

© IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

A Ground Truth for Iris Segmentation

Heinz Hofbauer *, Fernando Alonso-Fernandez †, Peter Wild *, Josef Bigun † and Andreas Uhl *

* {hhofbaue, uhl, pwild}@cosy.sbg.ac.at

Department of Computer Sciences

University of Salzburg

Jakob-Haringer-Str. 2

5020 Salzburg, AUSTRIA

<http://wavelab.at>

† {feralo, josef.bigun}@hh.se

Halmstad University

Box 823, SE 301-18

Halmstad, Sweden

<http://islab.hh.se>

Abstract—Classical iris biometric systems assume ideal environmental conditions and cooperative users for image acquisition. When conditions are less ideal or users are uncooperative or unaware of their biometrics being taken the image acquisition quality suffers. This makes it harder for iris localization and segmentation algorithms to properly segment the acquired image into iris and non-iris parts. Segmentation is a critical part in iris recognition systems, since errors in this initial stage are propagated to subsequent processing stages. Therefore, the performance of iris segmentation algorithms is paramount to the performance of the overall system. In order to properly evaluate and develop iris segmentation algorithm, especially under difficult conditions like off angle and significant occlusions or bad lighting, it is beneficial to directly assess the segmentation algorithm. Currently, when evaluating the performance of iris segmentation algorithms this is mostly done by utilizing the recognition rate, and consequently the overall performance of the biometric system. In order to streamline the development and assessment of iris segmentation algorithms with the dependence on the whole biometric system we have generated a iris segmentation ground truth database. We will show a method for evaluating iris segmentation performance base on this ground truth database and give examples of how to identify problematic cases in order to further analyse the segmentation algorithms.

I. INTRODUCTION

A generic iris recognition system [1] consists of the following stages: iris image acquisition, image preprocessing, iris texture feature extraction and feature matching. The image preprocessing stage consist of iris segmentation, i.e. localization of the iris and boundary detection, and normalization, unrolling and histogram equalization, and occlusion detection and masking. The correct localization and segmentation of the iris is of critical importance to the overall performance of the biometric recognition system since errors during iris segmentation cannot be corrected at a later stage in the processing chain.

Classic iris recognition algorithms assume an ideal environment and cooperative users. Traditionally, iris segmentations algorithms are based on circular iris boundaries [2], [3]. Circular iris detection is a good assumption given ideal environments and cooperative users and a consequent frontal iris image recording. However, unaware or uncooperative users (e.g. a surveillance scenario) or non optimal environments (e.g. biometrics on the move, or gate based access systems) can result in off angle iris images with bad illumination and or stronger occlusions due to hair or cilia. For such cases the iris segmentation becomes a harder task and assumptions like circular iris boundaries no longer hold.

Iris image databases like UBIRIS [4], [5] or MobBIO [6] provide iris images acquired under non ideal conditions. Databases like UBIRIS can aid in the development of newer image segmentation algorithms which can cope with non ideal environments. However, the typical way to test such systems is to look at the performance of the full iris recognition system. With this way of testing, a failure of the iris segmentation subsystem becomes readily apparent. However, a suboptimal performance might not be so clearly noticeable [7]. The overall recognition performance is not only affected by the segmentation accuracy, but also by the performance of the other subsystems based on possible suboptimal segmentation of the iris. As such it is difficult to differentiate between defects in the iris segmentation system and effects which might be introduced later in the system.

To allow the possibility of assessing iris segmentation systems on their own, a ground truth for the iris segmentation is needed. In order to provide such a ground truth we have segmented a total of 12621 iris images from 7 Databases. This data will be made publicly available, and can be used to analyse existing and test new iris segmentation algorithms.

The rest of the paper is structured as follows. Section II gives information about the segmentation ground truth databases as well as the iris image databases it is based on. In the description we include the links to both the ground truth and the iris image databases as well as relevant papers. In section III we will give a possible set of methods to evaluate iris segmentation performance and use them to evaluate a number of iris segmentation algorithms, which are also publicly available. Section IV will conclude the paper.

