

Andreas Uhl and Peter Wild, “Enhancing Iris Matching Using Levenshtein Distance with Alignment Constraints”, In G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. Chung and R. Hammoud, editors, *Advances in Visual Computing: 6th International Symposium, (ISVC 2010)*, pp. 469–479, LNCS, 6453, Springer Verlag, 2010

© Springer Verlag. The copyright for this contribution is held by Springer Verlag. The original publication is available at www.springerlink.com.

<http://www.springerlink.com/content/k85264341nqq6274/>

Enhancing Iris Matching Using Levenshtein Distance with Alignment Constraints

Andreas Uhl and Peter Wild

Department of Computer Sciences,
University of Salzburg, A-5020 Salzburg, Austria

Abstract. Iris recognition from surveillance-type imagery is an active research topic in biometrics. However, iris identification in unconstrained conditions raises many problems related to localization and alignment, and typically leads to degraded recognition rates. While development has mainly focused on more robust preprocessing, this work highlights the possibility to account for distortions at matching stage. We propose a constrained version of the Levenshtein Distance (LD) for matching of binary iris-codes as an alternative to the widely accepted Hamming Distance (HD) to account for iris texture distortions by e.g. segmentation errors or pupil dilation. Constrained LD will be shown to outperform HD-based matching on CASIA (third version) and ICE (2005 edition) datasets. By introducing LD alignment constraints, the matching problem can be solved in $O(n \cdot s)$ time and $O(n + s)$ space with n and s being the number of bits and shifts, respectively.

1 Introduction

Unconstrained iris recognition is a relatively new branch in iris-based identification. It is driven by the demands to push biometric image acquisition towards an extraction of biometric signals with the subject of interest moving or being at-a-distance from biometric sensors. Advantages of such systems comprise better usability, higher throughput, and the ability to acquire biometric measurements without required cooperation. While iris recognition in reasonably constrained environments provides high confidence authentication with equal error rates (EERs) of less than 1% [1], a reduction of constraints is quite challenging. First-generation prototype iris identification systems designed for stand-off video-based iris recognition, e.g. Sarnoff's Iris-on-the-move [2], or General Electric's Stand-off Iris Recognition system [3], have proven the feasibility of iris recognition from surveillance-type imagery. But also the need for better segmentation techniques than usually applied in still-image iris recognition to account for distortions like motion blur, defocusing or off-axis gaze direction has been identified as a main issue. Challenges like the Iris Challenge Evaluation (ICE) and Multiple Biometric Grand Challenge (MBGC) have provided standardized datasets to aid in finding solutions to these problems.

Most publications regarding iris recognition in unconstrained environments aim at more sophisticated preprocessing techniques to successfully localize and

segment images of the human eye. Proença et al. [4] identify the critical role of segmentation and observe a strong relationship between translational segmentation inaccuracies and recognition error rates. Matey et al. [5]. assess the effect of resolution, wavelength, occlusion and gaze as the most important factors for incorrect segmentation and give a survey of segmentation algorithms. While iris boundaries have been modeled as circles, ellipses and more complex shapes, still whatever model is used, the processing chain of almost all iris recognition algorithms resembles Daugman’s standard approach [6] very close: After successful determination of the inner and outer pupil centers, the iris-ring texture of a person’s eye is unwrapped and further processed by feature extraction modules. Refinements of this model usually refer to more sophisticated generation of noise-masks determining pixels containing eyelashes, eyelids, or other types of distortions. The majority of feature extraction approaches extracts binary output (iris-codes) from the obtained normalized textures [1], and employs the fractional Hamming Distance (HD) over different bit shifts (to account for rotational alignment) between iris-codes in order to determine a degree of similarity at matching stage. Indeed very few studies have proposed new or compared different binary similarity and distance measures, and it is common agreement, that HD is the best method for this task.

Similarity measure selection is a problem encountered in various fields. Cha et al. [7] compare several binary vector similarity measures including a new variable credit similarity measure (altering credit for zero-zero and one-one matches) for iris biometric authentication. Their proposed metric improved and generalized HD measures by introducing weights in order to give greater importance to error pixels in a neighborhood of error pixels. However, in order to determine parameters, a separate training stage is needed and the trained contributing factor was reported to vary considerably depending on the application data. A more exhaustive hierarchically clustered summary of binary similarity and distance measures can be found in Choi et al. [8].

