtain a binary code to represent that iris, and using the Hamming distances between such binary representations to determine whether two iris images match or do not match. However, some of the component operations may not work well when the iris images are of poor quality, perhaps as a result of the long distance between the camera and the subject. One approach to matching images with appearance variations is the use of correlation filters (CF). In this chapter, we discuss the use of CFs for iris recognition. CFs exhibit important benefits such as shift-invariance and graceful degradation and have proven worthy of consideration in other pattern recognition applications such as automatic target recognition. In this chapter, we will discuss the basics of CF design and show how CFs can be used for iris segmentation and matching.

17.1 Introduction

Among the many biometric modalities proposed to recognize individuals, iris recognition [1] has become a major research direction because of the excellent accuracy iris recognition methods seem to offer [2], particularly when the iris

B.V.K. Kumar (✉) • M. Savvides • V.N. Boddeti • J.M. Smereka
Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: kumar@ece.cmu.edu; marioss@andrew.cmu.edu; naresh@cmu.edu; jsmereka@andrew.cmu.edu

J. Thornton
MIT Lincoln Laboratory, Lexington, MA, USA
e-mail: jthornton@ll.mit.edu

Fig. 17.1 Iris is the
texture-rich region between
the *black* pupil and the *white*
sclera
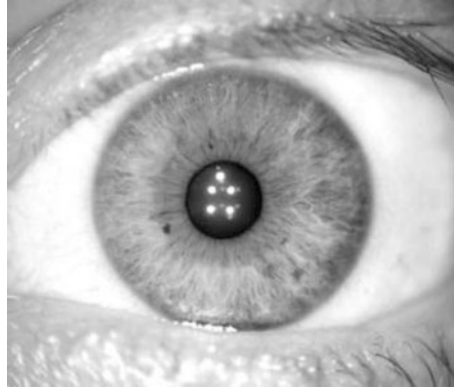
images are of high quality. As shown in Fig. 17.1, iris refers to the texture-rich
region of the eye surrounded by the black pupil (inside) and the white sclera region
(outside). It is believed that iris patterns do not change substantially over a lifetime
and that an iris pattern is unique to an eye, i.e., iris patterns from the left eye and the
right eye of the same individual are different.

The original iris recognition method pioneered by Daugman [3] consists of the
following major components:

- Segmenting the iris from the pupil and the sclera by determining its inner and
  outer boundaries
- Mapping the segmented iris pattern from Cartesian coordinates to polar coordi-
  nates to normalize for iris size variations caused by the pupil dilation and other
  factors
- Producing a binary code from the phases of the inner products of Gabor wavelets
  (of different widths, orientations, locations, and frequencies) with the mapped
  (also called unwrapped) iris pattern
- Using the normalized Hamming distance (NHD) between the binary code of an
  enrolled iris image and the binary code of a query (also called probe) iris image
  as an indicator of the quality of the match between the two iris patterns

The above iris recognition method based on comparing the resulting binary
codes is attractive for its high-speed matching and its excellent accuracy when the
training and testing iris images are of good quality. However, iris images require
near-infrared (NIR) illumination to produce images with good contrast, and NIR
illumination levels cannot be too high because of safety considerations. As a result
of the limitations on illumination levels, the iris images acquired from distance
are not expected to be of sufficiently good contrast. To add to this challenge, iris
images can also exhibit degradations such as occlusions caused by eyelids, non-
frontal gazes, and specular reflections. In such challenging iris images, traditional
binary code matching methods may not work well because iris regions cannot be
easily segmented from the pupil and the sclera regions.

One attractive method to deal with degraded query images is based on correlating the query images with a template carefully designed from training images [4]. This method, known as correlation filter (CF) [5], has proven useful in other challenging image matching applications such as automatic target recognition [6] and face recognition [7]. In this chapter, we will discuss the basics of CF design and show how CFs can be used for iris segmentation and matching.

The rest of this chapter is organized as follows. Section 17.2 provides a brief re view of CF design, and Sect. 17.3 shows how CFs can be used for iris segmentation. CFs can be used for matching segmented iris images as well as unsegmented images, and this is discussed in Sect. 17.4. Section 17.5 discusses using CF outputs in a Bayesian graphical model to improve the recognition performance, and Sect. 17.6 provides a summary.

17.2   Correlation Filter Background

A straightforward measure of the similarity between a probe image $p[m,n]$ and a reference image $r[m,n]$ is the inner product between the two two-dimensional arrays or the inner product between column vectors **p** and **r** obtained by lexicographically scanning $p[m,n]$ and $r[m,n]$. If this inner product (after appropriate normalization) is large, it indicates a high degree of similarity, whereas if this value is small, it suggests that the probe and the reference images are not a good match. However, in practice, the probe image $p[m,n]$ may be a shifted version of the reference $r[m,n]$ necessitating that the inner product be evaluated for all possible shifts between the two images. This leads to the following correlation output $c[x,y]$ which measures the similarity between $r[m-x,n-y]$ and $p[m,n]$ for all possible shifts x and y.

$$c[x,y] = \sum_m \sum_n p[m,n]\, r[m-x,n-y] \qquad (17.1)$$

It can be shown that if the probe image $p[m,n]$ is exactly equal to $r[m-x_0,n-y_0]$, then $c[x,y]$ will have its highest value (called *correlation peak*) at $x = x_0, y = y_0$. Thus, the relative shift between the probe image and the reference image can be determined by locating the peak in the correlation output. If the probe image is from a different class than the class of the reference image, then the resulting $c[x,y]$ will not exhibit a dominant correlation peak, indicating that the two images may come from different classes. The cross-correlation operation above is more efficiently implemented in the frequency domain, i.e., $c[x,y]$ is the 2-D inverse discrete Fourier Transform (DFT) of the product of $P[u,v]R[u,v]$ where $P[u,v]$ and $R[u,v]$ are the 2-D DFTs of $p[m,n]$ and $r[m,n]$, respectively. Here, u and v denote the spatial frequencies corresponding to m and n, respectively, and we use upper case italics to denote frequency-domain functions and lowercase italics to denote image domain functions.The 2-D DFTs are efficiently implemented using fast Fourier transform

(FFT) algorithm. This implementation of the cross-correlation operation via the frequency domain is the main reason for this operation to be termed *correlation filtering*.

