Isolating Iris Template Ageing in a Semi-controlled Environment

Heinz Hofbauer
University of Salzburg
Salzburg, Austria
hhofbaue@cosy.sbg.ac.at

Inmaculada Tomeo-Reyes
Queensland University of Technology
Brisbane, Australia
inma.tomeoreyes@qut.edu.au

Andreas Uhl
University of Salzburg
Salzburg, Austria
uhl@cosy.sbg.ac.at

***Abstract*—The ageing of the iris and its influence on recognition performance is a hotly debated topic. Template ageing in iris biometrics is often attributed to changes during acquisition rather than the natural ageing of the iris. In this paper, we attempt to isolate template ageing effects in iris recognition by using a semi-controlled environment (selecting a subset from a larger database) and controlling as many factors as possible in the biometric tool-chain. This includes manual segmentation and the use of a non-linear biomechanical model to alleviate the influence of pupillary dilation. Despite our efforts, we still find iris template ageing to be apparent in the results.**

I. Introduction

There is a certain controversy regarding template ageing in iris recognition. The IREX VI report [1] by Grother *et al.* claims that there is no evidence of an iris template ageing effect. Bowyer and Ortiz [2] critically evaluated the IREX VI report and Grother *et al.* [3] replied to the critical evaluation.

One of the points of disagreement was the definition and, consequently, the evaluation of iris ageing. The ISO/IEC 19795–1 standard on biometric performance testing and reporting defines template ageing as the increase in error rates caused by time-related changes, meaning changes in the biometric pattern, the presentation of the pattern and the sensor.

The IREX VI report excludes all the effects which are not related to "irreversible changes to the anatomy, primarily the iris texture" and states that "dilation should not be considered part of ageing because it varies stochastically and can be mitigated". It also defends that the increased rejection rates attributed to ageing can often be reduced to increased rejection rates due to blur, noise, occlusion and pupil dilation. Bowyer and Ortiz criticize this view and the fact that, of the assumed reversible effects, only dilation is considered in the report.

There are ways to control certain acquisition effects like illumination, which can be changed in a controlled way, or blinking, which can be prevented by requiring reacquisition. However, some of these effects are not isolated. Take illumination, for example. It is possible to provide a constant illumination of the iris texture, but by doing that, it is not possible to ascertain a constant pupillary dilation, because pupil size changes with age. It is well known from both the medical [4] and biometrics context [5] that pupil size is affected by age, with smaller pupils being predominant among the elderly population. This relates directly to the disregard of illumination and/or dilation changes as apparent ageing effects by the IREX VI report – either the illumination is controlled to produce a constant lighting and consistent exposure of the iris texture, or to unify the pupillary dilation. In the authors' view, it is correct to exclude reversible or preventable effects when evaluating iris ageing; however, it should not be assumed that the effects exist in isolation.

Overall, there are a number of questions which remain to be answered regarding iris template ageing. As Grother *et al.* in IREX VI, some authors claim that iris ageing is of a negligible significance [6], while results presented by other authors [7]–[9] suggest that the iris template ageing effect is significant. In this paper, we present a database which is a subset of the upcoming CASIA v5.0 subject ageing database, and we try to control as many factors as possible and analyze if ageing is still apparent. This is done by a careful selection of the images in the database and an analysis of the possible effects during acquisition. We provide a ground truth for segmentation which further helps to reduce the impact of acquisition effects on the results.

II. The CASIA Iris Ageing Database

The database used in this paper is a subset of the upcoming CASIA v5.0 Iris Database[1]. The full original database contains 120 images per eye and user from video sequences captured in 2009, and 20 images per eye and user from video sequences captured in 2013. Wild *et al.* [10] used the full database to analyse the impact of segmentation and quality on iris recognition performance. Since our goal is to eliminate as many acquisition effects as possible, we used the results from their work to guide our selection.