II. IRIS SEGMENTATION GROUND TRUTH DATABASE

The iris segmentation database (IRISSEG) contains a mask for each iris image in form of parameters and a method to extract the mask. The database is partitioned into two datasets based on the shapes used for segmenting the iris and eye lid, the CC and EP dataset. For the CC dataset the parameters define circles which give the iris boundaries and eye lid maskings. For the EP dataset the parameters define ellipses for the iris and polynomials for the eye lid. Note that the eye lid parameterization for both datasets was done in a way to ensure the best possible separation of iris and eye lid in the iris region, i.e. outside the iris region the parameterization is not necessarily accurate. The generation of the CC and EP datasets were done independently, the CC dataset was created at the Halmstad University, the EP dataset was created at the University of Salzburg. Note that the ground truth was

manually generated. The result time and cost requirements prevented the generation of the ground truth for the full content of all databases. However, the *casia4i* (see below) was segmented by two independent operators. This was done in order to facilitate the estimation of level of conformity of the ground truth generated by different operators.

The iris segmentation database is provided without the original eye image databases. In the description of the databases a link to the actual iris database is included to make it easier to obtain them. The dataset description also contains a list of relevant publications for each of the source iris image databases.

A. IRISSEG-CC Dataset

The CC Dataset is available at http://islab.hh.se/mediawiki/index.php/Iris_Segmentation_Groundtruth.

In the CC dataset the iris segmentation was done based on circles. Both the two iris circumferences and the two eyelids are modeled as circles, providing the radius and center of each one. For this purpose, three points of each circle are marked by an operator, which are used to compute the corresponding radius and center. An example of annotated images is shown in Figure 1.

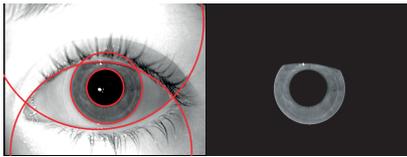


Fig. 1. Sample iris images with circles for both the iris circumferences and the eyelids.

The CC dataset consists of ground truth pertaining to the following iris image datasets.

BioSec Baseline database¹ [8]

The BioSec (*biosec*) database has 3,200 iris images of 640×480 pixels from 200 subjects acquired with a LG IrisAccess EOU3000 close-up infrared iris camera. Here, we use a subset comprising data from 75 subjects (totalling 1,200 iris images), for which iris and eyelids segmentation ground truth is available. Segmentation was done by one operator. The remaining 125 subjects only contain iris segmentation ground truth, i.e. no eyelid segmentation, and therefore they are not used in our experiments (this data, however, is also available in the distribution).

CASIA Iris Image Database version 3.0²

The CASIA-Iris-Interval (*casia3i*) subset of the CASIA v3.0 database, containing 2655 iris images of 320×280 pixels from 249 subjects, was fully segmented. Images were acquired with a close-up infrared iris camera in an indoor environment, having images with very clear iris texture details thanks to a circular NIR LED array. The segmentation of iris images was performed by one operator.

¹<http://atvs.ii.uam.es/databases.jsp>

²<http://biometrics.idealtest.org>

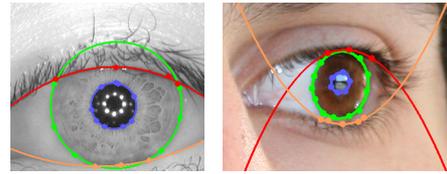


Fig. 2. Sample iris images with datapoints for lower and upper eyelid, inner and outer iris boundaries and the ellipses and polynomials fit to the datapoints.

MobbIO Database³ [6]

The iris (*mobbio*) training subset of the MobbIO database, containing 800 images of 240×200 pixels from 100 subjects, was fully segmented by one operator. Images were captured with the Asus Eee Pad Transformer TE300T Tablet (webcam in visible light) in two different lighting conditions, with variable eye orientations and occlusion levels, resulting in a large variability of acquisition conditions. Distance to the camera was kept constant, however.

B. IRISSEG-EP Dataset

The EP Dataset is available at <http://www.wavelab.at/sources>.

In the EP dataset the iris segmentation was done based on boundary points for the inner and outer iris circumference as well as for the lower and upper eye lids. The dataset provides at least 5 data points for inner and outer iris circumference and the actual boundary is given by an ellipse fitted with a least squares method [9]. For upper and lower eyelid the dataset provides at least 3 data points from which the eyelid boundary can be obtained by fitting a second order polynomial with a least squares method [10]. Figure 2 illustrates the points used for the least squares fit and the resulting boundary ellipses and polynomials.

