Fusion of Iris Segmentation Results

Andreas Uhl and Peter Wild

Multimedia Signal Processing and Security Lab.
Department of Computer Sciences, University of Salzburg, Austria
{uhl,pwild}@cosy.sbg.ac.at

Abstract. While combining more than one biometric sample, recognition algorithm, modality or sensor, commonly referred to as multi-biometrics, is common practice to improve accuracy of biometric systems, fusion at segmentation level has so far been neglected in literature. This paper introduces the concept of multi-segmentation fusion for combining independent iris segmentation results. Fusion at segmentation level is useful to (1) obtain more robust recognition rates compared to single segmentation; (2) avoid additional storage requirements compared to feature-level fusion, and (3) save processing time compared to employing parallel chains of feature-extractor dependent segmentation. As proof of concept, manually labeled segmentation results are combined using the proposed technique and shown to increase recognition accuracy for representative algorithms on the well-known CASIA-V4-Interval dataset.

1 Introduction

Aiming to bridge the performance gap of image-based biometric systems between highly accurate standardized cooperative applications and less constrained scenarios has attracted many researchers to propose algorithms improving preprocessing and segmentation techniques, which are reported to play an important role due to susceptibility to poor image quality [10]. The human iris is one of the most unique biometric identifiers, and also selected to be one of two modalities to be employed in the world's largest biometric deployment, Aadhaar, targeting biometric identification of each Indian citizen. It is clear, that such large-scale applications demand high accuracy to avoid misclassification. Furthermore, the discrepancy between users aware of the acquisition and the observed decreased rate when applied in unconstrained scenarios with reported VR (verification rate) as low as 44.6% [14] versus >99% VR at 0.1% FAR (false acceptance rate) for a series of iris biometric systems in constrained environments [1] support the claimed need for higher accuracy in less constrained applications.

A combination of multiple biometric information can increase accuracy at the cost of additional resources and is traditionally employed at the score or decision-level [15]. Such fusion rules unfortunately exhibit limitations: (1) many algorithms conduct the same or similar costly processing steps; (2) segmentation errors propagate along the biometric processing chain, and; (3) contradicting information may derogate system performance. This leads to the question: *Can*

J. Ruiz-Shulcloper and G. Sanniti di Baja (Eds.): CIARP 2013, Part II, LNCS 8259, pp. 310–317, 2013.

fusion at lower levels (segmentation) lead to more accurate (and faster) biometric systems? This paper is dedicated to providing a positive answer to the feasibility of fusion at segmentation stage, i.e. whether the combination of independent segmentation results lead to better system performance in terms of recognition accuracy, independent of the chosen feature extraction and comparison algorithm. Note, that the choice of methods may impact processing time.

The remainder of this paper is organized as follows: Section 2 reviews related work with respect to multi-biometric fusion. Section 3 formalizes the referred segmentation model and introduces the concept of fusion at segmentation stage. Section 4 introduces experimental setup and Section 5 presents an experimental evaluation of the proposed technique. Finally, Section 6 concludes this work.

2 Multibiometric Fusion

Multibiometric fusion refers to the "use of multiple pieces of evidence in order to deduce or verify human identity" [18] and can be applied at different stages in the biometric processing chain [15]:

1. *Data/Feature Level*: consolidating information from the raw biometric signal or after feature extraction from individual classifiers into a single high-dimensional template.
2. *Score Level*: consolidating comparison scores with density-based (using the likelihood ratio after modeling genuine and imposter score distributions), transformation-based and classifier-fusion-based (learning boundaries from observed data) solutions, this is probably the most-intensively studied type of fusion leaving other processing modules unaffected.
3. *Rank/Decision Level*: depending on whether biometric authentication is performed in identification mode (1-to-N comparison with all subjects registered with the system to determine an identity from a biometric sample) or verification mode (1-to-1 comparison to justify the authenticity of an identification claim as *genuine* or *imposter*), this fusion type consolidates the outcome of individual decision processes, i.e. ranking lists or class decisions.