In this paper, we propose the use of the Levenshtein Distance (LD) [9] for iris matching in order to tolerate segmentation inaccuracies and distortions caused by the linearity of Daugman’s normalization model [6]: Each iris point is mapped into a doubly dimensionless coordinate system to account for elastic deformation of the iris texture. Yuan et al. [10] identified drawbacks of this iris normalization model and confirm claims in [11] that a linear stretching seems not enough to catch the complex nature of pupil dilation. The novelty of our approach lies in considering deformation at the matching - and not at normalization - stage. We perceive a significant improvement of recognition accuracy using this method compared to traditional HD at only moderate additional time overhead. Improvements to common implementations of LD computation, i.e. alignment constraints, are suggested and evaluated for two different open iris databases with respect to recognition accuracy and time complexity. As a by-product of our evaluation, we also highlight the impact of shifts on iris recognition accuracy.

The remainder of this paper is structured as follows: An introduction to LD and its application in iris recognition is given in Sect. 2. Experimental setup,

employed feature extraction algorithms and datasets are presented in Sect. 3. Evaluations are discussed in Sect. 4 and Sect. 5 concludes this work.

2 Levenshtein Distance in Iris Recognition

The Levenshtein Distance dates back to the 1960s [9] and is a well-known classical distance measure in pattern recognition for sequences of possibly different length. The idea to use LD for biometrics is not new, e.g. Schimke et al. [12] employ an adapted version of LD based on event-string modelling and a nearest neighbor classifier for online signature verification. In this work, we assess the usability of LD for iris recognition in order to cope with imperfect normalization of iris textures. We employ LD at the matching stage to enhance iris recognition accuracy, as outlined in Fig. 1.

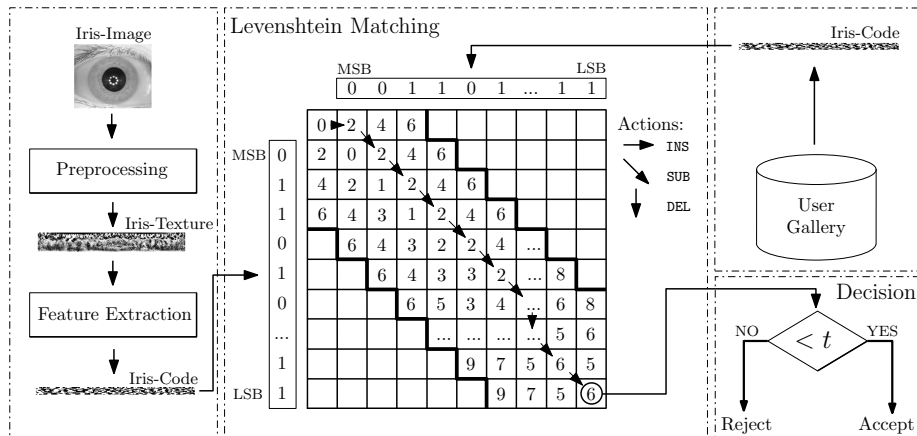


Fig. 1. System Architecture: the basic operation mode of the proposed system.

The inherent idea of LD is to employ inexact matching allowing a sequence to exhibit additional or lack parts of another similar sequence. Similarity is defined by an optimal transformation of one sequence into the other by three operations: *INS* (insert), *DEL* (delete) and *SUB* (substitute). Each operation is associated with a cost, c_{INS} and c_{DEL} are scalar values, $c_{SUB}(a, b)$ is a function depending on the symbols $a, b \in \{0, 1\}$ at specific positions of the two sequences to be compared (in iris recognition we consider binary sequences only). Typically, $c_{SUB}(a, b) = 0 \Leftrightarrow a = b$ for symbol b replacing a . The LD is also called Edit Distance and can be calculated by Dynamic Time Warping (DTW) [13], a dynamic programming algorithm to align the sequences. Let $A \in \{0, 1\}^m$ and $B \in \{0, 1\}^n$ be binary sequences of length m and n (as we apply LD to iris-codes of fixed length, typically $m = n$ holds). DTW uses a matrix D of size $(m + 1) \times (n + 1)$,

which is incrementally computed:

$$D[0, 0] := 0 \tag{1}$$

$$\forall i > 0 : D[i, 0] := D[i - 1, 0] + c_{DEL}, \tag{2}$$

$$\forall j > 0 : D[0, j] := D[0, j - 1] + c_{INS}, \tag{3}$$

$$\begin{aligned} \forall i, j > 0 : D[i, j] := \min(&D[i - 1, j] + c_{DEL}, \\ &D[i, j - 1] + c_{INS}, \\ &D[i - 1, j - 1] + c_{SUB}(A[i], B[j])). \end{aligned} \tag{4}$$

