AI Challenges in Iris Recognition. Processing Tools for Bath Iris Image Database

NICOLAIE POPESCU-BODORIN, *Member*, *IEEE*
'*Spiru Haret*' University
Dept. of Mathematics and Computer Science
13 Ion Ghica, Bucharest 3
ROMANIA
bodorin@ieee.org

VALENTINA E. BALAS, *Senior Member*, *IEEE*
'*Aurel Vlaicu*' University
Faculty of Engineering
2 Elena Dragoi, Arad 31033
ROMANIA
balas@drbalas.ro

Abstract: - The present paper aims to bring some interesting iris recognition related problems in attention of the artificial intelligence research community and also to mark the release of the first publicly available set of processing tools for University of Bath Iris Image Database (UBIID), tools that can be used to generate test data sets, without losing precious time. New recognition results for UBIID are also presented here.

Key-Words: - iris recognition, University of Bath Iris Image Database

1 Introduction

The present paper is a proposal written for artificial intelligence research community and it also aims to be a fast introduction to some interesting research subjects related to iris recognition.

The motivation behind this work is the belief that the future major improvements in iris recognition will come from the field of artificial intelligence. As an argument, at least two sub-problems of iris recognition, namely iris segmentation and occlusion removal, are NP-problems because they are reducible to certain Boolean satisfiability problems posed in terms of finding a particular spread of chromatic values in relation to particular chromatic gradients (textures) in their neighborhood. Therefore, it is very clear that finding simultaneously *fast* and *accurate* segmentation or occlusion removal procedures is another facet of the open problem *P vs. NP*. Consequently, there are only two choices that could guarantee speed in iris image processing: *model simplification* and *heuristic algorithms*. This fact inevitably qualifies iris recognition as an interesting subject of artificial intelligence.

We will also present here one of the best recognition results announced so far for the University of Bath Iris Image Database (UBIID). They are obtained by following a recently proposed approach to iris image processing ([1], [2]) which proves that unexplored paths still exist in iris recognition, and by using a package of Matlab processing tools, further referred to as the *Processing Toolbox for UBIID* (PT-UBIID, [3]). Toolbox release will take place simultaneously with the publication of this article.

Because of the limited space available here, we would like to refer to Bowyer et al. for a survey of iris recognition [4], rather than including our own survey which would be neither shorter nor better. There are also newer developments regarding the concepts of fragile bits [5], [6] and template aging [7]. We will also mention (in chronological order) the pioneering works of Wildes [8] and Daugman [9], iris recognition projects maintained at CASIA [10], the studies undertaken by Monro [11], Iris Challenge Evaluation (ICE) and Iris Exchange (IREX) projects conducted by NIST [12], and Noisy Iris Challenge Evaluation (NICE) projects managed by SOCIA Lab [13].

The second section of this paper describes the main sub-problems of iris recognition. A first set of nine artificial intelligence challenges is given here. As an example of heuristic approach to iris recognition, a set of processing tools designed for UBIID is further described in the third section which also contains two of the latest results announced for this database. The results of a test based on multi-enrollment scenario are given in the fourth section. Iris alignment challenge is formulated in the fifth part of the paper.

2 Iris Recognition Sub-Problems

In brief, iris recognition is a succession of operations designed to extract a binary iris code (or, more generally, a feature vector) from an eye image.

A particular acquisition procedure could require a special iris code extraction routine, but usually, there are three main sub-problems of iris recognition: iris segmentation, iris binary encoding (or more general, iris texture analysis and features encoding) and iris code matching (features matching). Hence, the formulation of iris recognition in terms of artificial intelligence means finding heuristic algorithms and/or neural architectures able to answer the following challenges:

(C1): Crop the eye image around the pupil.
(C2): Find the pupil center and pupillary boundary.
(C3): Crop the eye image around the iris.
(C4): Find the iris center and the limbic boundary.
(C5): Collapse the search space when looking for the pupillary or limbic boundary.
(C6): Identify iris texture occlusions (eyelashes, eyelids, specular lights) if any.
(C7): Encode and match the iris codes.

