Assessment of iris recognition reliability for eyes affected by ocular pathologies

Mateusz Trokielewicz†‡, Adam Czajka†‡, Piotr Maciejewicz*†
†Biometrics Laboratory
Research and Academic Computer Network
Wawozowa 18, 02-796 Warsaw, Poland
‡Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warsaw, Poland
mateusz.trokielewicz,adam.czajka@nask.pl

Piotr Maciejewicz*
*Department of Ophthalmology
Medical University of Warsaw
Lindleya 4, 02-005 Warsaw, Poland
piotr.maciejewicz@wum.edu.pl

Abstract

This paper presents an analysis of how the iris recognition is impacted by eye diseases and an appropriate dataset comprising 2996 iris images of 230 distinct eyes (including 184 illness-affected eyes representing more than 20 different eye conditions). The images were collected in near infrared and visible light during a routine ophthalmological practice. The experimental study shows four valuable results. First, the enrollment process is highly sensitive to those eye conditions that make the iris obstructed or introduce geometrical distortions. Second, even those conditions that do not produce visible changes to the iris structure may increase the dissimilarity among samples of the same eyes. Third, eye conditions affecting iris geometry, its tissue structure or producing obstructions significantly decrease the iris recognition reliability. Fourth, for eyes afflicted by a disease, the most prominent effect of the disease on iris recognition is to cause segmentation errors. To our knowledge this is the first database of iris images for disease-affected eyes made publicly available to researchers, and the most comprehensive study of what we can expect when the iris recognition is deployed for non-healthy eyes.

1. Introduction

Intricate and unique iris patterns provide attractive biometric features that, when calculated for healthy eyes, allow to build large biometric applications, such as India’s AADHAAR‡ with more than 600 million people enrolled to date, or the CANPASS§ system maintained by Canadian Border Services Agency (CBSA) and providing an efficient entry into Canada for frequent travelers. However, there are various medical conditions affecting the eye structures, in particular the iris, and possibly deteriorating the recognition reliability. When investigating past work (Sec. 2) it is clear that due to the lack of large, heterogeneous, and publicly available databases appropriate to this subject, we are still far from full understanding of how various eye conditions impact iris recognition. This paper aims at answering four crucial questions related to eye conditions and their impact on iris recognition, and delivers an appropriate database of non-healthy eye images to researchers.

Question 1: Do ocular pathologies impact the enrollment process? If so, presence of which structural impairments translates into an increase in Failure To Enroll rate (FTE), i.e., the proportion of samples that could not be enrolled to the overall number of samples?

Question 2: Does the iris recognition perform worse for non-healthy eyes with no visible impairments when photographed in near infrared (NIR) light, when compared to the healthy irides? In other words, can we assume that there are some properties of the iris image, not revealed in near infrared, that prevent iris recognition from achieving optimal performance?

Question 3: What kind of visible impairments in non-healthy irides impact the iris recognition to the highest extent?

Question 4: What are the main reasons for a bad performance when iris recognition is applied to non-healthy eyes?

To answer these questions a dataset of iris images representing more than twenty different eye diseases was built with the use of a professional iris recognition camera operating in NIR light, along with an ophthalmological commentary (Sec. 3). Most of the NIR samples are accompanied by color images to make independent ophthalmological interpretations possible. Experimental study done for

https://uidai.gov.in
http://www.cbsa-asfc.gc.ca/prog/canpass/canpassair-eng.html
three different iris recognition methods is presented in Sec. 4 To our knowledge, this paper offers the largest published dataset of NIR and color images for non-healthy eyes with a professional, ophthalmological commentary, and the most extensive study of how different groups of diseases may influence iris recognition.