The invariant maintained throughout the algorithm is, that the subsequence $A[1..i]$ can be transformed into $B[1..j]$ using a minimum of $D[i, j]$ operations. The Levenshtein Distance of A and B is $LD(A, B) := D(m, n)$. A traceback algorithm is used to find the optimal alignment, i.e. an alignment path results from joining nodes and depending on the local direction an optimal (not necessary unique) sequence of transformation operations is derived, see Fig. 1. However, as the matching stage needs the distance measure only, we can avoid a storage of the entire matrix and reduce space complexity from $O(n^2)$ (assuming $n = m$ in our application domain) to $O(n)$ with column-wise computation. Still, time complexity stays at $O(n^2)$ with this modification, which is not useful for commercial applications. The traditional HD dissimilarity measure needs $O(n \cdot s)$ time and $O(n + s)$ space with s being the number of bit shifts, which is usually a very small constant number.

In order to further reduce LD time requirements, we define additional constraints: Computations are restricted to an evaluation region S on matrix D shaped as a stripe from top-left to bottom-right (see Fig. 1), i.e. technically we define $S := \{(x, y) : |x - y| \leq s\}$ and set:

$$\forall (i, j) \notin S : D[i, j] = 0. \tag{5}$$

By this modification, we enforce a maximum local deviation of the patterns by s shifts. Like in the previously refined implementation, we keep track of the last column only. The resulting final algorithm solves the matching problem in $O(n \cdot s)$ time and $O(n + s)$ space with n and s being the number of bits and shifts, respectively. It is worth to notice, that the last optimization does no longer deliver the exact Levenshtein Distance. However, it is a very natural constraint to give an upper limit on relative shifts, since it should not be too difficult to get estimators for eye tilt (e.g. localization of eye lids).

Finally, in order to obtain a normalized distance value, we note that for $m = n$ the condition $LD(A, B) \leq HD(A, B)$ holds, and we therefore divide the result by the code size n . Note, that for LD the triangle inequality holds. The advantage of LD over traditional HD lies in its ability of non-linear alignment. For this reason, it is widely accepted in computational biology, e.g. to estimate alignments of DNA.

3 Experimental Setup

In order to test the performance of the proposed iris matching technique, we employ existing iris recognition systems and replace the iris matching module by the implementation outlined in Sect. 2, see Fig. 1. The transparent application in existing iris biometric solutions (no re-enrollment of the user gallery is necessary) is a key advantage of the proposed approach. All used system components as well as employed biometric databases are described in more detail as follows.

3.1 Databases

For experimental evaluation we employ two different iris databases:

- *CASIA*: we select all 1307 left eyes out of a collection of 2655 NIR illuminated indoor images with 320×280 pixel resolution of the open CASIA-V3-Interval¹ iris database. This dataset reflects performance on high quality input.
- *ICE*: this dataset contains all 1425 right eyes out of 2953 NIR illuminated images with 640×480 pixel resolution of the open ICE-2005² iris database. We selected this test set for lower quality input, as some images have noticeable blur and occlusions.

In case of the *CASIA* dataset, 4028 genuine (intra-class) and 15576 imposter (inter-class) comparisons were executed, for the *ICE* database results refer to 12214 genuine and 7626 imposter comparisons, i.e. we matched all genuine pairs, but only the first template between users.

3.2 Normalization

Depending on the type of input data, we employ two different iris normalization techniques. For the processing of *CASIA* images, we use a custom normalization software applying Canny edge detection and Hough circle detection to localize inner and outer pupil boundaries. This localization is followed by Daugman’s rubber sheet model [6] using a circular boundary model to transform the iris texture into a rectangular 512×64 pixel area. Finally, the texture is enhanced using blockwise brightness estimation.

While the first method was tuned to deliver accurate segmentation for the *CASIA* dataset, we obtained the open iris recognition software OSIRIS³ (version 2.01) in order to segment images of the *ICE* dataset (OSIRIS also comes with an official evaluation on ICE-2005). The OSIRIS segmentation module uses again a circular Hough transform, but also an active contour approach to detect the

¹ The Center of Biometrics and Security Research, CASIA Iris Image Database, <http://www.sinobiometrics.com>

² National Institute of Standards and Technology, Iris Challenge Evaluation (ICE) 2005, <http://iris.nist.gov/ice/>

³ BioSecure Project, Open Source for IRIS reference system, http://svnext.it-sudparis.eu/svnview2-eph/ref_syst/Iris_Osiris

contours of iris and pupil. In order to get similar input textures for the feature extraction module, we employed adaptive histogram equalization after normalization on the 512×64 pixel sized textures with a window size of 32 pixels.