In practice, the probe image $p[m,n]$ will differ from the reference image $r[m,n]$ in multiple ways including additive noise, shifts, rotations, scale differences, illumination changes, partial occlusions, and other differences. All these differences cause the correlation peaks to become lower and broader, making it harder to determine the similarity between the two images and to determine the location of the reference image in the probe scene. One method developed to deal with the appearance changes due to such distortions is to design a correlation filter $H[u,v]$ (or equivalently the template $h[m,n]$ in the image domain) that exhibits the following properties of the correlation output $c[x,y]$:

- Correlation output $c[x,y]$ should exhibit large and consistent values at the center in response to centered training images from the authentic class.
- Correlation output $c[x,y]$ should exhibit small values throughout the correlation plane in response to training images from the impostor class.
- Correlation output $c[x,y]$ should take on small values in the rest of the correlation plane so that the controlled values stand out (i.e., yield peaks) in response to centered images from the authentic class.
- Correlation output $c[x,y]$ should exhibit low sensitivity to noise in the probe or query image input.

Over the years, many correlation filter designs have been developed to achieve the above desired properties. For reasons of space, we will focus on the design of one type of correlation filter known as *optimal trade-off synthetic discriminant function* (OTSDF) filter [8]. More details about CF designs can be found elsewhere [5].

Let $r_1[m,n], r_2[m,n], \ldots, r_L[m,n]$ denote L training images (assumed to be centered) from the authentic class with each image of size $M \times N$. The 2-D DFTs of these images are lexicographically scanned to yield column vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_L$, each containing $d = MN$ elements. Matrix $\mathbf{R}$ of size $d \times L$ contains $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_L$ as its column vectors. Similarly, let $\mathbf{h}$ denote a d-dimensional column vector, containing as its elements the scanned version of the correlation filter $H[u,v]$. The goal is to find CF vector $\mathbf{h}$ to meet the above objectives.

The OTSDF design requires that the filter vector $\mathbf{h}$ satisfy the following inner product constraints:

$$\mathbf{h}^T \mathbf{r}_i = c_i, \quad i = 1, 2, \ldots, L \tag{17.2}$$

Typically c_i is set to 1 for all training images from the authentic class and to 0 for training images from the other classes. It is expected that the resulting CF will yield correlation output values (at the origin) close to 1 for centered non-training (i.e., test) images from the authentic class and values close to 0 for other images. The linear constraints in Eq. (17.2) are under-determined in that there are L constraints and d

unknowns, where d (the number of pixels in the training images) is usually much larger than L (the number of training images). Thus, there are infinite solutions to Eq. (17.2). Among these infinite solutions, the OTSDF design tries to find the vector **h** that leads to sharp correlation peaks (sharp peaks make it easy to detect and locate the objects) for authentic class images and that have good tolerance to input noise.

One way to obtain sharp correlation peaks is to force the correlation output values to be small while constraining the value of the peak (i.e., the correlation output at origin given that the input image is a centered training image from the authentic class) to be 1. Mahalanobis et al. [9] tackled this problem by minimizing the average correlation energy (ACE) defined as follows:

$$
\begin{aligned}
\text{ACE} &= \frac{1}{L}\sum_{i=1}^{L}\sum_{x}\sum_{y}|c_i(x,y)|^2 \propto \frac{1}{L}\sum_{i=1}^{L}\sum_{u}\sum_{v}|C_i(u,v)|^2 \\
&= \frac{1}{L}\sum_{i=1}^{L}\sum_{u}\sum_{v}|H(u,v)R_i(u,v)|^2 \\
&= \sum_{u}\sum_{v}|H(u,v)|^2\left\{\frac{1}{L}\sum_{i=1}^{L}|R_i(u,v)|^2\right\} \\
&= \sum_{u}\sum_{v}|H(u,v)|^2 D(u,v)
\end{aligned}
$$

$$
\text{where} \quad D(u,v) = \frac{1}{L}\sum_{i=1}^{L}|R_i(u,v)|^2 \tag{17.3}
$$

The ACE term in Eq. (17.3) can be more compactly expressed as $\mathbf{h}^T\mathbf{D}\mathbf{h}$ where **h** is a d-dimensional column vector containing the filter $H(u,v)$ and **D** is a $d \times d$ diagonal matrix whose diagonal entries are $D(u,v)$. Minimizing ACE in Eq. (17.3) subject to the peak constraints in Eq. (17.2) leads to the following CF known as the *minimum average correlation energy* (MACE) filter [9]:

$$
\mathbf{h} = \mathbf{D}^{-1}\mathbf{R}\left(\mathbf{R}^{+}\mathbf{D}^{-1}\mathbf{R}\right)^{-1}\mathbf{c} \tag{17.4}
$$

where superscript $+$ denotes conjugate transpose and where the L-dimensional column vector **c** is defined as $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_L \end{bmatrix}^T$. Since **D** is diagonal, forming $\mathbf{D}^{-1}$ is easy and the computation of $\left(\mathbf{R}^{+}\mathbf{D}^{-1}\mathbf{R}\right)^{-1}$ involves an $L \times L$ matrix.

