Detection of Contact-Lens-Based Iris Biometric Spoofs Using Stereo Imaging

Ken Hughes
School of Engineering and Computer Science, University of the Pacific
khughes@pacific.edu

Kevin W. Bowyer
Department of Computer Science and Engineering, University of Notre Dame
kwb@cse.nd.edu

Abstract

Cosmetic contact lenses can be used to spoof an iris biometric system, either to evade being matched to a watch list or in principle even to masquerade as a selected other person. Existing approaches to detecting whether or not a person is wearing cosmetic contact lenses either are limited to detecting lenses created by a particular manufacturing technology, assume knowledge of the particular pattern printed in/on the lens, or require a sequence of images. We present proof-of-concept results for a method of detecting cosmetic contact lenses that is general, in the sense that it assumes nothing about the manufacturing technique or texture pattern of the lens, and that requires only a “snapshot” instance of imaging. The “snapshot” is a stereo pair of images, from which the shape of the surface of the iris texture region is estimated. In the absence of contacts or the presence of clear contacts, the iris region presents a coarse planar surface. In the presence of cosmetic contacts, the iris region presents a convex surface. Thus the problem of determining if a person is wearing a cosmetic contact lens is transformed into the problem of classifying the estimated surface shape for the iris region. This is the first approach to analyze iris biometric images in the context of 3D shape.

1. Introduction

The National Science and Technology Council’s report titled “The National Biometrics Challenge” outlines Enterprise and E-Government Services as one of four primary driving forces for biometric technology [25]. Industry reports project iris biometric technology to be the fastest-growing of the major biometric

Figure 1 – Images from LG 2200 iris sensor of the same eye without (above) and with (below) cosmetic contact lens (subject 04780 in ND dataset).
modalities over the period 2009-2014; see Table 1 of [8].

Iris biometric technology was pioneered by John Daugman [1,2]. The basic idea underlying iris biometrics is that the texture of the iris can be used as the basis for generating a code that is unique to the particular iris and that is robust to a range of imaging conditions. Iris biometrics has become an increasingly active research area in recent years [3,4]. Major deployments such as the use in border control in the United Arab Emirates [5,6] and as part of India’s Unique ID program [7] have attracted a great deal of attention.

The wearing of various types of contact lenses presents challenges for iris biometric systems [9]. The term “cosmetic contact lenses” is used to refer to contact lenses that change the appearance of the eye in some way; e.g., allowing a person with dark-colored irises to appear to have light-colored irises. Attempts to match an image of an eye not wearing a cosmetic contact lens with an image of the same eye wearing a cosmetic contact lens result in a very high false non-match rate [9]. This could be exploited by a person who is on a watchlist to evade detection simply by wearing cosmetic contacts. It is also conceivable that an attacker wanting to impersonate a particular person could have cosmetic contact lenses made to give them the target person’s iris texture. Either of these scenarios might be referred to informally as “spoofing” the iris biometric system.

The problem of automatically detecting if a person is wearing cosmetic contact lenses has attracted a number of research efforts. Previous efforts are generally either too specific to the particular type of cosmetic contact lens, and / or require some interaction between the user and the biometric system. We present proof-of-concept results of the first approach to exploit image analysis of the 3D shape of the iris. Our approach is based on capturing a stereo pair of iris images, and estimating the 3D shape of the imaged iris region. In the case of an eye that is not wearing contacts or that is wearing clear contacts, the iris region has a coarse, flat shape. In the case of an eye that is wearing cosmetic contacts, the iris region has a smooth, convex shape.

2. Literature review

Baker et al [9] present results of experiments that investigate the effects of contact lenses on the accuracy of iris biometrics. They conclude that wearing contact lenses has the general effect of degrading recognition accuracy by broadening the authentic distribution and increasing the false non-match rate. They find that cosmetic contact lenses – those with some type of texture pattern embedded in the lens – result in essentially a 100% false non-match rate.