2. Related work

Roizenblatt et al. [2] were probably the first to analyze how eye pathologies might influence the performance of iris recognition. They report False Non-Match Rate (FNMR) equal to 11% when post-cataract-surgery iris images are compared to those obtained beforehand. A deterioration in FNMR in such scenario is also reported by Seyeddain et al. [8]. Trocielewicz et al. [10] report worse performance for three different iris recognition methods when pre-surgery, cataract-affected eyes are used in the system instead of healthy ones. Yuan et al. [11] claim that laser procedures used to ablate the corneal tissue to compensate for refractive pathologies such as myopia, hypermetropia and astigmatism have little effect on iris recognition accuracy. Aslam et al. [1] present the first known to us work examining biometric performance in relation to more than one eye condition, including corneal and scleral pathologies, glaucoma, scleritis and conjunctivitis. A single iris recognition methodology used by the authors turned out to be resilient for most diseases except the anterior uveitis (iritis), in which case the authors report FNMR=21%. McConnon et al. [4] examine the iris segmentation process (yet not the iris matching) for color images acquired by an ophthalmoscope, reporting variations by more than 2 pixels from the “ground truth” obtained by manual segmentation in about half of images. Borgen et al. [2] build a synthetic set gathering digitally modified UBIRIS samples so that to resemble selected changes to the eye structures such as keratitis, corneal infiltrates, blurring and dulling of the cornea, corneal scarring and surgery, angiogenesis, iridectomy, iris depigmentation, tumors and melanoma. High false non-match rates (32.8% – 86.8%) are reported for all modifications except for the pathological vascularization (6.6%), changes in iris color (0.5%) and iridectomy (0%). It is worth noticing that most of the presented studies use a single iris recognition methodology when generating final conclusions, and all of them are based on small datasets, acquired from no more than 100 subjects.

3. Database of iris images

3.1. Eye pathologies

The iris, located inside the eyeball, is separated from the outside environment by the eyes protective system – the eyelids – also known as palpebra, the cornea covered with tear film and the aqueous humor which fills the enclosed cavity between the cornea and the iris. In a state of normal anatomy – open palpebral fissure, transparent cornea and clean anterior chamber of the eye – undisturbed observation of the iridial surface is possible. However, temporary or permanent changes in the anterior part of the eye may inhibit obtaining iris images of appropriate quality. In this subsection we discuss ophthalmic pathologies represented in the database, which may prevent the use of iris biometrics technologies in practice, in respect to parts of the eye they impact most.

The cornea. Chemical injury can deal extensive damage to the ocular surface epithelium, the cornea, and the anterior segment of the eye. It can lead to opacification and neovascularization of the cornea, formation of a symblepharon and cicatrical ectropion or entropion. If significant corneal scarring is present, a corneal transplant may be required to restore vision. Pterygium – a benign growth of the conjunctiva, commonly forms from the nasal side of the sclera to the center of the cornea. This fibrovascular proliferation often occludes a part of the iris. Bacterial keratitis is an erosion or an open sore in the outer layer of the cornea with stromal infiltration, edema and hypopyon. Common pathogens that may lead to corneal ulcers include: Streptococcus pyogenes, Acanthamoeba, Herpes simplex, or fungal infections mainly caused by use of non-sterilized contact lenses. Acute glaucoma with increased pressure inside the eye occurs suddenly when the iris is pushed or pulled forward. High intraocular pressure produces symptoms such as corneal edema, shallowness of the anterior chamber and dilatation of the pupil which may become oval in shape. Corneal laceration usually requires placement of corneal sutures.

The anterior chamber. Hyphema is a condition in which blood is present in the anterior chamber of the eye and may partially or completely obstruct the view of the iris. Hyphemas are frequently caused by injuries but may also occur spontaneously. A long-standing hyphema may result in hemosiderosis and heterochromia in form of partial changes in the coloration of the iris. Hypopyon is a leukocytic exudate present in the anterior chamber of the eye – partially obstructing the iris, usually accompanied by redness of the conjunctiva. It is a sign of an iridial inflammation.

The iris. Rubeosis iridis is a medical condition of the iris in which new, abnormal blood vessels are found on the surface of the iris. It is usually associated with disease processes in the retina. Iris sphincter tear is a frequent concomitant of both laceration and blunt trauma of the anterior segment. Iridodialysis is defined as a rupture of the iris at its thinnest area, the iris root, manifested as a separation or tearing of the iris from its attachment to the ciliary body. It is usually caused by blunt trauma to the eye, but may also be caused by penetrating eye injuries or as a complication of an intraocular surgery. Iridodialyses can be often repaired us-
ing suturing techniques. Synechiae are adhesions between the iris and other structures in the eye. Iris bombe occurs when there is a complete adhesion (posterior synechiae) between the iris and the anterior capsule of the lens creating a 360-degree area of the adhesion.