While the MACE filter in Eq. (17.4) produces sharp correlation peaks in response to training images from the authentic class, it exhibits high sensitivity to input noise and other appearance variations. This is mainly because MACE filters tend to be high-frequency-emphasis filters (in order to produce sharp correlation peaks) which end up amplifying noise. One way to reduce the noise sensitivity of a CF is to reduce the output noise variance (ONV) defined as follows.If the input

image is corrupted by a wide-sense stationary noise $w[m,n]$ with power spectral density $P_w[u,v]$, then the variance of the correlation output is given as

$$\text{ONV} = \text{var}\{c(x,y)\} = \sum_u \sum_v P_c(u,v)$$
$$= \sum_u \sum_v P_w(u,v)|H(u,v)|^2 \qquad (17.5)$$

where $P_c[u,v]$ is the power spectral density of the noise in the correlation output. Once again, the ONV can be expressed as $\mathbf{h}^T\mathbf{Ph}$ where $\mathbf{P}$ is a $d \times d$ diagonal matrix whose diagonal entries are $P_w[u,v]$. Minimizing ONV in Eq. (17.5) subject to the constraints in Eq. (17.2) leads to the following CF known as the minimum variance synthetic discriminant function (MVSDF) filter [10]:

$$\mathbf{h} = \mathbf{P}^{-1}\mathbf{R}\left(\mathbf{R}^+\mathbf{P}^{-1}\mathbf{R}\right)^{-1}\mathbf{c} \qquad (17.6)$$

While the MACE filter in Eq. (17.4) exhibits sharp correlation peaks and high noise sensitivity, the MVSDF filter in Eq. (17.6) typically exhibits broad correlation peaks and good noise tolerance. Refregier [8] introduced the following optimal trade-off synthetic discriminant function (OTSDF) filter formulation that trades off peak sharpness for noise tolerance:

$$\mathbf{h} = \mathbf{T}^{-1}\mathbf{R}\left(\mathbf{R}^+\mathbf{T}^{-1}\mathbf{R}\right)^{-1}\mathbf{c} \qquad (17.7)$$

where $\mathbf{T} = \alpha\mathbf{D} + \sqrt{1-\alpha^2}\mathbf{P}$ with $0 \leq \alpha \leq 1$ is a scalar that controls the trade-off between peak sharpness and noise tolerance. For $\alpha = 0$, the OTSDF filter is same as the MVSDF filter, and for $\alpha = 1$, the OTSDF filter is same as the MACE filter. For other values of α, we achieve a compromise between the two extremes. In practice, we find that α values close to but not equal to 1 (e.g., 0.999) usually produce the best results.

Over the past two decades, CF designs have advanced in many ways [5]. Some examples of these advances are the following:

- Relaxing the hard constraints in Eq. (17.2) by maximizing the average correlation output value (at the origin) rather than requiring that the correlation output take on specific values
- Designing the CF based on the entire correlation output rather than just the value at the origin
- Applying nonlinear operations to input images in the form of point nonlinearities and in the form of quadratic correlation filters
- Combining the shift-invariance properties of the CFs with the good generalization properties of support vector machine (SVM) classifiers in the form of maximum margin correlation filters (MMCFs)

In the next section, we discuss how CFs can be used for iris segmentation.

17.3   Iris Segmentation

Before two iris patterns can be compared, they need to be segmented from the rest of the image. Most iris segmentation approaches rely on the fact that, in gray-scale images, the pupil (i.e., interior) to the iris is usually darker than the iris region and the sclera (on the outside) is brighter than the iris. So, iris boundaries can be detected by looking for regions with large gradient magnitudes (e.g., from pupil to iris and from iris to sclera) as was proposed originally in using integrodifferential operators for iris boundary detection. Another useful feature of iris boundaries is that they may be nearly circular, suggesting the use of circular Hough transforms to identify iris boundaries. More recently, improved iris segmentation results have been obtained using active contour techniques. In this section, we discuss a cross-correlation-based method for iris segmentation.

As discussed above, one way to locate the iris boundaries is to determine regions of high radial gradients of the circular Hough transform [11]. For eye image $E[m,n]$, the circular Hough transform is defined as

$$Z[m,n,r] = \sum_{\theta \in I_\theta} E[m + r\cos\theta, n + r\sin\theta] \qquad (17.8)$$

where I_θ is a sub-interval of $\theta \in [0°, 360°]$. In practice, we do not integrate across the entire circle because the upper and lower regions are unreliable due to eyelid interference. Instead, we integrate over symmetric left and right lateral regions extending from $45°$ to $150°$ from the top of the circle. This exclusion of the eyelid region leads to a more robust edge detector. Since the iris may not be centered in the image, we need to compute $Z[m,n,r]$ for every possible triplet $[m,n,r]$. This can be computationally prohibitive, as its complexity is order $O(M^4)$ for an $M \times M$ image. As an example, consider the following naive approach to computing the discrete circular Hough transform: for every possible center location $[m,n]$, the polar transform (requiring $O(M^2)$ operations) of the image is obtained using $[m,n]$ as the origin, followed by summing over the angle within the specified angular interval I_θ. This requires M^2 repetitions of the polar transform leading to $O(M^4)$ complexity. This complexity can be reduced by using cross-correlation operation to produce an approximation to the Hough transform, as described below.

For a given triplet $[m_0, n_0, r_0]$, the value of the discrete Hough transform is a sum of pixel intensities that fall along the circular arc of radius r_0 centered at $[m_0, n_0]$. This summation can be approximated as the inner product of the image $E[m,n]$ with a binary template $C_r[m,n]$ that equals 1 along this contour and 0 everywhere else, as shown in Fig. 17.2. The inner product of the eye image with the binary template in Fig. 17.2 yields one value $Z[m_0, n_0, r_0]$.