Due to the development of embedded solutions and with the rise of new biometric modalities focusing on specific parts and/or scales, the original classification of fusion scenarios in [15] into (1) multiple sensors, (2) multiple biometrics, (3) multiple units, (4) multiple snapshots, (5) multiple matchers is less strict and new scenarios emerge [16]. While an integration at early level is claimed to be more effective [15], it is more complex to design. The majority of proposed multibiometric techniques targeting biometric surveillance (e.g., [19,14,5]) are score-level fusion approaches. Only few data/feature level fusion techniques exist: [4] is the first signal-level fusion approach in iris recognition creating a single high-resolution image from multiple frames in video outperforming score-level fusion techniques. Their proposed technique is essentially an image fusion of iris images at the pixel level. Our approach is different in targeting not multiple snapshots but a single-snapshot only and combining the result of multiple segmentation

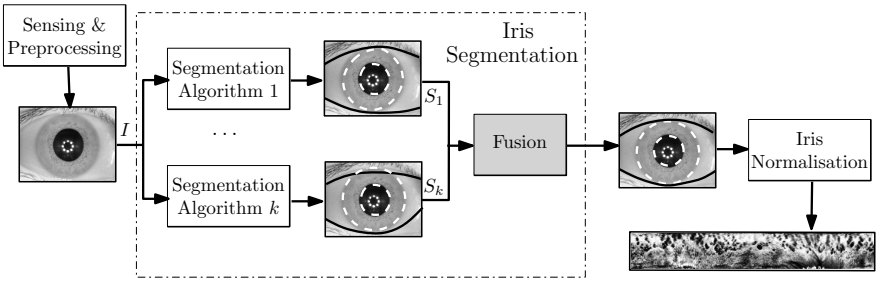

Fig. 1. Basic operation mode of novel proposed iris segmentation fusion

results in order to improve recognition accuracy. This is the first approach on combining segmentation results for improved iris recognition.

3 Iris Segmentation Fusion

This paper proposes to allow for a combination of multiple segmentation results $S_1, S_2, \ldots S_k$ of the same input iris image I using multiple segmentation algorithms, see Fig 1 for an illustration on how iris segmentation fusion can be integrated into iris processing chains between sensing and normalization. Since not all iris feature extraction techniques require the same preprocessing tasks, the proposed fusion technique uses segmentation results by employing Daugman's normalization [2], which serves as the basis for most commercial applications [12]. A good reference work for practices on image segmentation classifier combination is [6].

3.1 Daugman's Iris Normalization Model

In Daugman's algorithm [2], binary features are extracted after mapping iris texture between inner pupillary and outer limbic boundary into a representation called "Faberge coordinates" applying a rubbersheet transform, see Fig. 2. This process involves essentially two tasks [12]: (1) iris segmentation detecting the two (originally circular, but extensible to arbitrarily shaped) boundaries, pupillary and limbic polar curves $P, L : [0, 2\pi) \rightarrow [0, m] \times [0, n]$, for the eye instance in the $m \times n$ input image (we assume, that eye detection and quality checks indicate exactly one such instance is present and of sufficient quality); and (2) iris normalization, which creates a normalized representation of the iris texture, invariant under pupillary dilation and facilitating for rotational alignment via simple pixel-shifts using angular θ and pupil-to-limbic radial r coordinates:

$$R : [0, 2\pi) \times [0, 1] \rightarrow [0, m] \times [0, n]. \quad R(\theta, r) := (1 - r) \cdot P(\theta) + r \cdot L(\theta). \quad (1)$$

Besides the mapping in doubly dimensionless coordinates using R, due to eyelids and reflections, the resulting rectangular area does not only contain iris texture, but also areas, which should be masked out during feature extraction