The lens. With anterior lens luxation, the lens enters the anterior chamber of the eye. This can cause damage to the cornea, swelling and progression of lens opacity, so the iris image can become blurry. Phacolytic glaucoma is an inflammatory glaucoma caused by the leakage of lens protein through the capsule of a hyper-mature cataract. Escalating corneal edema and milky aqueous humor in the anterior chamber also blur the iris image.

Pars plana vitrectomy is a general term used to describe a group of surgical procedures performed in the deeper part of the eye, behind the lens. Silicone oil is used as an intraocular tamponade in the repair of retinal detachment or diabetic retinopathy. Sometimes it may relocate itself to the front of the iris, causing a blurry iridal image.

3.2. Data collection protocol

For this study a new database was designed and collected specifically for the assessment of how the iris recognition is immune, or prone, to ocular pathologies. The dataset comprises images collected from patients during routine ophthalmology examinations. All patients participating in the study have been provided with detailed information on the research and a written consent has been obtained from each volunteer.

Data collection lasted approximately 16 months. During each visit, both NIR-illuminated images (compliant with the ISO/IEC 19794-6:2011 recommendations) and standard color photographs (for selected cases) were acquired to perform visual inspection of the illness in samples showing significant alterations to the eye. The data have been acquired with three commercial cameras: 1) the IrisGuard AD100 for NIR images, 2) Canon EOS1000D with EF-S 18-55 mm f/3.5-5.6 lens equipped with a Raynox DCR-250 macro converter and a ring flashlight suited for macrophotography, and 3) an ophthalmology slit-lamp camera Topcon DC3, Tab. 1.

3.3. Database statistics

The entire dataset comprises 2996 images of 230 distinct irides, Tab. 1. Every class contains NIR-illuminated images, while for some of them visible spectrum photographs have also been taken (in cases when visual inspection revealed significant changes to the structures of the eye). Fig. 1 shows samples of five different eyes obtained using all three devices.

For 184 images there is one sample collection session, for 38 classes there are two sessions, for 6 classes there are three sessions, and finally for 2 classes there are four different acquisition sessions. Usually, the second and subsequent sessions contain images obtained after some kind of medical procedure, e.g., a cataract surgery. Detailed information, including precise description of medical conditions and procedures performed in each case, is disclosed in the metadata that accompany the dataset. No data censoring was performed when issuing the dataset, except for removing images that did not show an eye at all.

Table 1. Formats and numbers of images collected by each device.

| Device          | Image format     | Number of images |
|-----------------|------------------|------------------|
| IrisGuard AD100 | grayscale, 640x480 BMP | 1793             |
| Canon EOS1000D  | color, 10 Mpixel JPEG  | 868              |
| Topcon DC3      | color, 8 Mpixel JPEG  | 335              |

4. Experimental study

4.1. Dividing the data

Ophthalmological practice shows that most non-healthy eyes suffer from more than one condition, which are often unrelated and impacting the eye in different ways. While some illnesses cause the pupil to distort and deviate from its usual circular shape, other pathologies impact the iris directly or cause changes to other parts of the eyeball, such as the uvea, the cornea, the anterior chamber, or even to the retina. Hence, conducting an insightful analysis can be challenging. The data were therefore partitioned respectfully to the type of influence that given ocular pathology inflicts on the eye. This allows us to devise five different partitions: Healthy partition, comprising healthy eyes only, Clear partition made up of eyes with a disease present, but not affecting the eye structures perceivably, Geometry partition (eyes with pupil geometry distorted by the pathology), Tissue partition (eyes with damage inflicted on the iris tissue) and Obstructions partition, encompassing the eyes with obstructions present in front of the iris. Table 2 presents numbers of classes (i.e., different eyes) and images in each category. Figure 2 shows sample images belonging to each partition.