The EP dataset contains the ground truth for the following iris databases.

CASIA Iris Image Database version 4.0⁴

The CASIA-Iris-Interval (*casia4i*) subset of the CASIA v4.0 database, which contains the same iris images as the *casia3i*. The segmentation of iris images was performed by two separate operators. Consequently two sets of ground truth are available for the iris images in the *casia4i*.

IIT Delhi Iris Database version 1.0⁵ [11], [12]

The IIT Delhi iris database (*iitd*), containing 2240 iris images of 320×240 pixels from 224 subjects, was fully segmented by one operator. The Images were acquired with a JIRIS, JPC1000, digital CMOS camera in the near infrared spectrum. The acquisition was performed in an indoor environment with a frontal view (no off angle).

ND-IRIS-0405 Database⁶ [13]

The acquisition of ground truth for the ND-IRIS-0405 (*notredame*) database was not completed. Of the 64980 iris images, with 640×480, pixels contained in notredame only 837, from 30 different subjects, were segmented. The ground

³<http://paginas.fe.up.pt/~mobbio2013>

⁴<http://biometrics.idealtest.org>

⁵http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm

⁶http://www3.nd.edu/~cvrl/CVRL/Data_Sets.html

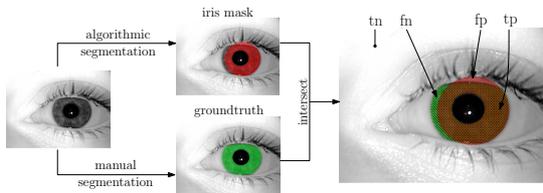


Fig. 3. Example of intersection for calculation of precision and recall.

truth of one operator is available. The image acquisition was done in the near infrared spectrum in an indoor environment with the LG 2200 iris biometric system. The iris images contain samples of 'real-world' type iris images, off-angle, blur, interlacing due to motion, occlusions and iris image cutoff by off center image acquisition.

UBIRIS.v2 Database⁷ [5]

The segmentation of the UBIRIS.v2 database (*ubiris*) was not completed. Of the 11101 iris images, with 400×300 pixels, contained in *ubiris* only 2250 iris images, from 50 different subjects, were segmented by one operator. The images were acquired with a Nikon E5700 camera and split into two sessions. The first session was performed under controlled conditions simulating the enrollment stage. A second session was performed using a real-world type setup, with a natural luminosity and corresponding heterogeneity in reflections, contrast and focus. The database contains frontal and off-angle iris images recorded from various distances with occlusions.

III. EXPERIMENTAL FRAMEWORK

A. Methodology

The goal of a segmentation algorithm is to retrieve the iris region, or iris pixels, from an image. In order to evaluate the performance of the algorithm we will utilize the precision (\mathcal{P}), recall (\mathcal{R}) and F-measure (\mathcal{F}), which are well known measures from the field of information retrieval [14]. We can split the mask produced by an algorithm into four result groups, illustrated in fig. 3: true positives (tp), the number of iris pixels which were correctly marked; false positives (fp), the number of non-iris pixels which were marked; false negative (fn), the number of unmarked iris pixels; and true negative (tn), the number of unmarked non-iris pixel.

The precision defined as

$$\mathcal{P} = \frac{tp}{tp + fp},$$

gives the percentage of retrieved iris pixels which are correct. The recall

$$\mathcal{R} = \frac{tp}{tp + fn},$$

gives the percentage of iris pixels in the ground truth which were correctly retrieved. Since the target is to optimize both recall and precision these two scores are combined by the F-measure, which is the harmonic mean of \mathcal{P} and \mathcal{R} ,

$$\mathcal{F} = \frac{2\mathcal{R}\mathcal{P}}{\mathcal{R} + \mathcal{P}}.$$

The recall, precision and F-measure can concisely describe the segmentation performance of a given algorithm based on ground truth. The measures are good when comparing algorithms. From the equations we can see that recall is a measure for the original iris content retrieved by an algorithm, it can also be maximized by overestimating the iris. Precision on the other hand can be optimized by underestimating the iris and is a measure of the non-iris content of the retrieved iris mask. The F-measure combines precision and recall in a way that will prevent optimization of results by over fitting or under fitting the iris.