Designing iris recognition systems could be viewed as a heuristic way of solving a specific constraint satisfaction problem in which the goal is to choose a set of techniques and the values of the calibration variables (for example: iris code size, the number of enrolled templates, the type of the similarity measure used to compare iris codes, other internal control variables tuning the balance between speed and precision, etc.) so as to meet some conditions imposed in terms of accuracy / reliability (a small False Accept Rate) and comfort / correctitude (a small and entirely motivated False Reject Rate). Hence, the hardest AI challenge related to iris recognition is to design an exploratory supervised intelligent agent enabled to evolve itself and also to control complex computations and to detach (to design, to test and to validate) child applications from itself:

(C8): Build an exploratory supervised intelligent agent for iris recognition.

Among all the above challenges, (C8) is particularly demanding but it is also equally rewarding. We are working on designing such a pilot application. A part of it is the proposed toolbox (PT-UBIID) which enabled us to obtain iris recognition results presented further in this paper. Because the toolbox is just a collection of functions, it is unable to run independently from our decisions. A possible improvement of this behavior could come with the answer to the following challenge:

(C9): Build at least a rudimentary control unit (including an inference engine) enabling the exploratory agent mentioned in (C8) to act independently much of the time based on its own decisions.

Two important steps toward solving this problem have been undertaken in [14] and [15] but further developments in terms of fuzzy logic are needed because all iris recognition results ever published reveal that there is no *crisp separation* between the class of matching scores and the class of non-matching scores (the classes are *intricate* and their separation is *fuzzy* because some similarity values can be obtained both as matching scores and as non-matching scores).

Another argument for fuzzification is that we use *heuristic* procedures for segmentation and occlusion removal. Inevitably, their results are *near solutions*, not *exact solutions*. Hence, quantifying their quality is a *matter of degree* [16].

The majority of the best approaches to iris recognition available these days use binary iris codes to retrieve phase information from the iris texture and compare the Hamming distance between iris codes to a recognition threshold. Modified Hamming distances are also used and different authors usually use different phase encoders.

There are also proprietary iris code formats, one of the most recent of them being owned by Daugman and tested by NIST in [17].

The quality of an iris recognition system is very often expressed through the False Reject and False Accept Rates (FRR and FAR) which illustrate the relation between user comfort and the risk of allowing identity confusions. It is currently accepted [17] that a reasonable balance between risk and comfort is given by a FAR of 10^{-6} and a FRR of 2-10%, depending on the average quality of the eye images used. It is not clear why the neural approaches to iris encoding and matching usually do not achieve the same performances.

In terms of artificial intelligence, a way to answer (C7) could be to define a neural network architecture or a heuristic algorithm able to replicate currently available iris recognition results obtained by comparing the iris codes directly. Such an approach would assume that each enrolled identity is stored as a trained memory or as a feature vector and would be able to classify candidate iris codes as well as possible by preserving the quality of the separation between genuine and imposter score distributions in terms of False Accept/Reject Rates.

3 Processing Tools for UBIID

Processing Toolbox for the University of Bath Iris Image Database (PT-UBIID-v.01, [3]) illustrates our way of dealing with the challenges (C1-C8). The main components of the package are:

- A fully automatic unsupervised heuristic iris segmentation procedure derived from Circular Fuzzy Iris Segmentation [1],[2];
- A binary encoder based on Hilbert

transform and derived from Gabor Analytic Iris Texture Binary Encoder [1];

- A statistical occlusion detection procedure based on standard deviation of the chromatic values;
- A single / multi-enrollment iris recognition simulator;
- Graphical display functions for generating Encapsulated PostScript files like Fig.1-Fig.7;

The latest iris recognition results obtained by using PT-UBIID are illustrated in Fig.2-Fig.8. The complete scenarios of three iris recognition tests are given in Tables I-III.

The acronyms used in Tables I-III have the following meanings: Equal Error Rate (EER), False Accept Rate (FAR), False Reject Rate (FRR), Odds of False Accept (OFA [2], [6]), Maximum Imposter Score (MIS), Recommended Recognition Threshold (RRT), Circular Fuzzy Iris Segmentation (CFIS [1], [2]), Gabor Analytic Iris Texture Binary Encoder (GAITBE [1], [6]), Mean-Deviation Similarity Score (MDSS [1], [6]).