In 2009, six different recording sessions were carried out, with different characteristics regarding pupillary dilation, occlusion, blur, and use of glasses. Wild *et al.* found that most sessions have a negative impact on the recognition rate, suggesting a rejection of all sessions except sessions 1 and 4. We based the selection of users on the *2009 session 1*, which contains images without glasses. In 2013, images were recorded in a more controlled setup, showing more consistent pupillary dilation and less frequency of glasses. We similarly used the session without glasses in 2013. From these sessions we removed some individuals: 0023, 1067 (no data in 2009);

[1] see http://www.biometrics.idealtest.org

1004 (no images without glasses in 2013); 0190, 0191 (no data in 2013).

Since images in the database are from a video sequence of the eyes, temporal proximity resulted in almost equal images. To avoid that, we chose ten images per eye, spaced apart as fas as possible in the temporal domain. Images in which large parts of the iris were occluded by the eyelids were not taken into consideration. The resulting database consists of 1880 images captured from both eyes of 47 users. There are 10 images per eye and year.

III. Analysis of the Database

To perform the experiments we use USITv2 (University of Salzburg Iris Toolkit v2.0.x [11], [12]), a publicly available iris recognition software package which comprises different algorithms for iris pre-processing, feature extraction, and comparison. *Segmentation* is performed using a manual method (`manuseg`) and a method based on contrast-adjusted Hough transform (`caht`) proposed by Rathgeb *et al.* [13]. Manual segmentation uses a single operator, similarly to IRISSEG-EP [14]. The inner and outer iris boundaries are segmented using an ellipse fitting, and the lower and upper eyelids using a polynomial fitting[2]. *Normalization* is performed using the rubber sheet model [15]. To counter dilation changes, we also use the nonlinear biomechanical-based normalization method proposed by Tomeo-Reyes *et al.* [16]. *Feature extraction* is based on 1D log-Gabor filters (`lg`), as proposed by Masek [17].

A. Impact of Acquisition Effects

The objective in this section is to analyze those acquisition effects which can lead to apparent iris template ageing. The following effects were considered:

- *Head tilt* was analysed and found to be in the range of ± 24 bit ($\pm 16.875°$). Rotation was compensated for in the experiments by using the relevant USIT [11] features.
- *Sensor ageing* has no significant influence as shown by Bergmüller *et al.* [18]. The database they used for their research is a superset of the database we are using, so the results clearly apply. They extracted an ageing model of the sensor based on the four-year gap and applied the effects of an artificial sensor-ageing to the images, which produced a negligible influence for up to 96 years of sensor-ageing.
- *Obfuscation of the iris by the eyelids*, due to blinking or similar effects, was minimized by selecting images from the database where this does not occur.
- *Obfuscation of the iris by glasses*, which might introduce additional reflections and distortions, was prevented by choosing images of subjects without glasses.
- *Illumination conditions* were analysed and found to be relatively stable over the two years. A comparison of the aggregate histograms of both years shows a close fit, with the main difference being a 1-bit shift in brightness towards white in 2013 compared to 2009. The overlap coefficient between the aggregate histograms of 2009 and 2013 is 73.13% (no shift) and 94.83% (correcting the 1-bit shift).
- *Blurring effects* and *visible interlacing* appear in both the 2009 and 2013 images, and increase the error rates.

[2]The list of iris images and manual segmentation data for the database used in this paper is available online at http://www.wavelab.at/sources/Hofbauer16d. Original images will be available in the upcoming CASIA v5.0 Iris Database.

B. Impact of Segmentation

This experiment aims at finding whether changes in iris recognition performance between 2009 and 2013 are due to template ageing effects or if they are linked to segmentation errors. Wild *et al.* [10] tested the impact of segmentation on the CASIA 5.0 database and found that segmentation algorithms "do not change the overall picture of observable impact of an increased timespan. However, depending on the choice of quality-based filtering, this impact is either more pronounced or less pronounced than other variability". The use of manual segmentation should remove this variability.

Using `caht` and manual segmentation, we performed a full comparison for each year (intra-year tests) and between years 2009 and 2013 (inter-year tests). The VeriEye tool by Neurotechnology was used to ensure that the results and the apparent template ageing are not a systemic problem of the USIT. The results are given in Table I.