While the overall performance over a database is a good performance measure, it is often more useful to know border cases when developing an algorithm. Iris images which are especially good or especially bad show faults and strength of an algorithm better than the overall performance. Especially during development this information can be utilized to further improve and algorithm. In order to find outliers we suggest the following methods.

Precision, recall and the F-measure are calculated for every image I in a given database \mathcal{D} . In order to find outliers we calculate the z-score for a given measure $m \in \{\mathcal{R}, \mathcal{P}, \mathcal{F}\}$ as

$$z(I, \mathcal{D}) = \frac{m(I) - \mu(m_{i \in \mathcal{D}}(i))}{\sigma(m_{i \in \mathcal{D}}(i))}.$$

A given mask is defined as an outlier if $|z| > 3$ ($\mu \pm 3\sigma$). This cannot only be done for individual iris images but also for users, i.e. find users which overall exhibit interesting properties regarding an algorithm. To do this we partition \mathcal{D} into groups \mathcal{G} , where each group contains all iris images of one user. Another possibility is to partition the groups by user and eye id, either left or right. In this case we can also calculate the group outliers, for $G \in \mathcal{G}$, with

$$z(G, \mathcal{G}) = \frac{\mu(m_{i \in G}(i)) - \mu(\mu_{g \in \mathcal{G}}(m_{i \in g}(i)))}{\sigma(\mu_{g \in \mathcal{G}}(m_{i \in g}(i)))}.$$

B. Evaluation

For the evaluation we utilize the following iris segmentation algorithms: CAHT [1], WAHET[15], Osiris[16], IFFP [17] and GST [18]. The algorithms generate mask images for the iris. These images are compared to the ground truth provided in the database as discussed in the previous section.

However, in order to be able to put the performance of the algorithms into context we need a suitable baseline. The notion that a parameterized boundary exactly matches the iris is wrong, it is however preferable to a pixel perfect mask for the subsequent rubbersheet normalisation. The (parameterized) circles, ellipses and polynomials cannot generate a pixel perfect iris mask. Consequently, if an algorithm would generate a perfect pixel mask the F-measure would be lower than 1, since the ground truth is not that exact. However, the addressed segmentation algorithms employ parameterized curves (circles and ellipses) unlikely to generate a pixel perfect iris mask. The question then becomes: is it still plausible to assume an F-measure of 1 to be an attainable maximum.

In order to establish a proper baseline we use the *casia4i* dataset for which we have the ground truth from two different operators. Since the operators independently assessed and marked the iris circumference the resulting iris mask will be

⁷<http://iris.di.ubi.pt/ubiris2.html>



Fig. 4. Sample of iris mask from the *notredame* database which illustrates the over- and underestimation of algorithms.

closely matched but not congruent. The predictive value of one operator to another, given in table I, can be used as a base line towards which algorithms should be judged. From the table it becomes clear that the target of a high performance segmentation algorithm should be a high F-measure as well as low variance over the predictions.

TABLE I. COMPARISON OF THE GROUND TRUTH ACCORDING TO OPERATOR A AND B ON *casia4i* IN ORDER TO ESTABLISH A BASELINE FOR \mathcal{R} , \mathcal{P} AND \mathcal{F} .

Operator	\mathcal{R}		\mathcal{P}		\mathcal{F}	
	μ	σ	μ	σ	μ	σ
A predicts B	96.45%	1.55%	98.66%	1.06%	97.53%	0.73%
B predicts A	98.66%	1.06%	96.45%	1.55%	97.53%	0.73%

The values for the algorithm, grouped by data set, are given in table II. For reasons of brevity only comparisons with one operator are given. When comparing these values to the baseline we can see that occasionally precision or recall reaches baseline levels. However, the standard deviation over any given dataset for the algorithms is significantly higher than the baseline values. Additionally, the F-measure is consistently lower than the baseline, which indicates that while \mathcal{R} or \mathcal{P} reach baseline levels the other is always lower. Basically, IFFP, CAHT and WAHET tend to overestimate the iris boundaries, resulting in higher recall than precision. Osiris on the other hand underestimates, leading to high precision but lower recall. GST includes eye lid separation but in less aggressive variant than Osiris, resulting in a more balanced precision and recall. The reason for this is that IFFP, CAHT and WAHET do not utilize eye lid separation while GST and Osiris try to mask out obstructions, illustrated in fig. 4.