3.3 Feature extraction

The first feature to be extracted is a custom implementation of the iris-code version by *Ma* et al. [14]. This algorithm extracts 10 one-dimensional horizontal signals averaged from pixels of 5 adjacent rows from the upper 50 pixel rows. Using dyadic wavelet transform, each of the 10 signals is analyzed, and from a total of 20 subbands (2 fixed bands per signal), local minima and maxima above a threshold define alternation points where the bitcode changes between successions of 0 and 1 bits. Finally, all 1024 bits per signal are concatenated and yield the resulting 10240 bits code.

The second applied feature is based on row-wise convolution with Log-Gabor filters, following an implementation by *Masek*⁴ resembling Daugman's feature extraction approach. Again, rows are averaged to form 10 signals of 512 pixels each. The bit-code results from quantizing the phase angle of the complex response with two bits and sticking together all 10240 bits.

3.4 Matching and Decision

In order to assess the performance of our adapted version of the Levenshtein Distance (*Constrained LD*), we compare its recognition accuracy as well as average matching time requirements with the traditional Hamming Distance (*Minimum HD*), as used in commercial systems today [1]. The latter employs the fractional HD (i.e., the number of disagreeing bits of both codes divided by the total number of bits) over a number of bit shifts, and returns the minimum as a degree of dissimilarity between iris codes. Finally, the decision module compares the outcome with a threshold to classify the match as either genuine or imposter.

4 Experimental Results

The following subsections will cover each a specific research question. Unless otherwise noted, LD refers to the in Sect. 2 introduced constrained version in our tests and HD refers to the minimum fractional HD over a fixed number of shifts. Experiments are carried out in verification mode using the equal error rate (EER) and receiver operating characteristics (ROC) as the main performance indicators.

⁴ L. Masek: Recognition of Human Iris Patterns for Biometric Identification, Master's thesis, University of Western Australia, 2003

4.1 Does LD enhance iris recognition accuracy compared to traditional minimum HD?

We have tested *Constrained LD* and *Minimum HD* on *CASIA* and *ICE* datasets using two different algorithms: *Ma* and *Masek*. From the obtained EERs summarized in Table 1 we can see, that LD clearly outperforms HD in all tested combinations. All rates refer to a fixed maximum amount of 20 shifts in order to ensure a fair comparison between LD and HD with respect to rotational tolerance. However, it seems that the feature extraction algorithm influences the amount of improvement. For *Ma*, relative EER improvements were more pronounced than for *Masek* on both databases: up to 40 percent improvement (from 8.6% to 4.96% EER in case of *ICE*) could be achieved. A possible explanation of this behaviour can be found in the manner this algorithm defines its iris-code bits: the alternating zero and one chains seem ideally suited for non-linear LD alignment. Still, even for *Masek* relative EER improvements of up to 10 percent (from 6.08% to 5.49% EER in case of *ICE*) could be reported for LD. It is remarkable that even in case of very high HD accuracy on *CASIA* (0.58% EER for *Ma*, 0.89% EER for *Masek*) LD can push forward recognition rates (to 0.44% EER for *Ma*, 0.81% for *Masek*). Whereas the EER reflects only a single point of operation, the better performance of LD over HD becomes even more visible, if we take a closer look at the ROC curves for *ICE* and *CASIA* datasets in Figs. 2, 3. Almost all LD curves are clearly superior to HD, except the one for *Ma* on *CASIA*, which also depends on the selected maximum shift count, as will be investigated in the next research question. Table 2 lists matching times.

Table 1. EERs of HD versus LD (20 shifts).

EER (%)	ICE		CASIA	
	Ma	Masek	Ma	Masek
Minimum HD	8.60	6.08	0.58	0.89
Constrained LD	4.96	5.49	0.44	0.81

Table 2. Average Matching Times of HD versus LD (20 shifts).