4.2. Methodology

To answer the first question formulated in the introductory part, the failure-to-enroll error rates are calculated for each database partition and three different iris recognition methods. To answer questions two and three, all possible genuine comparisons were generated and full cross comparisons were made to get all the possible impostor comparison scores for each dataset partition. One-tailed two-sample t-test at the significance level \( \alpha = 0.05 \) is used to compare averages of the comparison scores across partitions. That is, in all t-tests the null hypothesis states that the comparison scores calculated for samples belonging to two partitions...
being compared (for instance, Healthy and Clear samples) come from independent random samples with equal means (equality of variances is not assumed). Additionally to testing the equality of means, we run two-sample Kolmogorov-Smirnov test at the significance level $\alpha = 0.05$ to check whether samples belonging to two partitions come from the same distribution. These analyses were done independently for three iris recognition methods. Finally, the segmentation errors are analyzed and visual inspection of the eye samples resulting in the worst comparison scores are performed to answer the last, fourth question.

4.3. Iris recognition methods used in this research

Three independent iris matching methods are used in this work and briefly characterized in this subsection. MIRLIN (Monro Iris Recognition Library and INterface) is a commercial implementation of the concept proposed by Monro et al. [5]. In this approach the iris features are derived from zero-crossings of the differences between Discrete Cosine Transform (DCT) calculated in rectangular iris image patches. The resulting iris binary feature sets are compared by calculating fractional Hamming distance. Two matching irides should have this distance close to zero, while 0.5 is expected when two different eyes are being matched. OSIRIS (Open Source for IRIS) follows a well-known iris recognition concept proposed by Daugman and includes an iris image normalization into a dimensionless polar representation, Gabor-based filtering, quantization of the filtering results into a binary iris code, and code matching based on fractional Hamming distance [9]. As in MIRLIN method, close-to-zero comparison scores are expected when same-eye samples are compared, and 0.5 should be obtained for non-same-eye images. Details of the third method, VeriEye, are not revealed by the vendor, apart from the claim of using non-circular approximations of the iris and pupil boundaries, and coding which does not follow Gabor filtering [6]. The comparison score starts from zero (when totally different eyes are compared) to some, yet unknown high values when same-eye samples are compared.

5. Results

5.1. Enrollment performance (re: Question 1)

FTE rates obtained in each partition (Tab. 2) suggest that iris recognition performs particularly bad for samples included in Geometry and Obstructions partitions. Those partitions comprise images with pupil either distorted or not visible at all due to various types of occlusions. Hence, the answer to the question 1 is that the enrollment stage is highly sensitive to those conditions that make the pupil geometry distorted, or the iris pattern partially obstructed.

5.2. Matching performance (re: Questions 2 and 3)

To check whether the iris recognition performs worse for non-healthy eyes (although not revealing visible changes) than healthy eyes (question 2), the genuine and impostor scores for Healthy and Clear eyes are compared (Fig. 2, solid lines) for those irides that were correctly enrolled. For all iris coding methodologies we observe slightly worse mean genuine comparison scores in Clear partition when compared to the Healthy eyes. Mean values of the impostor scores differ not too much. Although t-test suggests that those differences in average of genuine scores are statistically insignificant for MIRLIN and OSIRIS methods ($p$-value $> \alpha = 0.05$), the Kolmogorov-Smirnov test points at statistically significant differences in distributions for all tested methods and both for the genuine and impostor comparison scores, Tab. 3. Hence, the answer to the question 2 is affirmative, but the observed differences are uneven.
Table 2. Columns 2 and 3 show the number of samples and unique irides in each dataset partition. Columns 4 through 6 present FTE rates obtained in each partition for three iris recognition methods used in this work. The worst results for each method are bolded.

| Data partition | Number of irides | Number of images | FTE for MIRLIN | FTE for VeriEye | FTE for OSIRIS |
|----------------|------------------|------------------|----------------|----------------|----------------|
| Healthy        | 35               | 216              | 1.85%          | 0%             | 0%             |
| Clear          | 87               | 568              | 4.40%          | 0%             | 1.23%          |
| Geometry       | 53               | 312              | 16.03%         | 5.13%          | 5.45%          |
| Tissue         | 8                | 50               | 2%             | 0%             | 0%             |
| Obstructions   | 36               | 207              | 18.36%         | 3.86%          | 8.21%          |

Figure 2. Receiver operating characteristics (ROC) for MIRLIN (left), OSIRIS (middle) and VeriEye (right) iris recognition methods.