The difference between intra- and inter-year comparisons is very pronounced, with worse performance in inter-year tests. Manual segmentation leads to better performance than `caht`. While this improvement affects all comparisons similarly, the apparent ageing effect is not removed, as shown by the fact that inter-year error rates are higher than intra-year. The segmentation and features of VeriEye are superior, which is known [12], but the system shows the same properties when it comes to apparent template ageing. Poor segmentation *might* be mistakenly identified as template ageing, but this is not the

TABLE I
Intra- and inter-year performance given as equal error rate (EER) and FNMR at operation point FMR=0.01% (OP 0.01). It may be noted that VeriEye uses its own segmentation and features.

	Manual		`caht`		VeriEye	
Comparison	EER [%]	OP 0.01 [%]	EER [%]	OP 0.01 [%]	EER [%]	OP 0.01 [%]
2009–2009	2.512	13.197	4.184	9.722	1.346	2.458
2009–2013	3.531	55.442	8.810	56.722	1.678	3.864
2013–2013	0.312	1.532	5.135	6.327	0.709	0.709

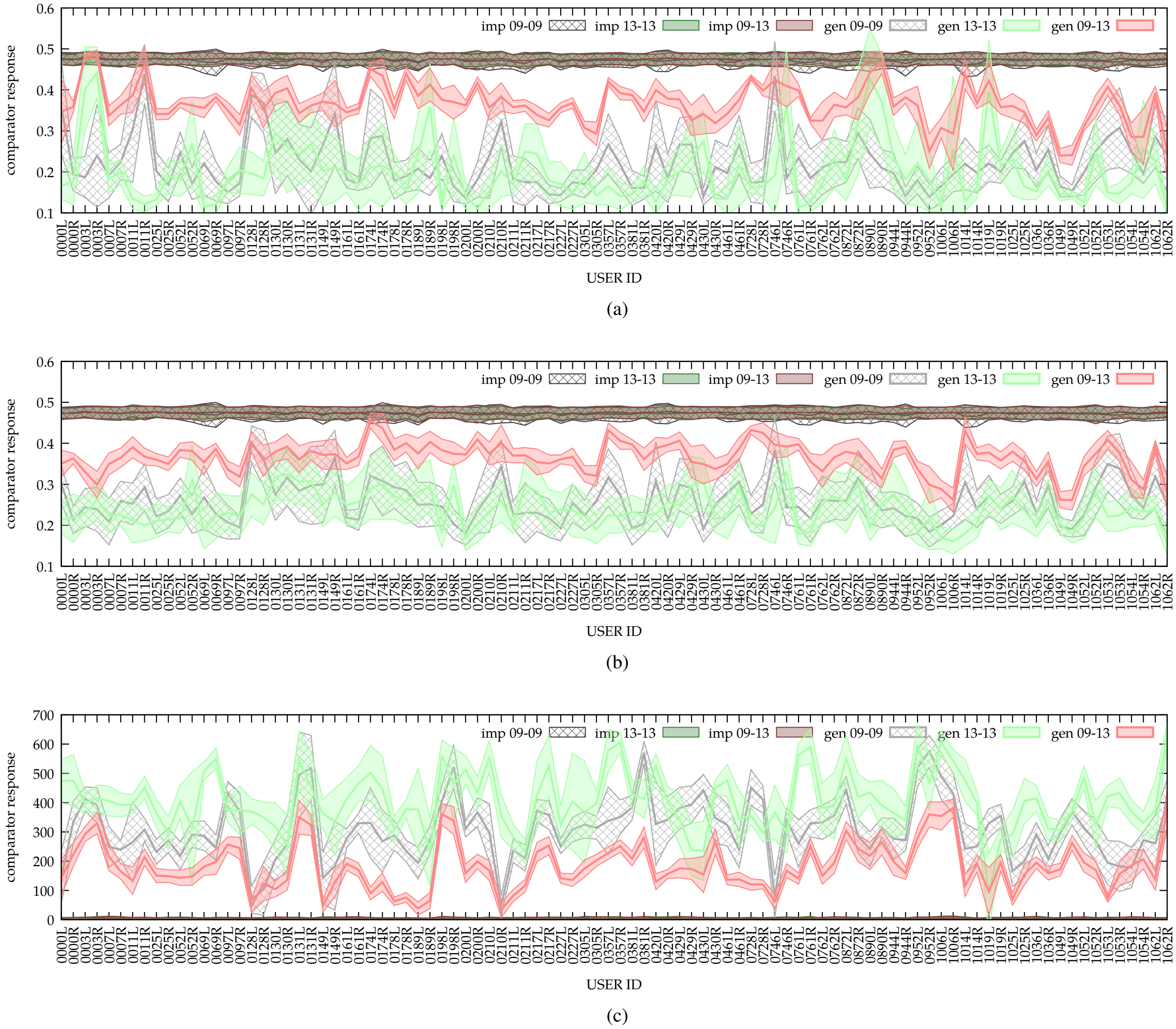

Fig. 1. Hamming scores, mean ± one standard deviation, for genuines and imposters grouped by user. (a) Segmentation with `caht` and `lg`-features. (b) Manual segmentation and `lg`-features. (c) Neurotechnology VeriEye. Note that, unlike (a) and (b), lower scores are worse in this case.

case here since even manual segmentation leads to a worse inter-year than intra-year performance.

To further support this conclusion, we plotted the mean and standard deviation of the Hamming scores per user for genuine and imposter comparisons in Figure 1. The improvement due to manual segmentation can be clearly seen, but the difference between intra- and inter-year statistics is very pronounced even with manual segmentation. In this experiment, it can also be observed that `caht` segmentation does not simply degrade all results; rather, there seem to be specific users for which the segmentation fails. An example of this is ID 0003 (both eyes), where `caht` clearly introduces errors. For other users like ID 0007 (both eyes), the difference between `caht` and manual segmentation is negligible.

Overall, the use of manual segmentation does not eliminate the apparent template ageing effects, but it does not introduce additional errors either, and in that sense it is preferable.