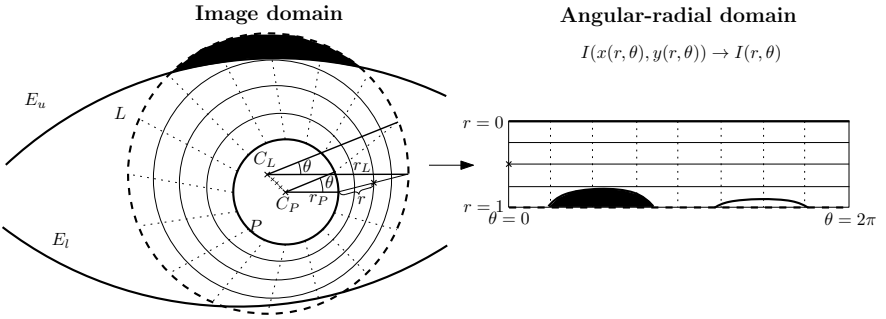

Fig. 2. Iris rubbersheet transform model with circular P, L and paraboloid E_u, E_l

and comparison. While [13] shows, that indeed, a reordering of pixels based on reliability has almost a similar effect like noise masks (and render their use less effective), traditional processing also creates a binary noise mask as part of the normalization task, $N : [0, 2\pi) \times [0, 1] \to \{0, 1\}$, marking areas occluded by eyelids, eyelashes or reflections. Usually, in order to build this noise mask, upper and lower eyelids are fitted by paraboloid or polynomial curves $E_u, E_l : [0, 1] \to [0, m] \times [0, n]$ to mask out occluded areas in the noise mask.

3.2 Combination of Segmentation Results

Motivated by the observation, that more generic alignment using Levenshtein instead of Hamming distance (HD) is able to increase recognition [17], the goal of the fusion module is to obtain a better pupillary and limbic boundary representation for minimizing the effect of mapping deformations due to inaccurately localized boundaries in the rubbersheet transform. While there are several different possibilities to accomplish this task (e.g., for practices on image segmentation classifier combination see [6]), we exemplary introduce two different techniques:

- **Sum-Rule Interpolation**: A very natural choice of a fusion rule combining multiple boundaries $B_1, B_2, \ldots B_k : [0, 2\pi) \to [0, m] \times [0, n]$ into a single boundary B is, in analogy to the sum rule in score-level-fusion, the arithmetic mean of sampled boundaries:

$$Sum\ Rule : B(\theta) := \frac{1}{k} \sum_{i=1}^{k} B_i(\theta) \tag{2}$$

  This interpolation is executed for $B = P$ and $B = L$ separately. The same method can be applied to interpolate between upper and lower eyelid approximations E_u, E_l to derive a common noisemask.
- **Augmented-Model Interpolation**: in case boundaries are rather different and/or the curves' sampling interval $[0, 2\pi]$ is not "equally spaced", i.e. for discretized equidistant samples of arguments $x_1, \ldots x_s \in [0, 2\pi]$ the boundary polygon $B(x_1), B(x_2), \ldots B(x_s)$ has large variation in the length of boundary line segments, sum rule interpolation may lead to inaccurate results. While in

this case a re-parametrization of boundary curves may be useful or necessary for sum-rule interpolation, an alternative approach to the fusion of boundary curves is fitting a model to the union of sampled edge points:

$$Aug\ Rule : B(\theta) := ModelFit\Big(\bigcup_{1\leq i\leq k}\bigcup_{1\leq j\leq s} B_i(x_j)\Big)(\theta) \tag{3}$$

where *ModelFit* is a fitting routine taking a set of points and providing a suitable shape (closed boundary curve) minimizing a model-error, e.g. Fitzgibbon's ellipse fitting method [3] in case of $B = P$ or $B = L$. For upper and lower eyelid curves $B = E_u, B = E_l$, input points can be used to fit a polygon of second order to the input points.