There are two important aspects that must be mentioned in order to enable future comparisons with other experimental results:

- firstly, the region used for comparison is as wide as possible (here, we do not use butterfly iris segment as in [1]) and the occlusions are masked.
- secondly, the phase content of the iris texture is retrieved from two different scales. Each iris code is a binary quantization of the instant phase computed for a sum of two analytic signals obtained by using two Hilbert filters of size 16 and 32, respectively.

TABLE I
THE RESULTS OF AN EXPLORATORY IRIS RECOGNITION TEST ON UBIID

System parameters:	
Iris code size	32x512 bits (**2KB**)
Hilbert filter size	16; 32
Enrolled templates	1
Similarity score	Hamming
Inter-class score distribution:	
Mean	0.5022
Standard deviation	0.0089
Degrees-of-freedom	3’150
Number of unique pairs	484’132
Intra-class score distribution:	
Mean	0.6513
Standard deviation	0.0515
Degrees-of-freedom	86
Number of unique pairs	9’389
Evaluation criteria:	
Decidability index	4.0353
Fisher’s ratio	8.1419
EER	6.155 E-03
MIS / FAR(MIS) / FRR(MIS)	0.55 / 2.0655 E-06 / 0.0151
RRT / OFA(RRT) / FRR(RRT)	0.57 / 5.6414 E-10 / 0.0398
Storage efficiency	19%
Segmentation failures	6‰

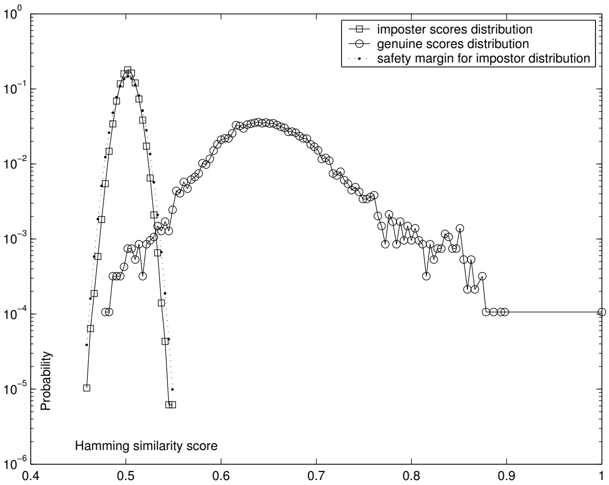

Figure 1: Inter-class (484’132 different pairs) and intra-class (9’389 different pairs) matching score distributions computed for UBIID (1000 images, 50 eyes) using fully automatic unsupervised iris segmentation (Circular Fuzzy Iris Segmentation, failure rate: 6/1000), Gabor Analytic Iris Texture Binary Encoder and statistical occlusion removal based on standard deviation of the chromatic values.

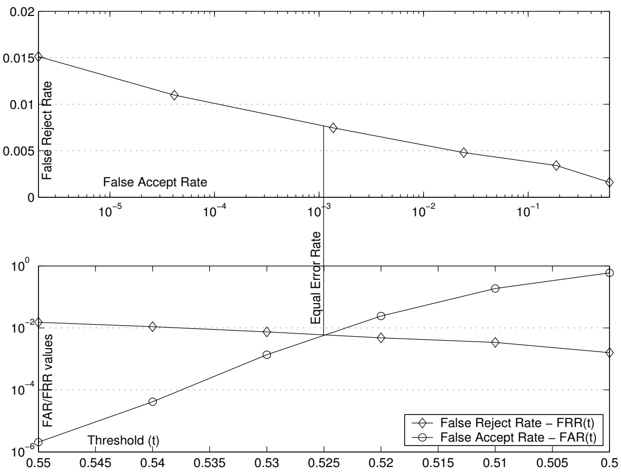

Figure 2: ROC (top), FAR, FRR curves and EER point for the first test (see also the Table I)