To determine $Z[m,n,r_0]$ for all $[m,n]$ values, we must compute inner products with all shifted versions of the binary contour, which is equivalent to spatial cross-correlation. So one 2-D cross section of the discrete Hough transform can be obtained as
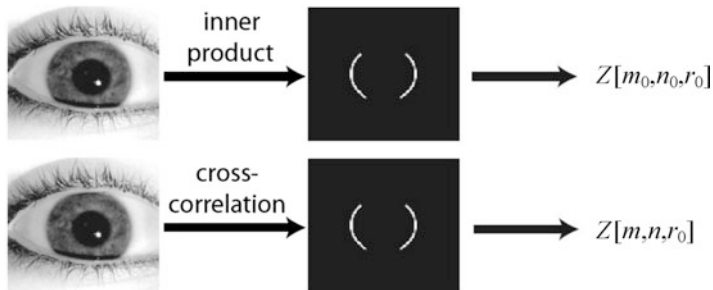
Fig. 17.2 Circular Hough transform approximations using contour filters. An inner product yields a single value, and a cross-correlation yields a 2-D cross section of values
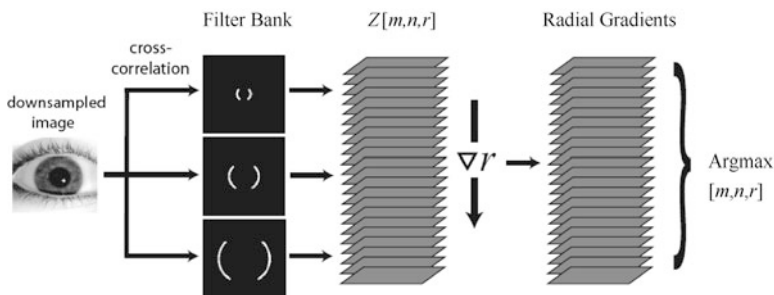
Fig. 17.3 Cross-correlation-based circular boundary detection

$$Z[m, n, r_0] = E[m, n] \otimes C_{r_0}[m, n] \tag{17.9}$$

where cross-correlation is denoted by the $\otimes$ symbol. The cross-correlation in Eq. (17.9) can be efficiently implemented using fast Fourier transform (FFT) algorithm. The computational complexity of the 2-D FFT is $O(M^2 \log_2 M)$. It should be noted that this technique only produces an approximation to the actual Hough transform values, since the use of discrete Fourier transform results in a *circular* cross-correlation. This means that the contour of integration will wrap around when shifted beyond the edge of the image. Fortunately, this does not present a problem in practice. We can either assume that the iris is completely contained within the eye image, or we can zero-pad the eye image if we expect part of the iris to be cut off.

To determine the iris boundaries, the starting point is to build a coarse approximation to the discrete Hough transform. Using a coarse approximation reduces computation time but allows for a reasonable initial estimate of the boundary locations. First, the input eye image is down-sampled to a low-resolution image, e.g., 100 by 100 pixels. Then it is passed through a correlation filter bank where each filter is a binary contour filter of different radius. The output from each filter yields the plane of all Hough transform values at a fixed radius, so concatenating all outputs together gives us the entire discrete Hough space. This is illustrated in Fig. 17.3.

Because the image dimension is low, we need only use 50–100 radial values. The resulting Hough transform is fairly coarse but is usually sufficient for detecting the approximate boundary locations. Each contour filter's frequency response is computed before hand and stored. Also, the eye image is converted to the frequency domain once, at the beginning of the filtering process. As a result, we only have to compute one inverse FFT for each contour filter. The entire computation has complexity $O\left(KL^2\log_2 L\right)$ when using K contour filters applied in the low-dimensional space of size $L \times L$. This is significantly better than the $O\left(M^4\right)$ complexity of straightforward computation of the circular Hough transform.

The approximate Hough transform is multiplied by $1/r$ to normalize for the circumference. Because we are detecting circular edges, we need the radial gradients of the normalized Hough space. A smoothed difference operator is applied along the radial direction to get these radial gradients. After this step, we have a 3-D set of values which indicate the presence of a circular iris boundary at a range of possible positions. First, a search is conducted for the inner iris boundary, which, depending upon the darkness of the pupil, typically produces the highest radial gradient. We locate the maxima of the gradient, with the minor constraint that the boundary cannot be very near the edge of the image since it is the inner boundary. Once established, the location of the first boundary places a prior on the location of the second boundary. This affects the search in two ways: (1) any potential boundary locations which do not completely surround the inner boundary by some minimal margin are ruled out, and (2) the gradient values are weighted by a Gaussian function, such that the Gaussian is centered on the inner boundary center. The second condition allows for slightly non-concentric boundaries but makes the detection of extremely non-concentric boundaries unlikely. Then the maximum of the weighted gradients is associated with the outer boundary.

At this point in the algorithm, we have coarse estimates for the location of both boundaries. In order to fine-tune these estimates, the detection process is repeated at a higher resolution. The higher resolution contour filters are more computationally expensive to apply, but we do not have to apply the entire filter bank. Instead, we only apply the few filters which have radii in the immediate neighborhood of the coarse estimates. This allows us to refine our estimates without adding significant computation. After deriving the final boundary estimates, the iris pattern is "unwrapped" into normalized pseudo-polar coordinates. We note that the objectives of cross-correlation-based segmentation algorithm are consistent with other iris segmentation algorithms that estimate the iris boundaries by finding the regions of high radial gradients. It is the use of correlation filter bank to efficiently obtain the Hough transform approximation, which differs from other implementations.

We tested the iris segmentation algorithm on an iris image database collected at CMU [12]. Some examples of iris images from CMU database are shown in Fig. 17.4. CMU iris image database contains high-resolution ($950 \times 1,419$) iris images acquired under visible wavelength illumination. This database contains 2,390 images from 101 different eyes with 20–25 images per class. Although the