4 Experimental Setup

In order to estimate the usability of the proposed new fusion framework, we assess its performance on manually segmented iris images. This test is useful, since (1) any dependencies between segmentation algorithms can be avoided in this case enabling a fair test of the fusion rule, (2) a positive outcome justifies its application in building high-confidence fused ground truth for evaluating segmentation algorithms, (3) manual segmentations are state-of-the-art (e.g. in the Noisy Iris Challenge Evaluation [11]) to evaluate segmentation techniques (i.e. considered superior to automated evaluations), therefore if segmentation fusion is able to improve manual segmentation, it is a positive result for also automated segmentation techniques, which are continuously improved to achieve close-to-manual performance.

For experiments we employ the entire *CASIA-V4-Interval*[1] dataset of high quality NIR illuminated indoor images with 320×280 pixel resolution (2639 images, 395 classes). For manual segmentations, a male (*Manual 1*) and female (*Manual 2*) expert manually labeled boundary points until the fitted elliptic inner pupillary and outer limbic boundaries sufficiently (according to the opinion of the expert) approximated the true possibly occluded iris boundary. The same procedure was also executed for upper and lower boundaries using a polynomial of order two as the curves' model. During manual segmentation, the expert could zoom in/out and see the original and fitted (segmented) image.

As feature extraction algorithms operating on normalized iris textures, three representative implementations available in USIT[2] were employed: *Masek* [8] is a feature extraction algorithm extracting phase angles from the row-wise convolution of the 1D intensity signals with scaled and oriented Log-Gabor kernels encoding each phase angle with 2 bits leading to a 10240 bits code. Fractional HD is employed for comparison. *Ma* [7] is a feature extraction algorithm tracking the positions of minima and maxima (switching bit sequences) after executing 1D

[1] The Center of Biometrics and Security Research, CASIA Iris Image Database, `http://biometrics.idealtest.org`

[2] University of Salzburg Iris Toolbox, `http://wavelab.at/sources/USIT/`

wavelet transform on the 10 one-dimensional signals, each one averaged from the pixels of 5 adjacent rows for each of two subbands. Again, fractional HD is the comparison criterion. Finally, *Monro* [9] employs a 1D discrete cosine transform (DCT) on diamond-shaped image patches using a Hanning window approach to locally summarize data. The final 2x128 bytes code tracks zero crossings of the differences between the DCT coefficients of adjacent patch vectors using first three DCT coefficients for a total of 7 shift positions $(0; \pm 4; \pm 8; \pm 12)$. Also for the other codes (*Masek, Ma*) the comparison routine employed 7 bit shifts in either direction for optimal alignment.

5 Results

We evaluate segmentation accuracy by assessing the impact on verification recognition accuracy, i.e. ROC curves plotting false acceptance rate (FAR) versus genuine acceptance rate (GAR), given in Figs. 3, 4 and 5. GAR at fixed FAR ($\leq 0.01\%$) for each of the two manual segmentations as well as fused results are reported in Table 1.

First, we can see that independent of the employed feature extraction algorithm, both manual segmentations exhibit the same order in performance over the entire operational ROC range: manual segmentation 2 delivers more accurate results with 97.64% GAR at FAR $\leq 0.01\%$ for Masek, 98.34% for Ma and 95.72% for DCT-based Monro versus 97.46% for Masek, 98.19% for Ma and 93.94% Monro in case of the first manual segmentation. This suggests, that manual segmentation 2 is more accurate/consistent. Both segmentations needed approximately 9 working days to segment the dataset.

The second important observation is an algorithm-dependent impact of segmentation on recognition accuracy. While typically, algorithms are compared by using their own segmentation technique, we can see that the sensitivity against segmentation among algorithms is quite different and should be considered when comparing algorithms. While for Masek performance differences are almost invisible (1.17% EER Manual 1 vs. 1.15% EER Manual 2, but still better performance for segmentation fused Sum Rule with 1.13% EER and Aug-Rule with 1.12% EER), differences for Monro are clearly present (1.84% EER Manual 1 vs. 1.62% EER Manual 2, vs. Sum Rule with 1.52% EER and Aug-Rule with 1.48% EER).