Cosmetic contact lenses cause highly variably changes in the observed texture of the iris region. One issue is that cosmetic contact lenses generally are not 100% opaque, but instead allow some amount of the natural iris texture to be visible. Thus the observed texture is dependent on both the natural iris texture and on the texture pattern embedded in the lens. Another issue is that contacts, both cosmetic and regular, naturally move around to some degree on the surface of the eye. Thus the observed iris texture, which is a blend of the natural texture and the lens texture, changes as the lens moves around on the surface of the eye. The result is that not only does the eye with the cosmetic lens not match the eye without the lens, images of the eye with the lens from two different times will generally not match each other.

Daugman presented an early approach to the detection of cosmetic contact lenses [10]. His approach is intended to work for lenses that are manufactured using a dot-matrix type technology. That is, the texture in the lens is implemented through dots on a regularly-spaced grid. However, this is not the only method used to manufacture cosmetic lenses, and even for dot-matrix type lenses this approach can fail if the printing resolution is much greater than the resolution of the camera in the iris sensor.

Bodade and Talbar [19] discuss approaches to detecting various types of iris spoofs, including contact lenses, images printed on paper, and video playback of an eye movie. They discuss three approaches to spoof detection: the frequency spectrum approach (Daugman’s approach, discussed above), the reflectance approach and the pupil dynamics approach. The reflectance approach involves illuminating the eye with multiple different wavelengths of light and comparing the relative response in the sclera and iris regions. The pupil dynamics approach involves acquiring several images while manipulating the illumination level, so as to check for a change in pupil dilation. They conclude that a combination of the approaches is needed. They present some results for detecting printed iris images, but do not analyze any images of cosmetic contact lenses.

He et al [14] propose an approach to iris spoof detection that is aimed in particular at cosmetic contact lenses. They learn a texture model and use it to classify images as real or spoof. Their model describes the texture in six regions of the iris: an inner, middle and outer band on each of the left and the right quadrant of the iris (ignoring the upper and lower quadrant). They compute multi-scale local binary pattern texture features and use Adaboost to learn the important features for distinguishing between classes.
Their experimental dataset includes “600 counterfeit iris images … the majority of them are printed color contact lens iris images” representing “20 kinds of different contact lens”, and a large number of real iris images selected from the CASIA and ICE datasets. From these, “300 counterfeit iris images and 6000 live iris images are randomly selected from the above data set for Adaboost training”. Thus it appears that the training data will on average contain about 15 sample images for each of the 20 types of contact lens. From this experiment structure, they report 0.67% false accept rate (presumably meaning accept a contact lens image as not a contact lens image) and 2.64% false reject rate (presumably meaning to reject an iris image as being a contact lens image).

Park [11] proposes an approach to iris spoof detection that involves a new camera design “with dual IR-LED and visible light illuminators”. The system can selectively turn on either of the 760 nm or the 880 nm near-IR LED illuminators. It is also capable of turning on a visible light, meant to prompt pupil contraction. The detection of cosmetic contact lenses is based on analysis of the iris images taken before and after the visible light is turned on. The inner and outer boundary of the iris is detected in each image, and “if the variation of (dilation) ratio does not exceed a predetermined threshold, we regard the input iris image as fake”. As a further check, the texture of the bands of the iris next to the pupil is compared in the two images. If the difference in texture between the two images exceeds a threshold, this is taken as indicating that the images represent a spoof attempt, with the contact lens texture seen bordering the dilated pupil and the real iris texture seen bordering the contracted pupil. The experimental evaluation includes two samples of patterns contact lenses.

The frequency analysis approach to detecting whether or not a person is wearing cosmetic contact lenses, of which Daugman’s work [10] is an example, can potentially detect lenses created with a dot-matrix type printing technique. However, it may have problems if the iris image is defocused or if the printing resolution in the lens is much finer than that in the iris image, and it cannot detects lenses created by other methods. The texture analysis approach of He et al [14] appears to require samples of the contact lens patterns to be detected as part of its training data. The approach proposed by Park [11] involves analyzing images taken before and after the system changes the ambient illumination level. Besides being inherently more complex and time-consuming, this approach may present difficulties for older users, as the response to a given change in illumination diminishes with age [20]. Bodade and Talbar [19] propose a combination of approaches, which will encounter the shortcomings of each of the component approaches.

The proposal that we investigate in this paper requires a second camera in the iris sensor, to allow stereo imaging, but does not require different illumination and both images would be acquired simultaneously. The approach does not require knowledge of, or samples of, the texture in the contact lens, as it is based simply on the fact that the contact lens has a roughly spherical shape, so as to fit the eye.  