Fig. 17.4 Sample images from CMU database
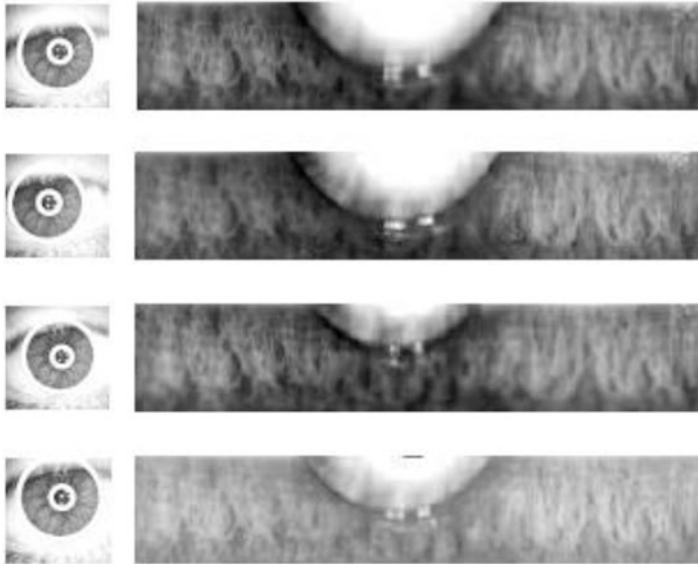
Fig. 17.5 Example cross-correlation-based iris segmentation results (*left*) and resulting unwrapped iris images (*right*)

CMU database images have high resolution, they tend to be more difficult for pattern matching because of greater intra-class variation (especially with regard to focus and occlusion). As can be seen from Fig. 17.4, the upper eyelid can cause partial occlusion in the iris image.

In Fig. 17.5, we show example segmentation results from the cross-correlation-based method as well as the corresponding unwrapped iris images. The white regions in the top-middle portion of the unwrapped images are due to the occlusion from the eyelids. The results of automatic segmentation were compared against manual segmentation, and it was observed that nearly 99% of the images were properly segmented by the cross-correlation-based algorithm. The few segmentation errors observed were mostly a result of heavy eyelid occlusion obscuring the iris boundaries.

This discussion about iris segmentation would not be complete without mentioning the real-world challenges in segmenting the iris images. Iris segmentation

is degraded by impairments such as non-frontal gaze, specular reflections, and
occlusion due to eyelids and eyelashes.

17.4 Iris Matching

Cross-correlation is a powerful tool for quantifying the similarity between two
images. So it can be used for matching iris images as illustrated in Fig. 17.6.
Segmented and unwrapped training iris patterns from one class (i.e., one eye)
are used to determine a correlation filter such as the MACE filter (Eq. 17.4) or
OTSDF filter (Eq. 17.7). When a query iris image is presented, it is also segmented
and unwrapped and then cross-correlated using the designed CF. The resulting
correlation output should contain a sharp peak if the query image is from the same
class as the training images used for designing the correlation filter, and no such
peak if the query image is from an impostor, as depicted in Fig. 17.6.

The sharpness of the correlation peak can be quantified by the peak-to-correlation
energy (PCE) defined as a ratio of square of the peak-to-correlation energy. Since
the PCE is a ratio, multiplying the input query image by any constant will not affect
the PCE, making it invariant to constant illumination changes. If the PCE is above
a prespecified threshold, the input image is classified as authentic and otherwise as
coming from an impostor. By varying the PCE threshold, one can trade off false
accept rate (FAR) for false reject rate (FRR).

We demonstrated [13] that correlation filters can offer excellent iris recognition
performance. We investigated the performance of OTSDF correlation filter on the
CMU iris database. Three images from each iris class were used as reference images
to define the class, and the rest were used for testing. The testing generated similarity
scores for a number of authentic and impostor comparisons. We measured the equal
error rate (EER), the point at which FRR equals FAR. The EER using correlation
filters was 0.61%, whereas the EER using Gabor wavelet-based binary codes was
1.04%.

Since correlation filters are applied to unwrapped iris images, the shift-invariance
of CFs corresponds to shifts in the polar domain, i.e., the CF method can handle
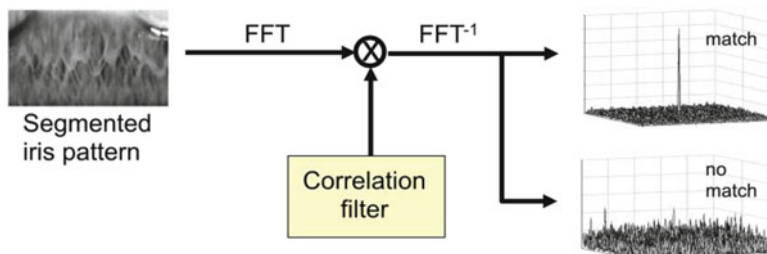

Fig. 17.6 Correlation-based matching of unwrapped iris images

Fig. 17.7 Sample ocular images from the Face and Ocular Challenge Set (FOCS) dataset [14]

in-plane rotations in the original image (before mapping to polar coordinates). However, the challenge is that iris regions have to be segmented from their surrounding regions before obtaining the unwrapped versions. One of the advantages of CFs is that they can be applied to the original eye images without any need for segmentation.

In Fig. 17.7, we show two example ocular images from the Face and Ocular Challenge Set (FOCS) [14] that contain the iris region as well as surrounding regions such as eyebrow, some part of nose bridge, skin near the eye, and part of the forehead region. Using these additional regions can improve the recognition rate. In Fig. 17.8, we show the correlation outputs for an authentic ocular image pair and an impostor ocular image pair. It is clear that correlation is stronger for the authentic pair. We also show the correlation output when the filter is same, but the probe image is an iris-occluded version of the authentic ocular image, and we can see that the correlation peak is still visible even though the iris is occluded.

17.5 Bayesian Graphical Models for Iris Recognition

One of the reasons for the degradation of match score between two iris images (acquired at different times) from the same eye is that the two images can exhibit nonlinear deformations, e.g., different regions of the unwrapped iris images may move differently, as illustrated in Fig. 17.9. Such nonlinear deformation can also be caused by slight differences in the segmentation boundaries produced for the two images. When segmentation boundaries differ, corresponding regions of the two unwrapped images may move by different amounts because the mapping from Cartesian coordinates to polar coordinates very much depends on where the inner and outer boundaries are.