C. *Impact of Pupillary Dilation*

To ascertain the influence of pupillary dilation on the performance of the iris recognition system, we used the biomechanical-based normalization scheme by Tomeo-Reyes *et al.* [16]. Results obtained with this scheme were compared with those obtained with the rubber sheet model (RSM) used by the USIT (following Daugman's approach [15]).

In [16], a biomechanical model (BMM) based on the work of Clark *et al.* [19] was used to define a non-linear normalization scheme that improves iris recognition under light-induced pupil dilation. The BMM was demonstrated to result in superior performance compared to RSM when matching iris images exhibiting significant differences in dilation. The performance of both models was shown to be quite similar when matching iris images with small dilation differences, with slightly better results obtained by the RSM.

Manual segmentation was used to perform the experiments

TABLE II

PERFORMANCE COMPARISON OF RSM AND BMM NORMALIZATION, WITH BOTH *BioC* AND *BioI* SETUP. THE TABLE SHOWS EQUAL ERROR RATE (EER) AND FNMR AT OPERATION POINT FMR=0.01% (OP 0.01).

	Rubber sheet		*BioC*		*BioI*	
Comparison	EER [%]	OP 0.01 [%]	EER [%]	OP 0.01 [%]	EER [%]	OP 0.01 [%]
gallery(2009) – 2009	1.773	9.235	1.843	13.404	2.178	12.191
gallery(2009) – 2013	2.548	46.393	4.778	62.281	2.979	53.813

reported in this section. To quantify the pupillary dilation, the ratio between the pupil and iris radii was used. This ratio, denoted as D, is referred to as dilation ratio ($D = R_{pupil}/R_{iris}$). Since the manual segmentation process uses ellipses to define the boundaries, the pupil and iris radii are calculated by using approximated circles that have the same area as the corresponding ellipses. The radius R of each circle is calculated as $R = \sqrt{a * b/4}$, where a and b are the major and minor axes of the ellipse.

The BMM estimates the non-linear displacement of the iris tissue using the difference in dilation ratio between the sample and a reference iris image (see [16] for further details). To evaluate a realistic scenario, a gallery of enrolled images was chosen. The gallery contains one image per eye from 2009, selected by visual inspection. The enrolled images are used as the reference images.