AMT (ms)	ICE		CASIA	
	Ma	Masek	Ma	Masek
Minimum HD	0.73	0.73	0.99	0.97
Constrained LD	4.07	3.71	4.04	3.66

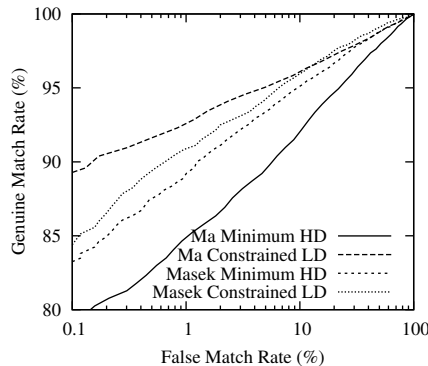


Fig. 2. HD versus LD: ROC for *Ma* and *Masek* (20 shifts) on *ICE* dataset.

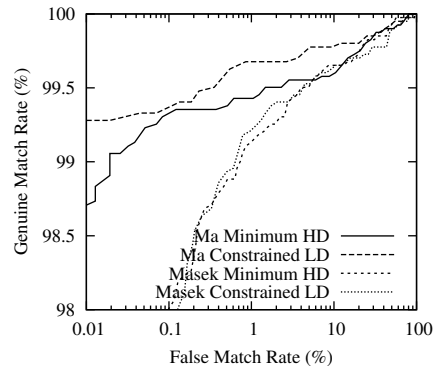


Fig. 3. HD versus LD: ROC for *Ma* and *Masek* (20 shifts) on *CASIA* dataset.

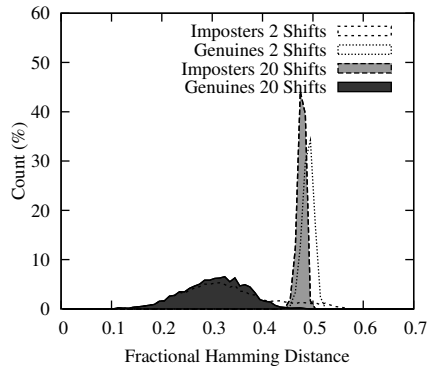


Fig. 4. Impact of max. HD shifts on *Ma*'s score distributions (*CASIA* dataset).

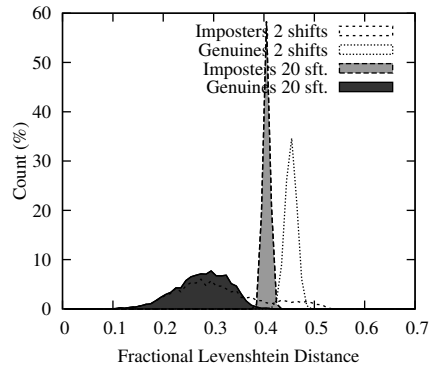


Fig. 5. Impact of max. LD shifts on *Ma*'s score distributions (*CASIA* dataset).

4.2 Which tradeoff exists between the maximum number of shifts, time complexity and recognition accuracy?

From Table 2 we can see, that LD requires additional matching time, however it theoretically lies in the same complexity class like HD. LD is on average 4-5 times slower than HD over different combinations of algorithms and datasets (results refer to the execution on a single processor at 2.8 GHz). While a single match using HD takes less than 1 ms, LD-based matching needs approximately 4 ms in case of 20 bit shifts. While assessing the performance of LD we noticed, that the number of shifts has an important impact on both recognition accuracy and certainly average matching time (due to additional comparisons). From the implementations it is easy to derive, that for both HD and LD in case of small bit shifts, doubling the number of shifts also results in approximately twice as much matching time. From the Figs. 4, 5 we can see, how larger maximum shifts cause inter-class (imposter) as well as intra-class (genuine) distributions to shift to the left, for both LD and HD.

The number of shifts is essential in order to cope with angular displacement of two iris textures to be matched. Furthermore, as shifts should be executed on iris-codes and not on iris-textures in order to avoid multiple extraction of features with resulting time overhead, it is important to consider the bit sampling rate with respect to texture pixels. In our tested algorithms the number of bits per texture row is equal to the number of pixels per row. A bit shift of one thus corresponds with shifting the texture a single pixel to the left or right. Fig. 6 illustrates the impact of shift count on recognition accuracy for the *CASIA* dataset. It is worth noticing, that for all algorithms a significant improvement of performance is achieved at around 7 and 14 shifts. Furthermore, we can see that for very few bit shifts, LD performs worse than HD, which makes sense, since in order to benefit from the better non-linear alignment, we need at least an amount of shifts in the order of angular displacement of genuine pairs. Finally, Fig. 7 illustrates the resulting tradeoff between EER and average matching time.