Another challenge in matching two iris images from the same eye is that one may exhibit more occlusions caused by eyelid and the other may exhibit less occlusion, affecting the match score. If such occluded regions can be excluded or weighted less in determining a match score, then that should lead to a more robust match
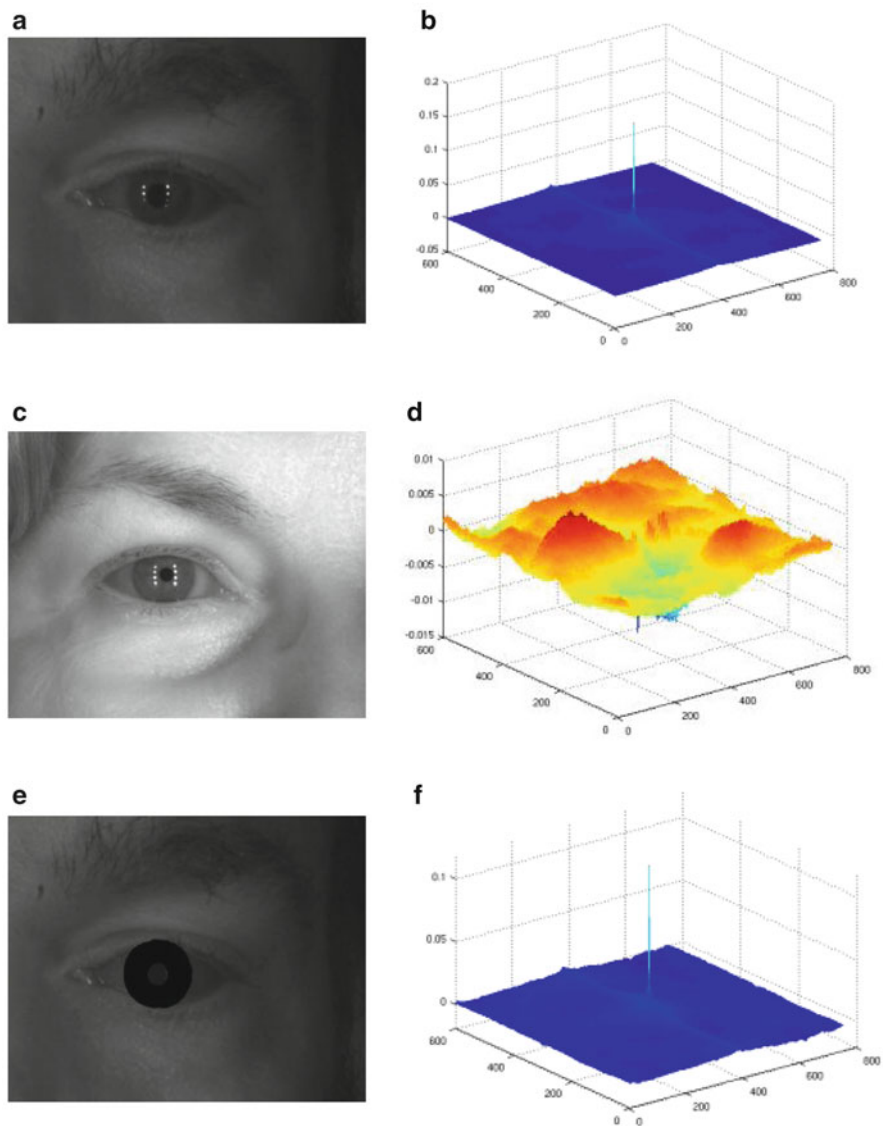
Fig. 17.8 (**a**) Authentic ocular image, (**b**) resulting correlation output, (**c**) an impostor ocular image, (**d**) resulting correlation output, (**e**) iris-occluded authentic ocular image, and (**f**) resulting correlation output

score. Toward this goal, Thornton et al. [15] proposed the use of Bayesian graphical models for improved iris matching.

The main idea of Bayesian graphical models can be summarized as follows. The two iris images being compared are divided into nonoverlapping patches as shown in
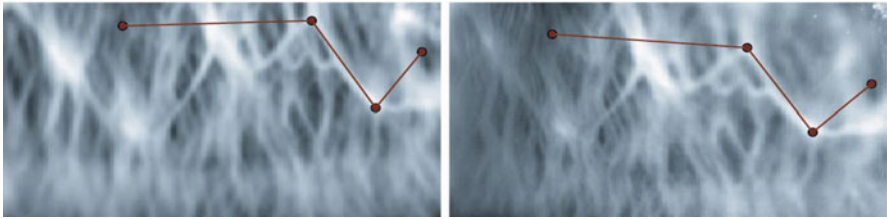
Fig. 17.9 Close-ups of segmented patterns from same eye (landmark points illustrate relative deformation)
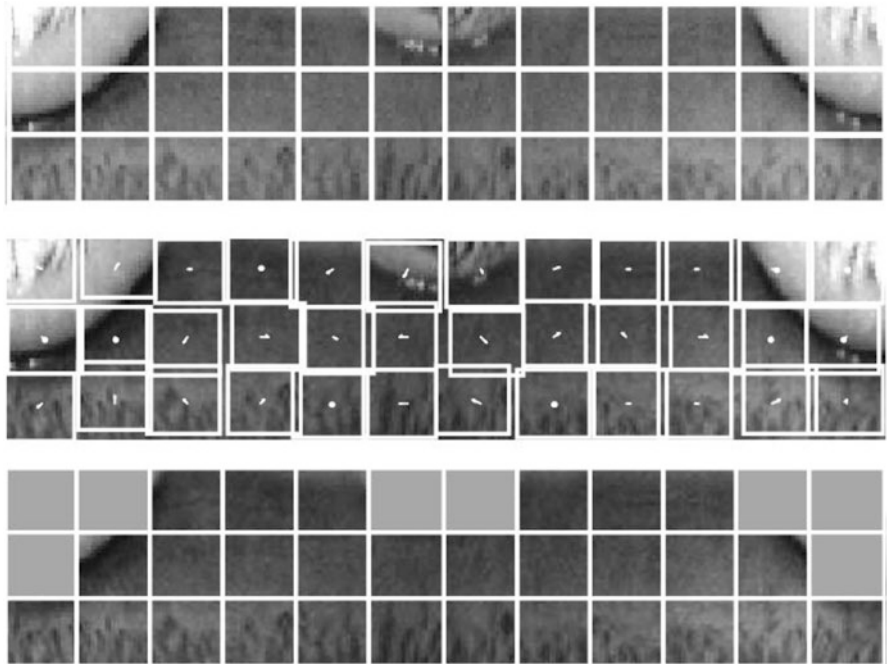
Fig. 17.10 Hidden states of the model: Sample iris plane partition (*top*), deformation vector field (*center*), and binary occlusion field (*bottom*)