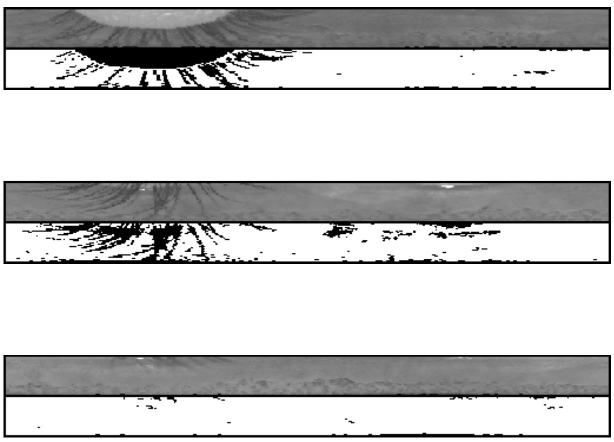

Figure 3: Statistical occlusion removal based on standard deviation of the chromatic values. From top to bottom: heavy / reasonable / minor occlusions.

Occlusion masks are obtained by excluding all pixels whose values exceed the confidence interval $[m-2s, m+1.5s]$, where 'm' and 's' are the mean and the standard deviation of the chromatic values found in the iris segment, respectively. Some examples of occlusion masks are given in Fig.3.

Storage efficiency is the ratio between the number of degrees-of-freedom computed for inter-class (imposter) score distribution and the size of the iris code (in bits).

An interesting comparison could be made between the results presented here (Fig.2, Fig.5, Fig.7 and Tables I-III) and those obtained in Iris Exchange tests and illustrated in Fig.11 from [17].

Fig.1, Fig.4 and Fig.6 show three distributions: imposter scores, genuine scores and *safety margin*. The last one is used to estimate the Odds of False Accept for those threshold values for which it is impossible to compute the False Accept Rate, i.e. for those similarity scores which do not appear in the experimental imposter distribution effectively (for example 0.57 in the first test and 0.59 in the second and third tests). The *safety margin* is a normal distribution having the same mean as the actual imposter distribution but a greater standard deviation (125% in Fig.2, 110% in Fig.4 and Fig.6).

TABLE II
SIMULATING A REAL-WORLD IRIS RECOGNITION APPLICATION

Iris codes:	
Size	8x256 bits (**2Kb**)
Binary	Yes
Interoperable	Yes
Simulation parameters:	
Identification	Yes
Comparison type	All-to-all
Similarity measure	Hamming
Total number of unique pairs (total number of comparisons)	477'753
Total comparison time	78 seconds
Average time / comparison	1.621E-4 seconds
Application type	Matlab (script)
Hardware	Pentium IV Prescott, 2.8GHz
Image source	Bath Database
Segmentation algorithm	CFIS
Phase Encoder	GAITBE
Inter-class score distribution:	
Mean	0.5048
Standard deviation	0.0157
Degrees-of-freedom	1'016
Number of unique pairs	468'637
Intra-class score distribution:	
Mean	0.6711
Standard deviation	0.0531
Degrees-of-freedom	73
Number of unique pairs	9'116
Evaluation criteria:	
Decidability index	4.2449
Fisher's ratio	9.0098
EER	1.856E-3
MIS / FAR(MIS) / FRR (MIS)	0.5791 / 2.134E-6 / 0.0177
RRT / OFA(RRT) / FRR(RRT)	0.59 / 3.699E-7 / 0.0377
	0.595 / 8.642E-8 / 0.0511
Storage efficiency	49.59%
Segmentation failures	22‰

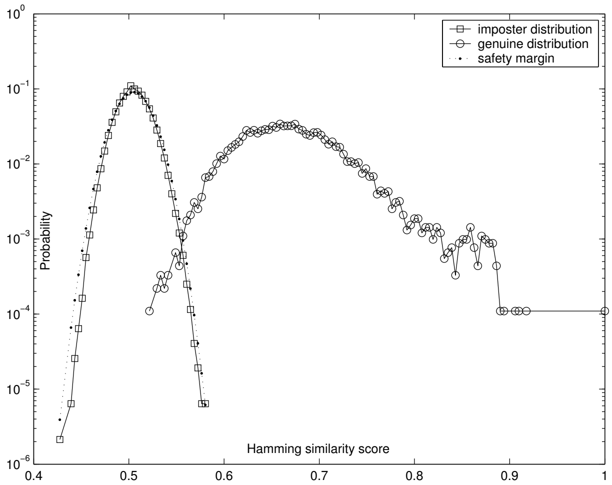

Figure 4: A second recognition test simulating a real-world application (see also the Table II).