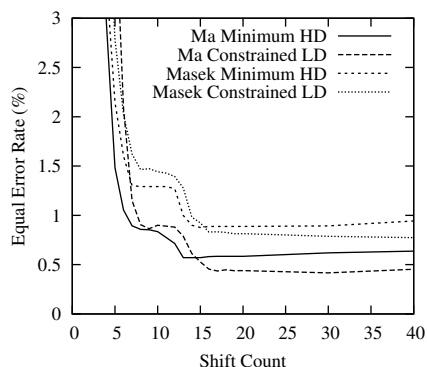


Fig. 6. Shifts-EER tradeoff for HD and LD on *CASIA* dataset.

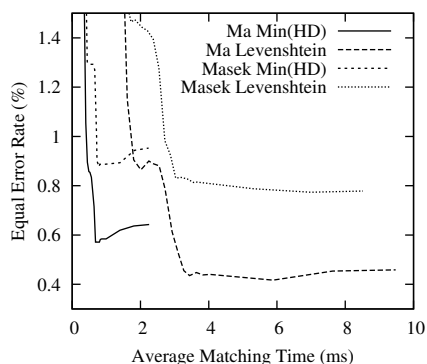


Fig. 7. Time-EER tradeoff for HD and LD on *CASIA* dataset.

5 Conclusion

In this paper we presented an adapted version of LD as a novel matching technique for iris recognition. Mislocation of pupil and iris centers cause significant and irrecoverable mapping distortions, which can not be overcome with simple HD-based matching. Even with accurate segmentation algorithms, inaccuracies are likely to occur in unconstrained biometrics. Due to its transparent integration, LD is a useful matching technique to tolerate non-linear deformations of iris textures and has been shown to reduce recognition rates drastically (from 8.6% to 4.96% EER for Ma in case of the challenging ICE dataset). Given a sufficient number of bit shifts, all tested algorithms and datasets reported a superior performance of LD over HD. With the introduction of alignment constraints, LD is only 4-5 times slower than HD with an equal amount of shifts. Existing work in unconstrained iris biometrics has mainly concentrated on normalization issues so far, while alternative matching techniques have largely been neglected. We demonstrated and highlighted the ability to tolerate distortions at the matching stage in this work.

Acknowledgment

This work has been supported by the Austrian Federal Ministry for Transport, Innovation and Technology (FIT-IT Trust in IT-Systems, project no. 819382).

References

1. Bowyer, K.W., Hollingsworth, K., Flynn, P.J.: Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding* **110** (2008) 281 – 307

2. Matey, J.R., Naroditsky, O., Hanna, K., Kolczynski, R., Lolacano, D.J., Mangru, S., Tinker, M., Zappia, T.M., Zhao, W.Y.: Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE* **94** (2006) 1936 – 1947
3. Wheeler, F.W., Perera, A.G., Abramovich, G., Yu, B., Tu, P.H.: Stand-off iris recognition system. In: *Proceedings of the IEEE Second International Conference on Biometrics: Theory, Applications and Systems (BATS)*. (2008) 1–7
4. Proença, H., Alexandre, L.A.: Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Elsevier Image and Vision Computing* **28** (2010) 202–206
5. Matey, J.R., Broussard, R., Kennell, L.: Iris image segmentation and sub-optimal images. *Image and Vision Computing* **28** (2010) 215 – 222
6. Daugman, J.: How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology* **14** (2004) 21–30
7. Cha, S.H., Tappert, C., Yoon, S.: Enhancing binary feature vector similarity measures. *Journal of Pattern Recognition Research* **1** (2006) 63–77
8. Choi, S., Cha, S., Tappert, C.: A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics* **8** (2010) 43–48
9. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* **10** (1966) 707–710
10. Yuan, X., Shi, P.: A non-linear normalization model for iris recognition. In: *Proceedings of the International Workshop on Biometric Recognition Systems (IW-BRS)*. (2005) 135–142
11. Wyatt, H.J.: A 'minimum-wear-and-tear' meshwork for the iris. *Vision Research* **40** (2000) 2167 – 2176
12. Schimke, S., Vielhauer, C., Dittmann, J.: Using adapted levenshtein distance for on-line signature authentication. In: *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04) Volume 2*, Washington, DC, USA, IEEE Computer Society (2004) 931–934
13. Myers, C.S., Rabiner, L.R.: A comparative study of several dynamic time-warping algorithms for connected word recognition. *The Bell System Technical Journal* **60** (1981) 1389–1409
14. Ma, L., Tan, T., Wang, Y., Zhang, D.: Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image Processing* **13** (2004) 739–750