Fig. 17.10 (top) where the unwrapped iris image is divided into 36 patches. Different patches from the probe image may be shifted by different amounts compared to the corresponding patches of the template, as shown in Fig. 17.10 (middle) by white arrows. The length of the arrow indicates the magnitude of the shift, and the direction of the arrow indicates the direction of the shift. Also, some of the probe image patches may be occluded, as shown by the gray squares in Fig. 17.10 (bottom). To estimate these patch shifts and to estimate whether a patch is occluded or not, the corresponding patches from the two regions are cross-correlated. If both patches are unoccluded and from the same eye, the resulting correlation peak should be large,
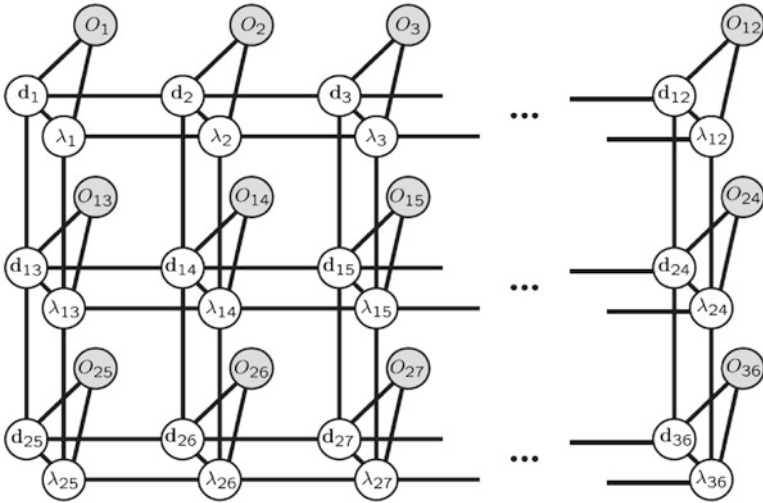
Fig. 17.11 Graphical model structure, for a 3 by 12 iris plane partition. *Shaded nodes* represent observed variables and *non-shaded nodes* represent hidden variables

and the location of the correlation peak should indicate the relative shift between the corresponding patches from the two images. Thus, the cross-correlations between the patches from the two images provide clues about patch shifts between the two images and occlusions.

The graphical models corresponding to the patch structure in Fig. 17.10 are shown in Fig. 17.11. Shaded nodes $O_1, O_2, \ldots, O_{36}$ denote the observations (e.g., PCE values and peak locations from patches) and represent evidence about both the similarity between the template and query iris patterns and the presence of eyelids across the iris plane, whereas hidden variables $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{36}$ indicate the shift of probe image patches from corresponding template image patches and binary-valued hidden variables $\lambda_1, \lambda_2, \ldots, \lambda_{36}$ indicate whether a patch is occluded (i.e., $\lambda = 1$) or not occluded (i.e., $\lambda = 0$).

Once the hidden variables (indicated by set H) and the observed variables (indicated by set O) are identified, the objective is to learn a joint probability distribution over these variables so that we may perform inference on the hidden states. In order to do this, we make some assumptions about the dependencies between variables. If we assume a fully connected model (i.e., every variable is directly dependent upon every other variable), learning and inference would be completely intractable. Therefore, we simplify the model by assuming direct dependence only between variables which have an intuitive or empirical statistical connection. This is a common practice in the field of probabilistic graphical models [16], which provides a general framework for working with complicated joint distributions. In keeping with conventional graphical model notation, a variable is represented visually as a node, and a direct dependence between two variables is represented as an edge connecting two nodes.
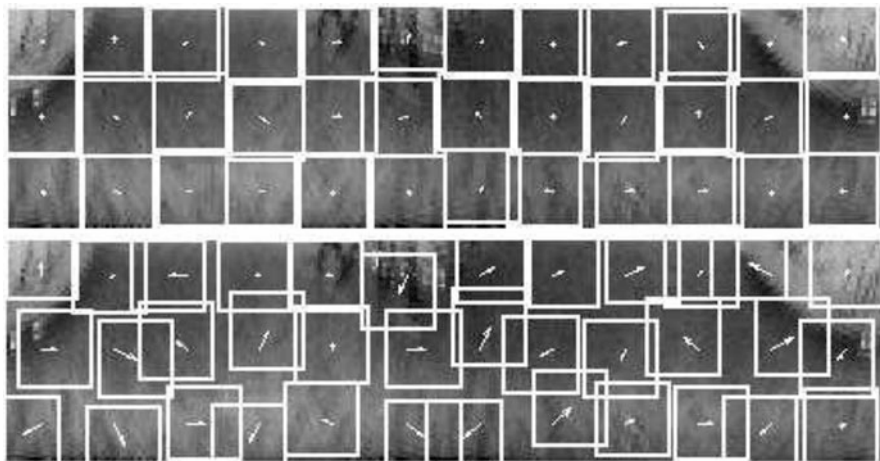
Fig. 17.12 Contrasting deformation field examples. *Top*: A "reasonable" field with higher likelihood in our model. *Bottom*: An "unreasonable" field with lower likelihood

We start by considering the relationship between the hidden deformation variables. Figure 17.12 shows two examples of deformation vector fields which might align a probe image to a template image. The alignment at the top is more "reasonable" than the alignment at the bottom, by which we mean that the top vector field is more likely to approximate the effect of real iris image movement. Part of what makes the second vector field more unlikely, besides the increased magnitude of the vectors, is the fact that many vectors from adjacent regions have conflicting directions. We would expect adjacent iris regions to exhibit similar motion, and we want a model capable of learning this tendency from iris data. Therefore, we allow for direct dependence between each vector and its spatial neighbors in the vector field. We formulate this dependence scheme as a Markov random field (MRF), which is an undirected graph structure on the variable set [17]. We arrange our variables in an MRF framework because it makes learning and inference more tractable. Specifically, we form a 2-D lattice MRF on the deformation variables, in which each node is connected to its neighbors on the lattice (illustrated in Fig. 17.11).