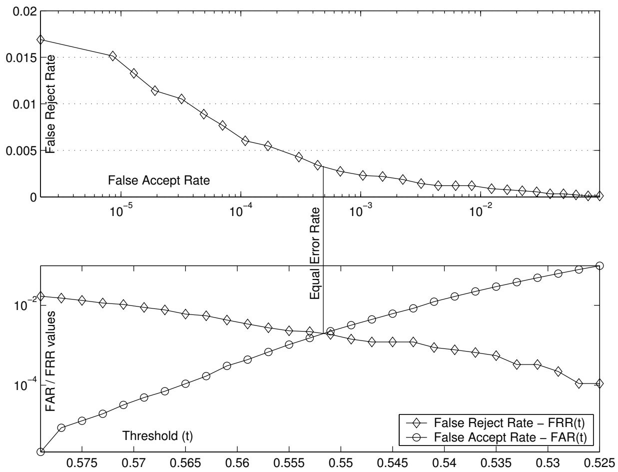

Figure 5: ROC (top), FAR, FRR curves and EER point for the second test (see also the Table II).

The Odds of False Accept for a certain threshold are computed as the cumulative of the safety margin up to the given threshold [6].

The proposed toolbox allows real world simulations to be derived easily. A telling example is given here as the test described in Table II, Fig.4, and Fig.5. This second test illustrates a different way of negotiating application speed, the risk of allowing identity confusion, user comfort and storage efficiency (see also Table I, Fig.2 and Fig.3 for comparison to the first iris recognition test). The most important differences between these two tests are: iris code length (2 Kilobytes vs. 2 Kilobits), storage efficiency (19.22% vs. 49.59%) and speed.

4 Closer to Real World

At first glance, iris recognition tells us something about how all iris codes representing the same eye tend to group in a cluster which defines a binary identity. Apparently, iris recognition theory affirms that these clusters are stationary (they do not move through the iris code space) and contain sufficiently larger and mutually exclusive kernels.

Baker, Bowyer and Flynn [7] have documented very recently a hypostasis of *template aging* in which the matching between candidate and gallery templates tends to degenerate when the current candidate iris codes are compared to old gallery templates extracted years ago. It means that the applications designed to handle this kind of situations should assume that the clusters which encode the identities are, in fact, non-stationary. Hence, in a real-world iris recognition system, multi-enrollment is mandatory because otherwise it could not be possible to trace binary identity changes over time. This is the motivation behind the third test presented here (Table III, Fig.6 and Fig.7) which assumes multi-enrollment scenario with four enrolled templates for each identity.

TABLE III
MULTI-ENROLLMENT IRIS RECOGNITION APPLICATION

Iris codes:	
Size	8x256 bits (**2Kb**)
Binary	Yes
Interoperable	Yes
Simulation parameters:	
Identification	Yes
Enrolled identities (unique eyes)	50
Enrolled templates / identity	4
Templates in gallery	200
Candidate templates	778
Comparison type	Current candidate to all enrolled identities
Similarity measure	MDSS[1]
Total number of comparisons	38'900
Total comparison time	19 seconds
Average time / comparison	4.884E-4 seconds
Application type	Matlab (script)
Hardware	Pentium IV Prescott, 2.8GHz
Image source	Bath Database
Segmentation algorithm	CFIS
Phase Encoder	GAITBE
Inter-class score distribution:	
Mean	0.5067
Standard deviation	0.0134
Degrees-of-freedom	1'386
Number of comparisons	38'122
Intra-class score distribution:	
Mean	0.6757
Standard deviation	0.0350
Degrees-of-freedom	178
Number of comparisons	778
Evaluation criteria:	
Decidability index	6.3674
Fisher's ratio	20.2724
EER	<1.3E-3
MIS / FAR(MIS) / FRR (MIS)	0.5627 / 2.62E-5 / 1.29E-3
RRT / OFA(RRT) / FRR(RRT)	0.58 / 3.56E-7 / 3.86E-3
	0.59 / 8.74E-9 / 5.14E-3
	0.60 / 1.4E-10 / 0.01156
Storage efficiency	67.70%
Segmentation failures	22‰

[1] MDSS stands for Mean-Deviation Similarity Score [1], [5].