Further, and more specific, analysis can also be done by utilizing outliers to find special cases. A full analysis of the results would exceed the limited space of the paper but additional plots for outlier detection are provided by a technical report [19].

An example of outlier detection is given in figure 5. Note that only the worst 7 eyes are labeled to prevent cluttering of the plots. What can clearly be seen from the plot is the fact that there are users, and images, where the algorithm does perform significantly worse than over the average of the database. Such cases are of importance for the analysis and improvement of segmentation algorithms since they provide border cases.

In fig. 5 certain cases produce an \mathcal{F} of zero, which is an indicator that either the algorithm could not produce a mask or that the mask is completely off. Circles mark cases where the average result of a subject is significantly lower than the average over all users. This could indicate users with eye properties which are problematic. Since no single user exhibits this property for all algorithms we can exclude the case that

TABLE II. THE \mathcal{P} , \mathcal{R} AND \mathcal{F} VALUES FOR THE ALGORITHMIC GENERATED MASKS WHEN COMPARED TO GROUND TRUTH, OPERATOR C FOR *casia3i*, *biosec* AND *mobbio*, OTHERWISE OPERATOR A.

data-base	algo-rithm	\mathcal{R}		\mathcal{P}		\mathcal{F}	
		μ	σ	μ	σ	μ	σ
<i>biosec</i>	CAHT	86.60%	25.45%	77.47%	27.18%	80.49%	26.15%
	GST	92.62%	8.96%	90.03%	7.10%	91.01%	7.07%
	IFFP	91.26%	13.68%	80.68%	15.32%	85.17%	14.37%
	Osiris	89.30%	6.26%	94.62%	4.19%	91.73%	4.18%
	WAHET	94.28%	9.27%	84.47%	10.96%	88.79%	9.55%
<i>casia3i</i>	CAHT	97.05%	4.41%	83.07%	9.47%	89.16%	6.48%
	GST	84.81%	18.48%	91.13%	7.30%	86.44%	12.06%
	IFFP	91.30%	14.49%	83.90%	13.69%	86.89%	13.00%
	Osiris	86.16%	8.27%	92.61%	5.00%	89.02%	5.70%
	WAHET	94.09%	8.84%	85.66%	9.21%	89.03%	8.20%
<i>casia4i</i>	CAHT	97.68%	4.56%	82.89%	9.95%	89.27%	6.67%
	GST	85.19%	18.00%	89.91%	7.37%	86.16%	11.53%
	IFFP	91.74%	14.74%	83.50%	14.26%	86.86%	13.27%
	Osiris	87.32%	7.93%	93.03%	4.95%	89.85%	5.47%
	WAHET	94.72%	9.01%	85.44%	9.67%	89.13%	8.39%
<i>itid</i>	CAHT	96.80%	11.20%	78.87%	13.25%	86.28%	11.39%
	GST	90.06%	16.65%	85.86%	10.46%	86.60%	11.87%
	IFFP	93.92%	10.68%	79.76%	11.42%	85.53%	9.54%
	Osiris	94.04%	6.43%	91.01%	7.61%	92.23%	5.80%
	WAHET	97.43%	8.12%	79.42%	12.41%	87.02%	9.72%
<i>mobbio</i>	CAHT	28.37%	26.84%	20.96%	25.14%	22.15%	24.39%
	GST	42.21%	35.31%	45.79%	31.49%	42.09%	31.57%
	IFFP	50.58%	39.31%	32.50%	32.28%	37.76%	35.10%
	Osiris	20.08%	22.34%	5.67%	10.52%	8.26%	12.85%
	WAHET	44.27%	37.99%	38.52%	34.46%	39.94%	35.60%
<i>notredame</i>	CAHT	91.32%	20.11%	72.71%	18.62%	80.51%	18.57%
	GST	91.07%	12.57%	83.72%	10.83%	86.44%	9.57%
	IFFP	92.62%	9.83%	79.55%	13.93%	85.06%	11.05%
	Osiris	90.00%	6.68%	95.08%	5.23%	92.27%	4.61%
	WAHET	93.65%	12.12%	82.24%	14.03%	86.79%	11.85%
<i>ubiris</i>	CAHT	18.02%	26.04%	11.39%	19.89%	12.55%	20.06%
	GST	42.59%	40.53%	39.91%	34.39%	39.93%	36.55%
	IFFP	44.38%	37.42%	23.71%	26.94%	28.52%	29.68%
	Osiris	26.46%	31.52%	17.29%	28.81%	18.65%	28.47%
	WAHET	27.40%	34.48%	22.76%	31.03%	23.68%	31.90%

the eye images are faulty. This means that the algorithm cannot deal with a certain property of the iris of the given user. This knowledge can be used to improve the algorithm and make it overall more robust.