Third, with respect to the targeted feasibility study of segmentation fusion we can report, that fusion algorithms were able to increase accuracy of both segmentation results, independent of the chosen feature extraction algorithm - a result which is not self-evident and justifies its future investigation with existing segmentation algorithms. Sum Rule Interpolation, which has the advantage of being fast in computing an averaged segmentation representation, could increase GAR from 97.46% to 97.84% for Masek, from 98.19 to 98.57% for Ma, and from 93.94% to 96.74% for the Monro implementation, which did not consider noise masks. Relative performance differences to the Augmented Model Interpolation were insignificant (97.84% GAR for Masek, 98.51% Ma and 96.8% for Monro), i.e. both fusion rules performed almost equally well.

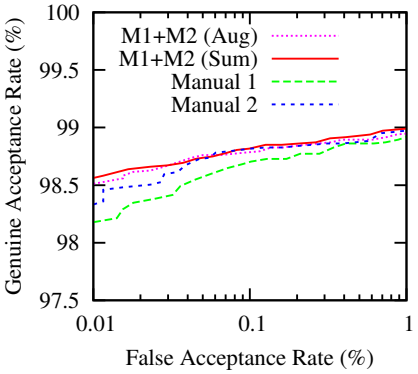

Fig. 3. ROC for the *Ma* feature

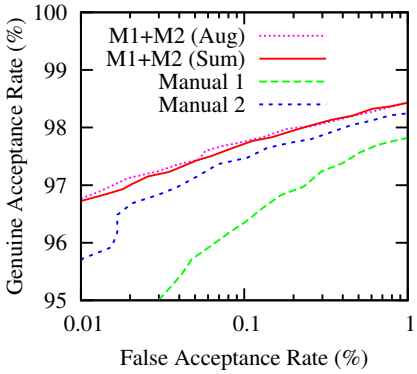

Fig. 4. ROC for the *Monro* feature

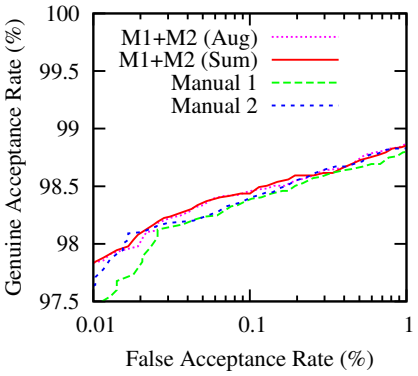

Fig. 5. ROC for the *Masek* feature

Table 1. Recognition accuracy in Genuine Acceptance Rate at $\leq 0.01\%$ False Acceptance Rate.

Algorithm	GAR at FAR $\leq 0.01\%$		
	Masek	Ma	Monro
Manual 1	97.46	98.19	93.94
Manual 2	97.64	98.34	95.72
Sum-Rule	**97.84**	**98.57**	**96.74**
Aug-Model	**97.84**	**98.51**	**96.80**

6 Conclusion

Recent challenges like the Noisy Iris Challenge Evaluation (NICE) or Multiple Biometrics Grand Challenge (MBGC) have put a strong focus on the segmentation problem of challenging iris images. But so far, there has been no systematic framework of combining segmentation results from different algorithms. We showed, that besides combining outcomes of biometric feature extraction or comparison algorithms, it may be useful to combine segmentation and normalization information from multiple sources. Evaluations using manual segmentation on CASIA-V4-Interval revealed improvement by segmentation fusion for each of the employed feature extraction algorithms and fusion rules. Segmentation-fused recognition was as high as 96.8% GAR at $\leq 0.01\%$ FAR (Aug-Rule) vs. 93.94% and 95.72% for individual segmentations in case of Monro's feature. Future work will focus on automatic segmentation algorithms, more challenging datasets, and quality-related information assisting fusion rule selection.