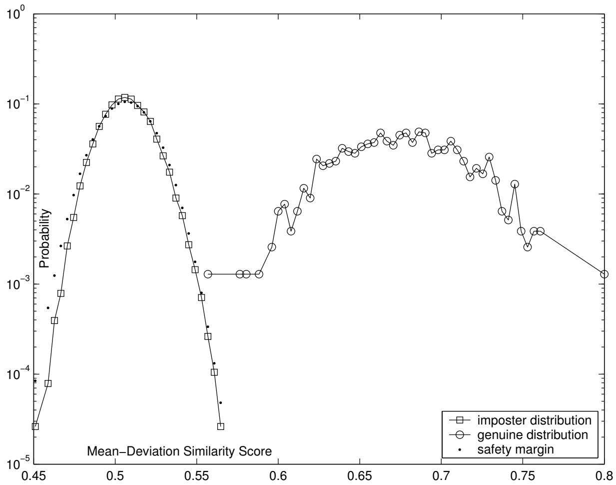

Figure 6: A third recognition test simulating a real-world multi-enrollment application.

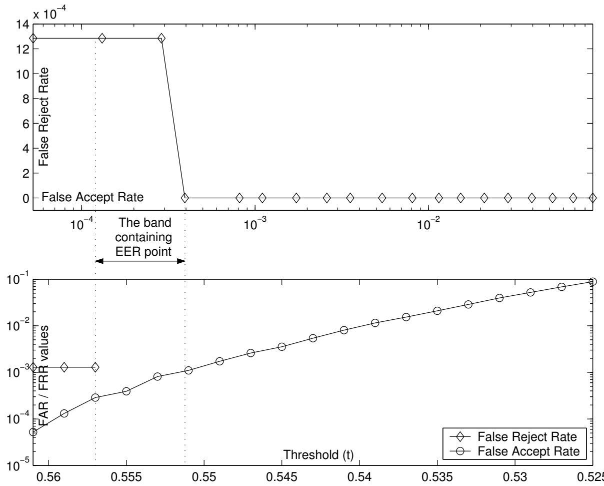

Figure 7: ROC (top), FAR, FRR curves and EER point for the third test (see also the Table III).

5 One More Challenge

The extreme scores within genuine distributions displayed in Fig.1, Fig.4 and Fig.6 are determined by perfectly identical iris images (on the right-side of the distributions), and by incorrect segmentation or incorrect angular/radial iris alignment (on the left-side). The analysis of the extreme scores within the genuine distribution gave us the opportunity to see that some irides are more dramatically influenced by pupil dilation than others. Usually, this is the cause of incorrect radial alignment. The problem with the wrong radial alignment is that it is not surmountable through an elastic deformation. Hence, there is one more challenge to be answered in iris recognition:

> *(C10): Describe radial iris movement through approximate equations formulated in the framework of discrete (digital) image topology.*

One reason for which the formal verification of iris recognition 'theory' fails is the nature of the problem itself. A simple iris recognition 'theory' written in crisp binary logic would say that:

if (a) the quality of eye images is good enough, and if (b) the localization/segmentation of the pupil is correct, and if (c) the localization/segmentation of the iris is correct, and if (d) the captured iris images presents all irides in nearly the same standard posture, and if (e) all iris images show comparable pupil dilation and if (f) all iris segments extracted are normalized at the same dimension, and if (g) all binary iris codes are generated using the same encoder, and if (h) all iris codes are compared using the same distance (Hamming, for example),

then:

(i) the greatest imposter score is smaller than the smallest genuine score (where the scores reflect the similitude between iris codes, not distance):

$$(a \wedge b \wedge c \wedge d \wedge e \wedge f \wedge g \wedge h) \rightarrow i$$

Even when working on the UBIID (an ideal database) this simple theory is contradicted by practice (see Fig.1, Fig.4, Fig.6, and also Fig. 11 in [17]). The problem is that only the assertions f, g, and h can be quantized using binary truth values. Quantifying all of the others assertions is a *matter of degree*.