However, it is not always clear why an algorithm does not perform as intended. The ground truth can only guide us to special cases, but the reason why the algorithm fails is not implicitly clear. An example of this is shown in figure 6 for subject with ID 1148 (*casia4i*). For this user both the CAHT and Osiris have their lowest \mathcal{F} for a certain eye image while performing well for others. Interesting about these samples is the fact that for an iris of similar size the algorithm works well (for each algorithm the second shown iris image is segmented correctly). This precludes the wrong choice of circular boundary parameters. Furthermore, the algorithm did not detect the boundary of the collarette and the rest of the ciliary zone, which at least is distinct. What exactly misdirects the segmentation algorithm can much easier be analysed using these types of errors. However, the refinement of algorithms is beyond the scope of this paper and subject to future detailed investigations.

Another interesting result is the group outlier for subject id 1206 for the WAHET algorithm. The iris images with the mask produced by the WAHET algorithm are shown in table III. The three cases where the iris is closely modelled exhibit an

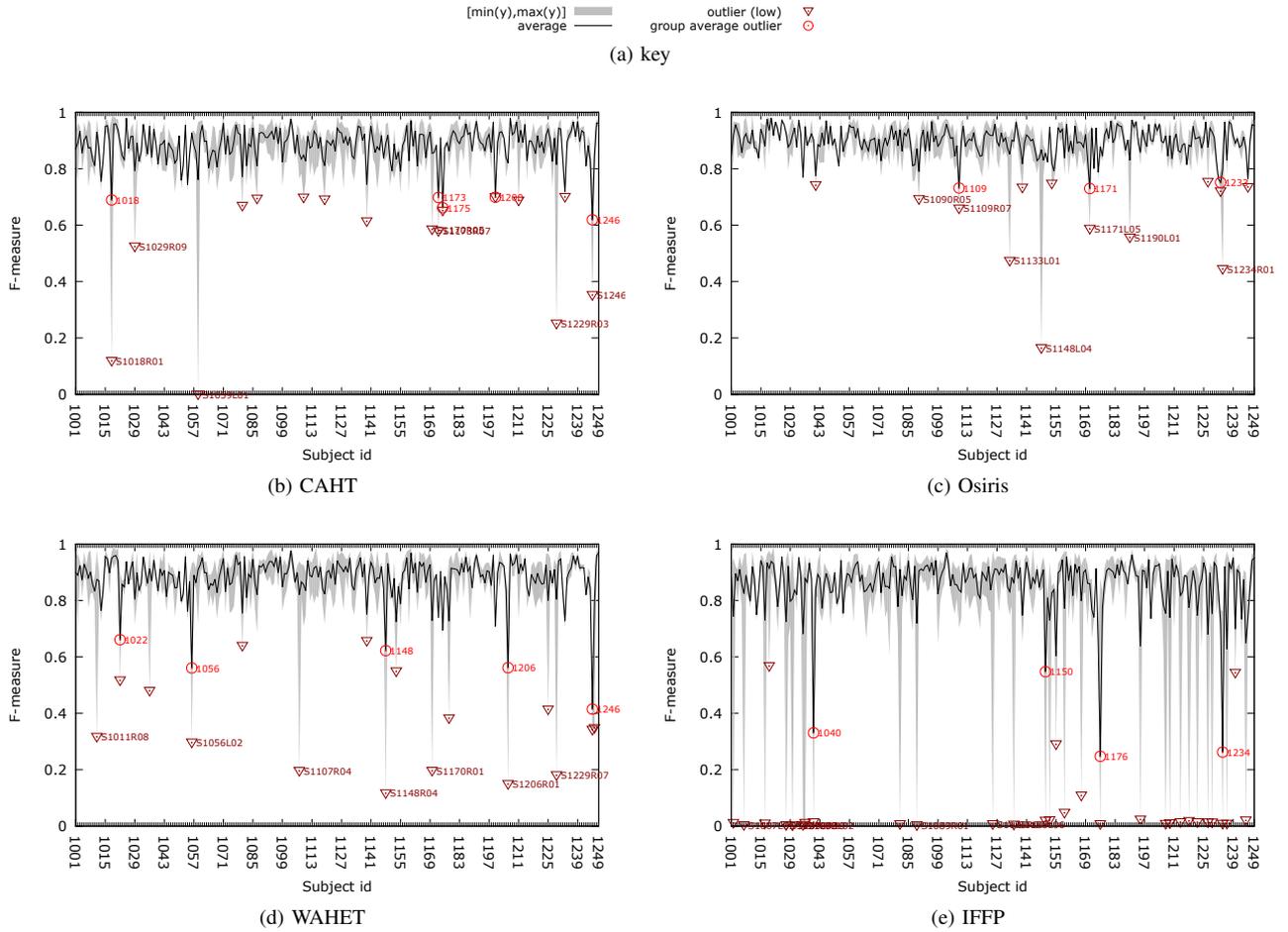