References

1. Bowyer, K.W., Hollingsworth, K.P., Flynn, P.J.: A survey of iris biometrics research: 2008-2010. In: Burge, M., Bowyer, K. (eds.) Handbook of Iris Recognition, pp. 15–54. Springer, New York (2012)
2. Daugman, J.: How iris recognition works. IEEE Trans. on Circiuts and Systems for Video Technology 14(1), 21–30 (2004)
3. Fitzgibbon, A., Pilu, M., Fisher, R.B.: Direct least square fitting of ellipses. IEEE Trans. Pattern Anal. Mach. Intell. 21(5), 476–480 (1999)
4. Hollingsworth, K., Peters, T., Bowyer, K., Flynn, P.: Iris recognition using signal-level fusion of frames from video. IEEE Trans. Inf. Forensics and Sec. 4(4), 837–848 (2009)
5. Jillela, R., Ross, A.: Mitigating effects of plastic surgery: Fusing face and ocular biometrics. In: Proc. 5th IEEE Int'l Conf. on Biometrics: Theory, App. and Syst., pp. 402–411 (2012)
6. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience, Hoboken (2004)
7. Ma, L., Tan, T., Wang, Y., Zhang, D.: Efficient iris recognition by characterizing key local variations. IEEE Transactions on Image Processing 13(6), 739–750 (2004)
8. Masek, L., Kovesi, P.: MATLAB Source Code for a Biometric Identification System Based on Iris Patterns (2003), `http://www.csse.uwa.edu.au/~pk/studentprojects/libor/sourcecode.html`
9. Monro, D.M., Rakshit, S., Zhang, D.: Dct-based iris recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(4), 586–595 (2007)
10. Proença, H., Alexandre, L.A.: Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. Image and Vision Comp. 28(1), 202–206 (2010)
11. Proença, H., Alexandre, L.: Toward covert iris biometric recognition: Experimental results from the nice contests. IEEE Trans. Inf. Forensics and Sec. 7(2), 798–808 (2012)
12. Rathgeb, C., Uhl, A., Wild, P.: Iris Recognition: From Segmentation to Template Security. Advances in Information Security, vol. 59. Springer, New York (2012)
13. Rathgeb, C., Uhl, A., Wild, P.: Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. In: Proc. Int'l Conf. on Biometrics: Theory App. and Syst., pp. 1–6. IEEE (2010)
14. Ross, A., Jillela, R., Smereka, J., Boddeti, V., Kumar, B., Barnard, R., Hu, X., Pauca, P., Plemmons, R.: Matching highly non-ideal ocular images: An information fusion approach. In: Proc. 5th Int'l Conf. on Biometrics, pp. 446–453 (2012)
15. Ross, A.A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. Springer, Secaucus (2006)
16. Uhl, A., Wild, P.: Single-sensor multi-instance fingerprint and eigenfinger recognition using (weighted) score combination methods. Int'l Journal on Biometrics 1(4), 442–462 (2009)
17. Uhl, A., Wild, P.: Enhancing iris matching using Levenshtein distance with alignment constraints. In: Bebis, G., et al. (eds.) ISVC 2010, Part I. LNCS, vol. 6453, pp. 469–478. Springer, Heidelberg (2010)
18. Vatsa, M., Singh, R., Noore, A., Ross, A.: On the dynamic selection of biometric fusion algorithms. IEEE Trans. on Inf. Forensics and Security 5(3), 470–479 (2010)
19. Woodard, D., Pundlik, S., Miller, P., Jillela, R., Ross, A.: On the fusion of periocular and iris biometrics in non-ideal imagery. In: Proc. 20th Int'l Conf. on Pattern Recognition, pp. 201–204 (2010)