The BMM is used in this paper to analyze whether changes in performance between 2009 and 2013 are due to template ageing effects or linked to pupillary dilation changes. It may be noted that the BMM assumes a positive radial displacement of the pupillary boundary from the resting state (given by the dilation ratio of the reference image), so the model is not optimized for constriction. Keeping this in mind, two experiments were defined:

- *BioC*: This experiment assumes that only iris codes are stored in the gallery, i.e., only the query images can be normalized using the BMM. In this case, all query images (regardless of the fact that they are dilated or constricted in comparison with the reference) are normalized using the BMM.
- *BioI*: This experiment assumes that iris images are stored in the gallery, and the BMM is applied to the gallery or query image depending on which one shows a higher dilation degree. In this case it is possible to properly account for dilation and constriction, since positive radial displacement of the pupillary boundary is always assumed.

In both cases the gallery was compared to the query images from 2009 (intra-year) and to query images from 2013 (inter-year). The results of these two experiments are given in Table II, together with a regular experiment using only RSM-based normalization.

From the *BioC* test we can see that the BMM does not perform as well as the RSM. For intra-year comparisons, the main reason is that the difference in dilation ratio between the reference and the sample is too small to reach the point where we can observe the non-linear behaviour of the iris tissue and use the BMM effectively. The RSM can handle these small variations. For inter-year comparisons, the performance degradation is related to the fact that constriction is not properly accounted for in this scenario. The *BioI* test improves on the inter-year comparison compared to *BioC* but reduces the intra-year performance. The intra-year performance suggests that the BMM is not effective when the difference in dilation degree between the reference and the sample is too small to result in non-linear iris tissue displacement. The improvement of the inter-year comparison suggests that the dilation difference over the four years is larger and shows a non-linear behaviour.

To further support these conclusions, additional information is presented in Table III, which shows the number of times where the RSM and BMM were used to normalize the query images in *BioI*. In the intra-year comparisons, the RSM and BMM were used relatively evenly, which suggests dilation changes resulting from blinking or other short-term changes in the behaviour or environment. On the contrary, the number of normalizations with the BMM dropped sharply in the inter-year comparisons. This means that 88% of the query images had a more constricted pupil than the 2009 reference. This is in accord with the medical study by Peterson *et al.* [4] which found that the pupillary radius decreases with age. Even so, the RSM performs slightly better than the BMM. As stated before, this result correlates well with the fact that the difference in dilation degree between the query and the gallery image is still too small to use the BMM effectively.

TABLE III

NUMBER OF TIMES THE IMAGES WERE NORMALIZED WITH THE RSM OR BMM FOR THE *BioI* TEST.

	Rubber sheet (RSM)	Biomechanical (BMM)	Total
gallery(2009) – 2009	53.43%	46.57%	79524
gallery(2009) – 2013	87.87%	12.13%	88360

An analysis of the change in dilation ratio D between years was also conducted. The one-way analysis of variance was calculated for each eye to check whether the difference in dilation ratio between years is significant. Of the 94 eyes, 72 showed a significant change with critical $p^* = 10^{-6}$ and 87 showed a significant difference with $p^*_{1\%} = 0.01$. The change in dilation ratio between years is $\mu_{\Delta D} = -0.125$ with $\sigma_{\Delta D} = 0.081$. While this decrease is statistically significant, none of the dilation ratios reach values within the range of high constriction or dilation, which can be under 0.2 and over 0.7, respectively (according to [16], these are the cases which are more effectively handled by the BMM).

The previous results support the main conclusions of this analysis, summarized as follows:

- *gallery(2009) - 2009*: The difference in dilation ratio between the gallery image and the query is too small to reach the point where we can observe the non-linear behaviour of the iris tissue and use the BMM effectively. The RSM can handle these small variations.
- *gallery(2009) - 2013*: While there is a clear decrease in dilation ratio (constriction), the results obtained with the RSM still outperform those from the BMM. The difference in dilation ratio between the gallery image and the query does not seem big enough to use the BMM effectively. Iris recognition performance is degraded regardless of the model used. This, coupled with the fact that the image acquisition process showed relatively constant illumination, suggests that the performance degradation is not only caused by the changes in pupil size, but probably by some other effect related to template ageing.