Fig. 5. Outlier detection for group averages and single iris images on the *casia4i* database, grouped by subject. Operator A was used as ground truth for F-measure calculation.

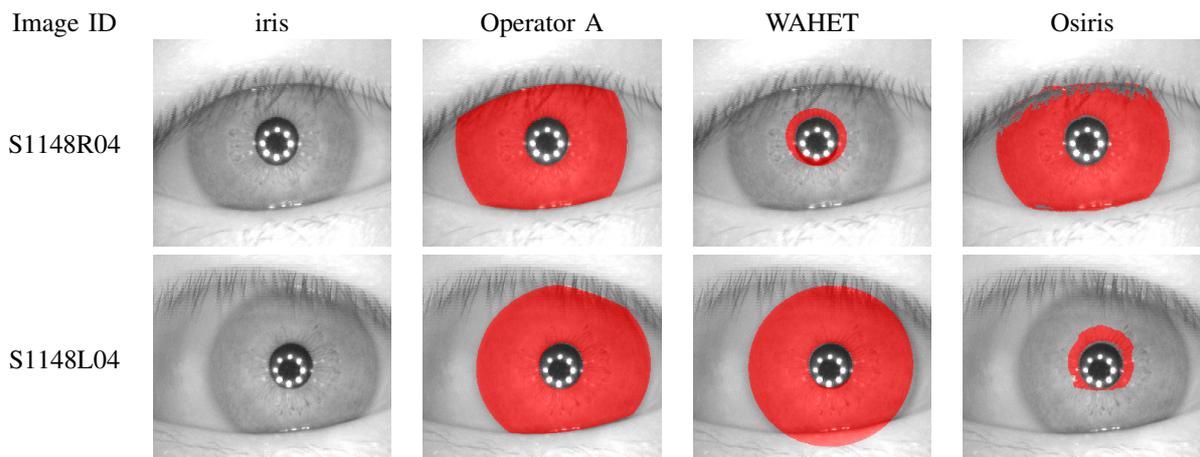


Fig. 6. Iris images for subject ID 1148 (*casia4i*) and mask images from the WAHET and Osiris algorithms.

TABLE III. OUTLIER GROUP FOR SUBJECT 1206 FROM THE *casia4i* DATABASE WITH MASK GENERATED BY THE WAHET ALGORITHM. COMPARISON FOR \mathcal{R} , \mathcal{P} AND \mathcal{F} WITH OPERATOR A.

Image ID	Iris Mask	\mathcal{R}	\mathcal{P}	\mathcal{F}
S1206R01		8.15%	86.84%	14.90%
S1206R02		87.61%	76.68%	81.79%
S1206R03		11.12%	97.96%	19.98%
S1206R04		89.06%	75.13%	81.50%
S1206R05		90.42%	76.25%	82.73%

instance where the circular approximation of iris boundaries is not exact, leading to a lower \mathcal{R} . This coupled with the large occlusion by the upper eye, resulting in the low values for \mathcal{P} , explain the relatively low F-measure. For the other two cases where the detection algorithm detects the collarette boundaries instead of iris boundaries are more curious. The authors assume that the closeness of the upper eye lid to the pupil throws off the boundary detection algorithm leading to the next logical match of the collarette instead of the actual iris boundary.