3. Background – eye and contact lenses

The outermost (anterior) layer of the eye is the cornea. The cornea is generally clear, and so contributes relatively little to the content of an image of the iris. The next layer behind the cornea is the aqueous humor. This layer is also generally clear, and so also contributes relatively little to the content of an iris image. The next layer is the iris (and pupil). Thus, the texture of the iris is imaged through the cornea and aqueous humor.

![Figure 2 – Cross-section diagram of the eye. The iris is approximately planar, whereas a contact lens rests on the cornea and sclera.](image)

The overall shape of the eyeball is roughly spherical. However, the outer surface of the iris is approximately a coarse planar surface. To a rough first approximation, the texture seen in an iris image can be considered to result from a coarse but roughly planar anterior surface of the iris.

A typical soft contact lens can be considered to rest on the surface of the cornea / sclera. The outermost band of the lens generally is in better direct contact with the cornea, whereas the central part of the lens generally has a shape that, for prescription lenses, is intentionally designed to provide the needed vision correction. To a first approximation, the surface of the contact lens can be considered to be spherical. Also, the contact lens is intentionally designed to be very
smooth, so as to not result in discomfort for either the sclera/cornea or the inner surface of the eyelid.

Consider that the region of an iris image that is segmented as representing “iris texture” may result from either the true iris surface or from a cosmetic contact lens. In the case that the iris texture region results from the true iris surface, it results from a rough but approximately planar surface. In the case that the iris texture region results from a cosmetic contact lens, it results from a smooth and approximately spherical surface. This suggests that it may be useful to consider the problem of distinguishing between “live” and “spoofed” iris texture as a problem of sensing 3D shape and categorizing into two basic categories—“coarse planar” versus “smooth spherical”.

4. Experimental materials and setup

At a conceptual level, our system operates as follows. First, images of the eye are obtained. The acquisition is designed so that the images enable us to estimate the 3D shape of the iris texture region. In our work, this is a stereo pair of images. Next, the 3D shape of the iris texture region is estimated. In our work, we find correspondences between the texture in the iris regions in the two images, and estimate a depth map based on the stereo correspondences. Finally, the estimated shape of the iris region is analyzed to determine if it arises from the natural iris surface or from texture in a contact lens.

4.1 Stereo camera rig

Our current proof-of-concept system uses two identical Nikon D90 cameras with Nikon 55-200mm f/4.5-5.6G ED AF-S DX zoom lenses. (The camera rig used in this proof-of-concept demonstration is larger than would be used in a production system. It represents more the equipment already available in our lab or available relatively inexpensively.) These cameras acquire images up to a resolution of 4288 x 2848. However, we determined that this resolution is not required, and we do not actually use the full-resolution images. The images are scaled to 1072 x 712 using Imagemagic's [16] convert function before processing by the stereo routines. The stereo reconstruction routines run more slowly for image sizes larger than this, but without giving significantly better 3D results.

The cameras are set to a manual exposure; 1/30 second shutter speed and F8 aperture. A smaller aperture gives a greater depth of field, which makes focusing easier and results in less blurring of the various depths. ISO is set to 200. The lenses are set to approximately 135mm using the marks on each lens to judge the approximate setting. Each camera lens is mounted on a 48mm extension tube, which allows additional enlargement of the images so that focusing can be accomplished more easily.

The cameras are mounted with approximately 5” to 6” of separation. (See Figure 3.) Using Manfrotto 3047 heads on a rig, the heads are arranged close together by mounting the cameras at a 90 degree angle and switching the tightening knobs on the right head so that the shortest knob sticks out toward the left head. The cameras are mounted to be as level relative to each other as possible. The goal is for the central axes of both lenses to be as close to as possible to planar. The ideal stereo configuration would have only a rotation about the Y (vertical) axis.

The minimum distance between the target and the camera mount screw (approximately where the sensor is located) using the 135mm lenses with the 48mm tubes is approximately 22”. For the calibration step, a calibration target is positioned at approximately 22” to 23” from each camera. The calibration process uses a very small chessboard pattern (0.7” x 1.0”, with 0.1” squares) to determine the relative positions of the two cameras in the stereo setup. Due to the small size and high magnification needed for the calibration, the pattern is printed on glossy photocopier paper, at 1200 dpi or higher.