5.3. Sources of errors (re: Question 4)

To seek for probable sources of erratic performance, sample pairs yielding genuine match scores below typical acceptance thresholds (i.e., 0.32 for MIRLIN and OSIRIS, and 30 for VeriEye) were visually inspected. The most likely cause or errors is a failed segmentation that causes the algorithm to encode non-iris image portions. These segmentation errors were caused mostly by either irregular pupil boundary, corneal occlusions that obstruct the pupil and the iris, or by damage to the iris tissue being misinterpreted by the pupil segmentation algorithms as the pupil itself. In particular, all of the sample pairs generating the worst OSIRIS scores (Geometry and Obstructions) devised their poor performance from the segmentation errors. This is similar when the MIRLIN matcher is involved, except for two images that produce poor scores in the Clear subset: one of which is blurred, and the other showing no particular flaw. For the VeriEye matcher there is no way of assessing if segmentation is done correctly, however, when looking at those samples that perform the worst, one may identify conditions responsible for errors, namely: significant geometrical distortions, severe corneal hazes, blurred boundary between the iris and the pupil. Thus, segmentation failures may be the case here as well.

6. Conclusions

This paper provides the most thorough analysis of iris recognition performance in the presence of various ocular pathologies to date. The fact that usually a few different and unrelated diseases might appear in a single eye, requires...
Table 3. Summary of statistical testing of differences in average comparison scores obtained for all iris recognition methods in five partitions of the dataset. In t-test the null hypotheses state that the comparison scores calculated for samples belonging to two partitions being compared come from distributions of the same mean. Alternative hypotheses are defined separately for each case and they are shown in rows labeled ‘H1 (t-test)’. In Koltmogorov-Smirnov test the null hypotheses state that the data of two compared partitions are from the same distribution. Alternative hypotheses state the opposite and are detailed in rows labeled ‘H1 (KS-test)’. In both tests, p-values not exceeding α = 0.05 denote that the null hypothesis was rejected and the corresponding alternative hypothesis was selected. g and i denote the genuine and impostor scores, correspondingly, while ¯ denotes inequality of distributions.

| Method | Genuine comparison scores | Impostor comparison scores |
|--------|---------------------------|---------------------------|
|        | Healthy (g) | Clear (g) | Geometry (g) | Tissue (g) | Obstr. (g) | Healthy (i) | Clear (i) | Geometry (i) | Tissue (i) | Obstr. (i) |
| MIRLIN | 0.0198 | 0.0236 | 0.0897 | 0.0773 | 0.1080 | 0.4053 | 0.4027 | 0.4050 | 0.4158 | 0.4120 |
|        | H1 (t-test) | g_{c} > g_{h} | g_{g} > g_{h} | g_{t} > g_{h} | g_{o} > g_{h} | i_{c} < i_{h} | i_{g} < i_{h} | i_{t} < i_{h} | i_{o} < i_{h} |
|        | p-value (t-test) | 0.0466 | -0 | 0 | -0 | 0.2453 | -0 | -0 | -0 |
|        | H1 (KS-test) | g_{c} ≁ g_{h} | g_{g} ≁ g_{h} | g_{t} ≁ g_{h} | g_{o} ≁ g_{h} | i_{c} ≁ i_{h} | i_{g} ≁ i_{h} | i_{t} ≁ i_{h} | i_{o} ≁ i_{h} |
|        | p-value (KS-test) | 0.0178 | -0 | -0 | -0 | 0.0022 | -0 | -0 | -0 |
| OSIRIS | mean | 0.2424 | 0.2483 | 0.3084 | 0.2701 | 0.3100 | 0.4688 | 0.4656 | 0.4675 | 0.4713 | 0.4700 |
|        | H1 (t-test) | g_{c} > g_{h} | g_{g} > g_{h} | g_{t} > g_{h} | g_{o} > g_{h} | i_{c} < i_{h} | i_{g} < i_{h} | i_{t} < i_{h} | i_{o} < i_{h} |
|        | p-value (t-test) | 0.0536 | -0 | 0.0012 | -0 | -0 | 0.002 | -0 | -0 |
|        | H1 (KS-test) | g_{c} ≁ g_{h} | g_{g} ≁ g_{h} | g_{t} ≁ g_{h} | g_{o} ≁ g_{h} | i_{c} ≁ i_{h} | i_{g} ≁ i_{h} | i_{t} ≁ i_{h} | i_{o} ≁ i_{h} |
|        | p-value (KS-test) | 0.0184 | -0 | 0.0024 | -0 | 0.0049 | -0 | -0 | -0 |
| VerEye | mean | 500.18 | 458.50 | 282.28 | 447.92 | 265.62 | 3.479 | 2.998 | 1.799 | 2.598 | 2.080 |
|        | H1 (t-test) | g_{c} < g_{h} | g_{g} < g_{h} | g_{t} < g_{h} | g_{o} < g_{h} | i_{c} > i_{h} | i_{g} > i_{h} | i_{t} > i_{h} | i_{o} > i_{h} |
|        | p-value (t-test) | -0 | -0 | 0.0001 | -0 | -0 | -0 | -0 | -0 |
|        | H1 (KS-test) | g_{c} ≁ g_{h} | g_{g} ≁ g_{h} | g_{t} ≁ g_{h} | g_{o} ≁ g_{h} | i_{c} ≁ i_{h} | i_{g} ≁ i_{h} | i_{t} ≁ i_{h} | i_{o} ≁ i_{h} |
|        | p-value (KS-test) | -0 | -0 | 0.0173 | -0 | -0 | -0 | -0 | -0 |