The second reason for which the formal verification of iris recognition fails is the fact that instead of comparing two normalized iris images (two matrices of unsigned 8-bit integers) directly, we test the similarity between two 'shadows' left by them in a space of binary matrices. Hence, finding two different irides such that their iris codes will match each other is just a matter of time and probabilities. Such a counter-example was already found during a study concerning long-term evolutions on the currency exchange global market [18]. Of course, the counter-example refers to our encoder (included in the proposed toolbox) but other encoders could be tested in a similar manner by their authors.

6 Conclusion and Future Works

The present paper presented the main iris recognition results obtained by the first author during its PhD research and replicated by the second author. The first publicly available Processing Toolbox for the UBIID is also released together with this paper in order to enable result replication and future comparisons. The future efforts will be directed to answering the challenges (C8)-(C10) mentioned above in this paper.

References:

[1] N. Popescu-Bodorin, "Exploring new directions in iris recognition", *Proc. 11th Int. Symp. on Symbolic and Numeric Algorithms for Scientific Computing*, IEEE Computer Society Press, pp. 384-391, 2009.

[2] N. Popescu-Bodorin, "A fuzzy view on k-means based signal quantization with application in iris segmentation", *The 17th Telecommunications Forum*, University of Belgrade, 2009.

[3] N. Popescu-Bodorin, Processing Toolbox for the University of Bath Iris Image Database (PT-UBIID-v.01), http://fmi.spiruharet.ro/bodorin/pt-ubiid/.

[4] K. W. Bowyer, K. Hollingsworth and P. J. Flynn, "Image understanding for iris biometrics: a survey", *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281-307, May 2008.

[5] K. P. Hollingsworth, K. W. Bowyer and P. J. Flynn, "The best bits in an iris code", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 6, pp. 964-973, June 2009.

[6] N. Popescu-Bodorin, "'Fragile bits' vs. multi-enrollment - A case study of iris recognition on Bath University Iris Image Database", *ROMAI Journal*, Nr.5, vol. 2, pp. 127-144, 2009.

[7] S. E. Baker, K. W. Bowyer and P. J. Flynn, "Empirical evidence for correct iris match score degradation with increased time-lapse between gallery and probe matches", *Proc. 3rd IEEE Int. Conf. on Biometrics*, L.N.C.S., Vol. 5558, pp. 1170-1179, 2009.

[8] R. Wildes, "Iris recognition - an emerging biometric technology", *Proc. of the IEEE*, vol. 85, no. 9, pp. 1348-1363, September 1997.

[9] J. Daugman, personal online collection of papers and articles, http://www.cl.cam.ac.uk/~jgd1000/, 2010.

[10] Center for Biometrics and Security Research, National Laboratory of Pattern Recognition, *CASIA Iris Images Databases* and online list of publications, www.cbsr.ia.ac.cn/IrisDatabase.htm.

[11] D. M. Monro, online list of publications, February, 2010, http://www.smartsensors.co.uk/pub.htm.

[12] Iris Challenge Evaluation, http://iris.nist.gov/ice/, February, 2010.

[13] SOCIA Lab., Noisy Iris Challenge Evaluation, http://nice1.di.ubi.pt/, February, 2010.

[14] N. Popescu-Bodorin, L. State, "Cognitive binary logic - the natural unified formal theory of propositional binary logic", *Proc. 4th WSEAS Conf. on Computational Intelligence*, pp. 135-142, WSEAS Press, 2010.

[15] N. Popescu-Bodorin, V. E. Bălaş, "From cognitive binary logic to cognitive intelligent agents", accepted in *14th IEEE Int. Conf. on Intelligent Engineering Systems* (INES'10), May 2010.

[16] L. Zadeh, "Test-score semantics for natural languages", *Proc. 9th Conf. on Computational Linguistics*, Vol. 1, pp. 425-430, 1982.

[17] P. Grother, E. Tabassi, G. Quinn and W. Salamon, *Interagency report 7629: IREX I - Performance of iris recognition algorithms on standard images*, N.I.S.T., October, 2009.

[18] N. Popescu-Bodorin, "Automatic detection of common long-term monetary policies on global exchange market using Gabor analytic phase binary encoder", accepted in *4th Int. Conf. Present Issues of Global Economy*, Constanța, 2010.