Figure 3 – Camera Calibration Setup.

The subject should be positioned so that the iris is approximately in the center of each of the two images, since optical distortion caused by the lenses is more likely at the edges of the image. Exact centering in both images is not mandatory, as compensation for horizontal displacement can be handled in the post processing. However, the ideal subject positioning to
minimize post-processing is for the irises to be located vertically near the center of the image, and for the iris in the right image to be farther left horizontally than in the left image. The lighting for the subject should minimize shadows on the iris and avoid any reflections on the iris itself, if possible. If using auto-focus, align the focus region-of-interest for each camera to be on the iris. If the auto-focus area is set, for example, on the eyelashes, this can place the iris outside the depth of field. In our acquisitions, each camera was set to take more than one image when the shutter release was pressed. The two cameras were triggered simultaneously and 4 or 5 images taken with each camera. Then one well-focused image was selected for each camera.

4.2 Stereo processing flow

Our processing flow for the stereo images is accomplished in four steps:
1. segmentation of the iris region in each image
2. estimation of 3D points from stereo correspondence
3. coarse segmentation and outlier removal
4. cluster extraction and classification

These steps are described in further detail in the next sections. Our proof-of-concept implementation performs some of these steps automatically and some with user interaction. Work is ongoing to make all steps automatic, and to reduce the processing to the minimum actually required to reliably estimate the 3D shape of the iris region.

4.3 Iris segmentation

Iris segmentation locates the region within both images that contains the iris. Depending on the magnification of the camera and lenses, only a small portion of the image is of interest for 3D reconstruction. A number of methods currently exist for segmenting the iris [21,22,23]. The x and y location of the iris center and the radius are needed, at a minimum. Our implementation currently uses hand segmentation by the user. The location of the center in each image is used by the stereo correspondence to determine the minimum disparity between the images, and how much vertical adjustment to the images may be necessary.

4.4 Point estimation

The next step is to extract depth estimates from a pair of iris images. We utilized the routine from the OpenCV libraries, cv.FindStereoCorrespondenceBM [24]. The images are converted to grayscale and, after loading the calibration matrices for the stereo rig, they are passed to the correspondence routine which constructs a disparity map. The disparity map is then projected into 3D to produce a point cloud of range estimates.

The results from the automatic iris segmentation in the previous step are used to determine an appropriate value for the minimum disparity parameter used by the stereo correspondence routine. This parameter gives the starting offset for the correspondence search within an epipolar line of the rectified images. The results are also used to segment the point cloud so that only the points near the iris are considered for further processing.

One of the more important parameters for the stereo correspondence is the sum of absolute differences (SAD) window size. Smaller window sizes allow greater variation in the range estimates but are susceptible to noise. Larger window sizes result in less variation but are less susceptible to noise. Our experiments used values ranging from 9 to 35; a value of 15 gave acceptable results in most cases.

Our implementation allows interactive manipulation of the results by adjusting the minimum disparity, and by shifting the offset of the right image of the stereo pair vertically. We observed that even after calibration and rectification, frequently only a small range of rows in the left and right images were epipolar aligned. Shifting one image up or down by one line would result in a new range of rows which aligned. The user can choose one or more offsets that give good range estimate results for use in the next step. The user can also interactively perform a hand segmentation of the iris if necessary.

Our implementation also performs some additional operations. If multiple range estimates are found for the same x,y image coordinate, they are averaged. This allows fusion of depth information from multiple image pairs if the data is available. The centroid for the point cloud is also calculated and all points are then translated to center the data at the origin. Finally, all the points are saved in Point Cloud Library (PCL) [17] format. We save both the 3D coordinates (x,y,z) and the color of the associated pixel from the left image (r,g,b). Stereo analysis is performed on the grayscale image; the color information is currently used only for visualization.