From the examples given in this section the authors hope to make clear the advantage of using a ground truth for iris segmentation. It provides the option to assess the segmentation results without advancing further into the iris recognition pipeline. The ground truth also makes it easy to identify problematic cases, be it certain users or singular instances of iris images.

IV. CONCLUSION

We have presented a ground truth database for iris segmentation and a methodology for utilizing the ground truth database in order to evaluate iris segmentation algorithm. Furthermore, we applied these methods to evaluate a number of iris segmentation algorithms and show their performance on the various iris image databases. We have presented methods to find special cases where the iris segmentation algorithms do not work as intended. While not in the scope of this paper, this special cases can be used to understand sensitivity with respect to parameter choice and lead to the development of more robust algorithms. The ground truth database will also be made available to the research community.

A. Future work

Future work includes improving a segmentation algorithm based on the ground truth data and evaluating if and how much the performance of the overall iris recognition is improved.

Another topic of research is to take a look at the available ground truth and estimating whether a certain measure is more

important for the overall performance of the iris recognition chain. For example is a high recall or precision more important in terms of EER for iris recognition? Also, it is expected that iris recognition algorithms based on different strategies may not be equally sensible to errors in the segmentation step [7].

REFERENCES

- [1] C. Rathgeb, A. Uhl, and P. Wild, *Iris Recognition: From Segmentation to Template Security*, ser. Advances in Information Security. Springer Verlag, 2013, vol. 59.
- [2] R. P. Wildes, "Iris recognition: an emerging biometric technology," in *Proceedings of the IEEE*, vol. 85, 1997, pp. 1348–1363.
- [3] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.
- [4] H. Proenca and L. Alexandre, "UBIRIS: a noisy iris image database," in *Image Analysis and Processing - ICIAP 2005*, ser. Lecture Notes on Computer Science, F. Roli and S. Vitulano, Eds., vol. 3617, 2005, pp. 970–977.
- [5] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. Alexandre, "The UBIRIS.v2: A database of visible wavelength images captured on-the-move and at-a-distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1529–1535, 2010.
- [6] J. C. Monteiro, H. P. Oliveira, A. Rebelo, and A. F. Sequeira, "Mobio 2013: 1st biometric recognition with portable devices competition," 2013.
- [7] F. Alonso-Fernandez and J. Bigun, "Quality factors affecting iris segmentation and matching," in *Biometrics (ICB), 2013 International Conference on*, 2013, pp. 1–6.
- [8] J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, and J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, no. 4, pp. 1389–1392, 2007.
- [9] A. Fitzgibbon and R. Fisher, "A buyer's guide to conic fitting," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 1995, pp. 51.1–51.10.
- [10] K. Atkinson, *An Introduction to Numerical Analysis*. John Wiley and Sons, 1989.
- [11] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal identification," in *Proceedings of the CVPR 2008*, 2008, pp. 21–27.
- [12] —, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern Recognition*, vol. 43, no. 3, pp. 1016–1026, 2010.
- [13] P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe, "FRVT 2006 and ICE 2006 large-scale experimental results," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 1–1, 2010.
- [14] C. J. van Rijsbergen, *Information retrieval*, 2nd ed. Butterworths, 1979.
- [15] A. Uhl and P. Wild, "Weighted adaptive hough and ellipsoidal transforms for real-time iris segmentation," in *Proceedings of the 5th IAPR/IEEE International Conference on Biometrics (ICB'12)*, 2012, pp. 1–8.
- [16] D. Petrovska and A. Mayoue, "Description and documentation of the biosecure software library," Project No IST-2002-507634 - BioSecure, Deliverable, 2007.
- [17] A. Uhl and P. Wild, "Multi-stage visible wavelength and near infrared iris segmentation framework," in *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR'12)*, ser. LNCS, 2012, pp. 1–10.
- [18] F. Alonso-Fernandez and J. Bigun, "Iris boundaries segmentation using the generalized structure tensor. an study on the effects of image degradation," in *Proceedings of the IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2012.
- [19] F. Alonso-Fernandez, J. Bigun, H. Hofbauer, A. Uhl, and P. Wild, "Evaluation of the irisseg datasets," University of Salzburg, Department of Computer Sciences, Technical Report: IRISSEG–2013–12, 2013. [Online]. Available: <http://www.cosy.sbg.ac.at/tr>