The graphical model parameters are learned using training images via expectation-maximization (EM) algorithm [18]. Use of the graphical model for an iris pattern comparison is a two-step process: (1) we infer distributions on the hidden variables of the model, and (2) we use this information to compute a match score between template and query. For space reasons, we do not discuss the details, but they can be found elsewhere [12].

The Bayesian graphical model approach was tested using CMU iris database containing high-resolution ($950 \times 1,419$) iris images acquired under visible wavelength illumination. This database contains 2,390 images from 101 different eyes with 20–25 images per class. Table 17.1 shows the false alarm rates (FAR) for

Table 17.1 False accept rate (*FAR*) for different false reject rates (*FRRS*)

Algorithm	FRR = 1%	FRR = 0.1%	FRR = 0.01%
Baseline	0.23%	1.16%	1.85%
BGM	0	0	0.46%

three different false reject rate (FRR) levels for the Bayesian graphical model (BGM) ap proach and the baseline algorithm. The baseline algorithm is the standard approach of iris segmentation followed by a binary code derived from the phases of Gabor wavelet inner products. As can be seen from Table 17.1, the Bayesian graphical model-based matching algorithm offers improved recognition performance.

17.6   Summary

Correlation filters (CFs) have long been researched for automatic target recognition applications where the targets can appear in different orientations, scales, locations, and with occlusions and obscurations. Advanced CF designs have been developed to deal with such image impairments. Also, CFs have built-in advantages such as shift-invariance and graceful degradation. One of the challenges in real-world iris recognition is that the iris images exhibit impairments such as non-frontal gaze, occlusions due to eyelids and eyelashes, specular reflections, and nonuniform illuminations. Correlation filters can be beneficial in dealing with such impairments. In this chapter, we discussed how correlation filters (CFs) can play an important role in iris recognition.

We discussed how CFs can provide an alternative method for iris segmentation and how they can be used for matching both unwrapped iris images (in polar coordinates) and iris images in the original Cartesian coordinates. The use of CFs can be extended to iris image patches so that the patch cross-correlations can provide information to a Bayesian graphical model (BGM) that outputs match scores that are adjusted for impairments such as nonlinear deformations and occlusions. We showed that BGM offers improved iris recognition performance compared to a baseline binary code-based matching.

References

1. Wildes, R.: Chapter 3: Iris recognition. In: Wayman, J., Jain, A., Maltoni, D., Maio, D. (eds.) Biometric Systems, pp. 63–95. Springer, Berlin (2005)
2. Daugman, J.: Probing the uniqueness and randomness of IrisCodes: results from 200 billion iris pair comparisons. Proc. IEEE **94**(11), 1927–1935 (2006)
3. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1148–1161 (1993)

4. Vijaya Kumar, B.V.K., Xie, C., Thornton, J.: Iris verification using correlation filters. In: Proceedings of 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA). LCNS 2688, pp. 697–705. Springer, Berlin/Heidelberg (2003)
5. Vijaya Kumar, B.V.K., Mahalanobis, A., Juday, R.: Correlation Pattern Recognition. Cambridge University Press, Cambridge (2005)
6. Mahalanobis, A., Ortiz, L., Vijaya Kumar, B.V.K.: Performance of the MACH/DCCF algorithms on the 10-class public release MSTAR data set. Proc. SPIE **3721**, 285–291 (1999)
7. Vijaya Kumar, B.V.K., Savvides, M., Xie, C.: Correlation pattern recognition for face recognition. Proc. IEEE **94**(11), 1963–1976 (2006)
8. Refregier, P.: Filter design for optical pattern recognition: multicriteria optimization approach. Opt. Lett. **15**(15), 854–856 (1990)
9. Mahalanobis, A., Vijaya Kumar, B.V.K., Casasent, D.: Minimum average correlation energy filters. Appl. Opt. **26**, 3630–3633 (1987)
10. Vijaya Kumar, B.V.K.: Minimum variance synthetic discriminant functions. JOSA-A **3**, 1579–1584 (1986)
11. Duda, R.D., Hart, P.E.: Use of the Hough transform to detect lines and curves in pictures. Commun. ACM **15**, 11–15 (1972)
12. Thornton, J.: Iris pattern matching: a probabilistic model based on discriminative cues. Ph.D. dissertation, Carnegie Mellon University (2007)
13. Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: Robust Iris recognition using advanced correlation techniques. In: Proceedings of International Conference on Image Analysis and Recognition (ICIAR), Image Analysis and Recognition, Lecture Notes in Computer Science #3656, pp. 1098–1105. Springer Berlin/Heidelberg (2005)
14. http://www.nist.gov/itl/iad/ig/focs.cfm, Face and ocular challenge series (FOCS).
15. Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: A unified Bayesian approach to deformed pattern matching of iris images. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 596–606 (2007)
16. Frey, B.J.: Graphical Models for Machine Learning and Digital Communication. MIT Press, Cambridge, MA (1998)
17. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Ana. Mach. Intell. **6**(6), 721–741 (1984)
18. Jordan, M.I.: Learning in Graphical Models. MIT Press, Cambridge, MA (1999)