4.5 Outlier removal and coarse segmentation

The next step, which is optional, is filtering to remove outlier noise in the point cloud. Even after iris segmentation, there are normally many small clusters of points that do not lie near the surface of the iris and
need to be removed. This is done using PCL classes `StatisticalOutlierRemoval` and `RadiusOutlierRemoval`. Statistical outlier removal calculates the mean distance from each point to all its neighbors. The distribution is assumed to be Gaussian with a mean and standard deviation. Outliers are all points whose mean distance is beyond a specified standard deviation (in this program, one SD). Radius outlier removal calculates the number of neighbors within a given radius for each point. Outliers are all points that have fewer than a given number of neighbors.

An initial coarse segmentation is done in order to identify the elements of the point could that correspond generally to the iris and sclera surface. This step first uses the class `pcl::SACSegmentation` to find the largest cluster of roughly planar points. A RANSAC segmentation looks for the best plane fit for the point cloud, using points within a user-specified threshold. The default threshold distance between the plane and inliers is 0.025. The inliers resulting from this step are clustered using `pcl::EuclideanClusterExtraction`. The default threshold distance for the clusters is 0.02. The default minimum cluster size is 250. The distance thresholds are related to the metric size of the eye and so are independent of the resolution of the original images. The minimum cluster size would potentially change based on the resolution of the original images.

4.5 Cluster extraction and classification

The result of the coarse segmentation is finally processed at a finer level to make a classification. Our implementation performs this analysis using the class `pcl::SACSegmentation` to find the largest cluster of planar points and also the largest cluster of spherical points. These two results are examined to determine whether they better match a model for an actual iris or for a cosmetic contact lens.

The threshold distance between the planar fit and outliers, and also the threshold distance between the spherical fit and outliers, are both set to 0.01. (Note that this threshold is substantially tighter than what is used in the initial coarse segmentation.) The default number of iterations for the RANSAC algorithm is 1000.

The point cloud found as the spherical cluster is analyzed by examining the coefficients of the spherical equation. The radius of the sphere should be smaller than that of a typical eyeball (e.g., less than 0.5”). The center's z coordinate should be negative, and the x and y coordinates should be “close” to 0.0. The closest points on the sphere should have positive z coordinates. Note that since the center of the contact lens is not textured and will not be detected, checking for the maximum z values will not work. However, comparing the radius to the center's z coordinate, the radius should be greater than the center's z coordinate.

The point cloud found as the planar cluster is analyzed to look for a wide “disc” of points. The centroid of the cloud is determined, and then the distance of each point from the centroid is calculated. Each point is then weighted by the inverse of its distance from the centroid. This is because the distribution of points increases with the square of the distance. The distribution of points is modeled roughly
as Gaussian and the mean and standard deviation is computed. The ratio of the mean versus standard deviation is computed and compared to a threshold (default value is 4.0) to determine if the cloud is planar and so represents a true iris texture. This ratio gives a measurement of the extent of the “disc” of points. An iris would be a wide disc with small central hole (for the pupil), so the ratio would be small. A cosmetic contact, with the plane slicing through it, would be a thin disc with a larger central hole, so the ratio would be larger.

Additional tests could be formulated to increase reliability of classification. The radius to the outer edge of the disc should be larger for a true iris region than for a textured contact lens. The coefficients of the plane equation can also be analyzed. The angle between the normal and the Z axis should be “small”, and the distance from the origin to the plane should be “small” (since the point cloud would have been centered about its centroid in an earlier step).

5. Experimental results

We acquired images and processed results for four subjects. All four were photographed without any contact lenses and two of the four were photographed once with clear contact lenses and once with cosmetic contacts. All subjects had light-colored irises since the cameras were photographing in normal visible spectrum (the cameras were not modified to take near-IR images). The images shown in Figures 5, 6 and 7 all show the same subject: Figure 5 shows the 3D data for the subject without contact lenses, Figure 6 with a clear contact lens, and Figure 7 with a cosmetic contact lens.

The iris regions for all images were segmented by hand in the stereo correspondence step, covering the entire iris and a small amount of the sclera. The resulting point cloud was then processed using the process flow described in the previous section. We applied the classification after each step in the process flow to observe its performance after each step.