IV. Conclusion

We provide an iris database with a four-year time lapse between captures which has been selected to exclude acquisition factors which could potentially result in apparent iris template ageing. We also provide a manual segmentation of the database to minimize the influence of these factors. Despite this controlled approach, a decrease in the pupillary radius with time is still detectable. This is consistent with recent medical research results which show that the pupillary radius decreases with age. We tried to compensate the effect of the pupillary change by using a biomechanical model for nonlinear iris normalization. Despite our efforts, iris template ageing effects are still apparent, with worse iris recognition performance obtained for inter-year comparisons compared to intra-year comparisons.

Acknowledgment

This work was partially supported by the Austrian Science Fund, project no. P26630.

References

[1] P. Grother, J. Matey, E. Tabassi, G. Quinn, and M. Chumakov, "IREX VI – temporal stability of iris recognition accuracy," National Institute of Standards and Technology (NIST), Tech. Rep. NIST Interagency Report 7948, 2013.

[2] K. W. Bowyer and E. Ortiz, "Critical examination of the irex vi results," *IET Biometrics*, vol. 4, no. 4, pp. 192–199, 2015.

[3] P. Grother, J. R. Matey, and G. W. Quinn, "Irex vi: mixed-effects longitudinal models for iris ageing: response to bowyer and ortiz," *IET Biometrics*, vol. 4, no. 4, pp. 200–205(5), 2015.

[4] J. R. Peterson, L. S. Blieden, A. Z. Chuang, L. A. Baker, M. Rigi, R. M. Feldman, and N. P. Bell, "Establishing age-adjusted reference ranges for iris-related parameters in open angle eyes with anterior segment optical coherence tomography," *PloS ONE*, vol. 11, no. 1, p. 12, 2016.

[5] M. Fairhurst, Ed., *Age Factors in Biometric Processing*. IET, 2013.

[6] S. Shchegrova, "Analysis of iris stability over time using statistical regression modeling," in *Proc. of Biometrics Consortium Conference*, 2012.

[7] S. P. Fenker and K. W. Bowyer, "Analysis of template aging in iris biometrics," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2012, pp. 45–51.

[8] A. Czajka, "Influence of iris template aging on recognition reliability," *Communications in Computer and Information Science*, vol. 452, pp. 284–299, 2014.

[9] M. Trokielewicz, "Linear regression analysis of template aging in iris recognition," in *Proc. of the IEEE Int'l Workshop on Biometrics and Forensics*, 2015.

[10] P. Wild, J. Ferryman, and A. Uhl, "Impact of (segmentation) quality on long vs. short-timespan assessments in iris recognition performance," *IET Biometrics*, vol. 4, no. 4, 2015.

[11] Various, "USIT – University of Salzburg iris toolkit," http://www.wavelab.at/sources/USIT, version 2.0.x.

[12] C. Rathgeb, A. Uhl, P. Wild, and H. Hofbauer, "Design decisions for an iris recognition sdk," in *Handbook of Iris Recognition*, second edition ed., ser. Advances in Computer Vision and Pattern Recognition, K. Bowyer and M. J. Burge, Eds., 2016.

[13] C. Rathgeb, A. Uhl, and P. Wild, *Iris Recognition: From Segmentation to Template Security*, ser. Advances in Information Security. Springer Verlag, 2013, vol. 59.

[14] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl, "A ground truth for iris segmentation," in *Proc. of the 22th Int'l Conf. on Pattern Recognition*, 2014.

[15] J. Daugman, "How iris recognition works," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.

[16] I. Tomeo-Reyes, A. Ross, A. D. Clark, and V. Chandran, "A biomechanical approach to iris normalization," in *Proc. of the IAPR/IEEE Int'l Conf. on Biometrics*, 2015.

[17] L. Masek, "Recognition of Human Iris Patterns for Biometric Identification, Master's thesis, University of Western Australia, 2003."

[18] T. Bergmüller, L. Debiasi, Z. Sun, and A. Uhl, "Impact of sensor ageing on iris recognition," in *Proc. of the IAPR/IEEE Int'l Joint Conf. on Biometrics (IJCB'14)*, 2014.

[19] A. D. Clark, S. A. Kulp, I. H. Herron, and A. A. Ross., "A theoretical model for describing iris dynamics," in *Handbook of Iris Recognition*, 2013, pp. 129–150.