Further information on how to get the data can be found at: [http://zbum.ia.pw.edu.pl/En]-Research-Databases

References

[1] T. M. Aslam, S. Z. Tan, and B. Dhillon. Iris recognition in the presence of ocular disease. J. R. Soc. Interface 2009, 6, 2009.

[2] H. Borgen, P. Bours, and S. D. Wohlhusen. Simulating the Influence of Aging and Ocular Disease on Biometric Recognition Performance. International Conference on Biometrics 2009, LNCS 5538, 8(8):857–867, 2009.

[3] ISO/IEC 19794-6:2011. Information technology – Biometric data interchange formats – Part 6: Iris image data, 2011.

[4] G. McConnon, F. Deravi, S. Hoque, K. Sirlantzis, and G. Howells. Impact of Common Ophthalmic Disorders on Iris Recognition. 2012 5th IAPR International Conference on Biometrics Compendium, IEEE, pages 277–282, 2012.

[5] D. M. Monro, S. Rakshit, and D. Zhang. DCT-based iris recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence – Special Issue on Biometrics: Progress and Directions, 29(4):586–595, April 2007.

[6] Neurotechnology. VeriEye SDK, v. 4.3 (accessed Feb 28, 2015): www.neurotechnology.com/verieye.html.

[7] R. Roizenblatt, P. Schor, F. Dante, J. Roizenblatt, and R. Belfort Jr. Iris recognition as a biometric method after cataract surgery. BioMedical Engineering Online, 3(2), 2004.

[8] O. Seydadd, H. Krakar, A. Redberger, A. K. Dexl, G. Grabner, and M. Emesz. Reliability of automatic biometric iris recognition after phacoemulsification or drug-induced pupil dilation. Eur J Ophthalmol, 24(1):58–62, 2014.

[9] G. Sutra, B. Dorizzi, S. Garcia-Salitcetti, and N. Othman. A biometric reference system for iris. OSIRIS version 4.1 (accessed Oct 1, 2014): http://svnext.it-sudparis.eu/svnview2-eph/ref_syst/iris_osiris_v4.1/

[10] M. Trokielewicz, A. Czajka, and P. Maciejewicz. Cataract influence on iris recognition performance. Proc. SPIE 9290, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2014, 9290, 2014.
[11] X. Yuan, H. Zhou, and P. Shi. Iris recognition: a biometric method after refractive surgery. *J Zhejiang Univ Sci A*, 8(8):1227–1231, 2007.