Our tests were performed on the three sets of data; (a) eight image pairs with no contact lenses, (b) two image pairs with prescription contact lenses, and (c) four image pairs with cosmetic contact lenses. Table 1 shows the calculated radius of the sphere for all tests. The spherical classifier correctly rejected the no contacts data and the clear contacts data as not being from cosmetic lenses. The texture that appears in the image when a person is wearing a clear contact lens represents the surface of the iris tissue. Thus the 3D surface found from the stereo processing is the iris surface rather than the surface of the clear contact lens.
Figure 6. Images For Clear Contact Lens; Left Camera Image (top), Output from Stereo, Output from Spherical Clustering (offset and side views).

Figure 7. Images For Cosmetic Lens; Left Camera Image (top), Output from Stereo, and Output from Spherical Clustering (offset and side views).
The results of the spherical classification for images with cosmetic contacts were not always consistent. Classification applied to the output from the stereo correspondence and coarse segmentation gave correct classifications for all four images. Classification applied to the cluster extraction was incorrect for two of the four images, with radii of 1.17” and 1.50”. Analysis of the point cloud for the two image pairs which were incorrectly classified showed that the data representing the cosmetic contact lens was sparse, which resulted in fragmentation when coarse segmentation was applied and thus too few points to accurately estimate the sphere. One of the images is shown in Figure 8.

### Table 1. Radii and classification results for the three data types. Gray shaded cells represent classification as cosmetic contact lenses. The two black cells were incorrectly classified.

<table>
<thead>
<tr>
<th>Contact type</th>
<th>None</th>
<th>Clear</th>
<th>Cosmetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>1 2 3 4 5 6 7 8</td>
<td>9 10</td>
<td>11 12 13 14</td>
</tr>
<tr>
<td>Original</td>
<td>1.15 1.29 0.99 1.29 1.22 1.76 1.06 1.19</td>
<td>1.54 1.37</td>
<td>0.31 0.32 0.32 0.37</td>
</tr>
<tr>
<td>-- Cluster ext.</td>
<td>1.27 1.24 1.30 1.24 1.70 2.13 1.25 88.41</td>
<td>1.61 0.89</td>
<td>0.31 0.34 0.32 0.41</td>
</tr>
<tr>
<td>Outliers/Segm</td>
<td>0.66 1.58 0.79 1.58 1.49 1.85 1.31 1.10</td>
<td>1.55 0.80</td>
<td>0.31 0.31 0.35 0.37</td>
</tr>
<tr>
<td>-- Cluster ext.</td>
<td>1.22 1.57 0.93 1.57 1.52 1.57 1.27 1.50</td>
<td>1.97 1.11</td>
<td>1.50 1.17 0.32 0.40</td>
</tr>
</tbody>
</table>

We have also shown that the 3D surface estimated in the case of an iris wearing a cosmetic contact lens is approximately spherical, and that this difference in 3D shape can be used to distinguish whether or not a cosmetic contact lens is present.

Our work has the potential to enable fast, accurate, automatic detection of whether or not a person using an iris recognition system is wearing a cosmetic contact lens. This would allow detection of attempted spoofs in which someone attempts to avoid being matched with their previously enrolled identity.

As is apparent from the sample images in Figure 1, it is generally possible to tell by manual visual inspection if cosmetic contacts are present. This is especially true if the lens is not aligned with the center of the iris, so that slivers of all-lens and all-iris texture are visible. However, for some types of cosmetic lenses, in images where they are well-aligned with the pupil center, it can be more difficult to detect them by manual inspection. We are interested in developing automated detection that would be faster and potentially more accurate than manual inspection.
Based on our results so far, it appears that the categorization tests can be applied to noisy data taken directly after the first two steps in our process (segmentation and stereo correspondence), which would significantly simplify the overall procedure. We plan to conduct further testing with larger data sets to determine to what extent this holds true. We also note from our results that the spherical cluster tests perform better than the planar cluster tests.

The range estimates from the stereo correspondence were very sensitive to reflections on the iris. During our experiments, it was noticed that "holes" in the point clouds were consistently located in the same regions of the iris. Closer examination of the original images discovered reflections of the nose and brow ridge in these images. In some images, even reflections of the eyebrows, eyelashes and cheeks were visible. Subsequent images taken using a dark cloth to mask areas of the face surrounding the eye resulted in a significant reduction of such holes. This points to the importance of designing the illumination in a production system to minimize specular inter-reflections.

7. References