Age prediction from iris biometrics

ERBILEK, M., FAIRHURST, M. and DA COSTA ABREU, Marjory
<http://orcid.org/0000-0001-7461-7570>

Published version

ERBILEK, M., FAIRHURST, M. and DA COSTA ABREU, Marjory (2013). Age prediction from iris biometrics. In: 5th International Conference on Imaging for Crime Detection and Prevention (ICDP 2013). Institution of Engineering and Technology.

Copyright and re-use policy

Age Prediction from Iris Biometrics

M. Erbilek*, M. Fairhurst* and M. C. D. C. Abreu†

* School of Engineering and Digital Arts, University of Kent, Canterbury, Kent CT2 7NT, UK.

*{M. Erbilek, M.C.Fairhurst}@kent.ac.uk

† Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte,

Natal, RN 59078-970, Brazil.

†marjory@dimap.ufrn.br

Keywords: iris age prediction, iris geometric features, intelligent classifier structures.

Abstract

This paper proposes and investigates experimentally an approach to age prediction from iris images by using a combination of a small number of very simple geometric features, and a more versatile and intelligent classifier structure which can achieve accuracies to 75%. To our knowledge, this is the first experimental study of three class age prediction from iris images.

1 Introduction

The field of biometrics – the identification of individuals from measurement of their physiological or behavioural characteristics – is now well established, offering practical solutions in a number of important application areas. In recent years, however, there has been a growing interest in exploiting the predictive properties of biometric data for scenarios in which full identification of a specific individual is not the primary aim. In such scenarios, the aim is more generally the prediction of a “soft” biometric marker – a piece of information which characterises, but is not unique to, an individual. A typical example might be to predict the age or gender of an individual, such information revealing a specific piece of useful identity-related information, but information which is common across a larger number of individuals. Despite the non-unique nature of its outcome, such a predictive capacity is nevertheless extremely valuable in a number of practical applications. This is most obvious in, for example, forensic analysis in criminal investigations, in providing security monitoring in electronic transactions, in subject profiling activities and many other areas, including the assessment of entitlement to age-restricted goods and services [1].

Biometric-based age estimation is perhaps the most common and valuable manifestation of this process. The literature shows that face [1, 2], speech [3, 4], and signature [5] biometrics, have received most attention in the research area of age prediction. However, even though it is thus possible to find some interesting work dealing with age estimation based on various different biometric modalities; there is only one reported study [6] concerning age estimation based on iris biometrics (although gender and ethnic group prediction from iris images is proposed in [7, 8] and [8, 9] respectively). However, iris recognition is widely regarded as one of the most reliable biometrics and there is currently a particular interest in the acquisition of samples for iris recognition systems designed to capture eye images at a distance or while the subject is mobile (i.e. iris "on the move"). Such scenarios broaden the scope and significantly enhance the potential usability of age prediction based on the iris modality for many important applications, including those noted above.

Specifically, in this paper we propose an approach to the age prediction task which, while demonstrating the viability of iris-based age prediction per se, also uses only simple geometric features extracted from iris images, and is thus fast and efficient in implementation in comparison with the previous study reported in [6].

2 Related work

As we have reported above, age prediction from iris biometrics is studied only by Sgroi *et. al.* [6]. This study proposes a classification technique which categorises a person into “young” or “old” age groups from the iris’s texture-based characteristics.

Iris biometric data used in this study have been collected at the University of Notre Dame and are not generally publicly available. Biometric data in the “young” group consists of 50 subjects (with 3 left and 3 right eye samples) whose age is between 22 and 25, and the “old” group consists of 48 subjects (with 3 left and 3 right eye samples) whose age is greater than 35. Hence, a total of 98 subjects with 6 samples (total of 600 images) are used for this experimental study. Initially, 630 features were computed from the segmented and normalised iris texture and then classification is performed by using the RandomForest algorithm from the Weka software using 300 trees. Experimental results have shown that the correct classification rate of this method is 64.68%.

A more relevant and earlier analysis of ageing issues in iris biometrics, which can be found in [10], shows that the physical ageing effects on iris are primarily the result of the physiology of pupil dilation mechanisms, with pupil dilation responsiveness decreasing with age. Hence, since pupil dilation is clearly related to the geometric appearance of the pupil and the iris, this suggests that geometric features of the iris may provide useful information for the age prediction task.

Also, in [11] (and in [12], although from a different modality), Erbilek and Fairhurst have investigated, analysed and documented the effects of different age-band assignment, in order to guide and enhance the management of age-related data and take a step towards the possibility of more objectively determining optimal age-bands which offer a greater possibility of minimising the sensitivity of a system which relies on such information. According to the results presented there, it is suggested that a structure which divides a test population into the three age bands defined by the boundaries ‘<25’, ’25-60’ and ’>60’ is one which best reflects age-related trends and provides useful information to support both the analysis and practical management of age-related factors in iris-based biometric systems.

Hence, on the basis of the discussion presented, in this paper, we propose an approach to the age prediction task which uses only five simple geometric features extracted from iris images. This not only reflects fundamental iris properties, but provides the basis of a technique which is simpler and computationally both less expensive and faster than the requirements of computing texture-based features of the iris (in comparison, for example with the previous study reported in [6] which used 630 texture features). For our proposed iris-based age estimation system, the age groups are defined to be based on three broad age groupings, which may be described as relating to the general categories “young”, “middle” and “older”, which is a wider experimental study than that reported in [6] which evaluated only a two class problem. Also we have performed our experiments using a commercially available and larger database than in [6].

3 Proposed age prediction approach

This section describes the basic processing of biometric data in our iris-based age prediction approach, as illustrated in Figure 1.

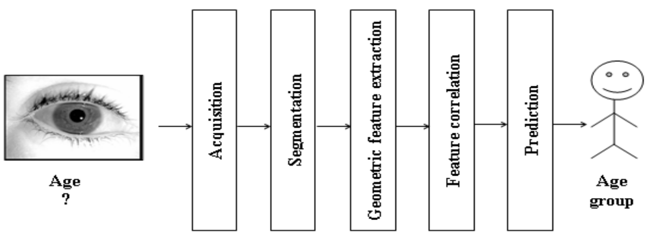

Figure 1: Proposed age prediction approach

The processing is based around the following:

- An eye image is captured in the Acquisition step. The commercially available data Set 2 (DS2) of the BioSecure Multimodal Database (BMDB) [13] is utilised in this study for this stage. Eye images in this particular database were acquired in a standard “office” environment managed by a supervisor and using the LG Iris Access EOU3000 system. During the acquisition, spectacles were not allowed to be worn by subjects, although contact lenses were allowed. Four eye images (two left and two right) were acquired in two different sessions with a resolution of 640*480 pixels. The 210 subjects providing the samples contained in this database are within the age range of 18-73. The iris samples of 10 subjects were found to be incorrectly labelled in this database (some of the left eye samples labelled as right or right eye samples labelled as left). Hence, this decreased the available number of subjects to 200.

- The Segmentation step localises the iris region from the acquired eye image. This step involves detection of the sclera/iris and pupil/iris boundaries. Each eye sample is first segmented using a robust segmentation algorithm as described in [10]. As the purpose of the study is to analyse the age prediction power of a more intelligent technique but with the of use simpler features, the wrongly segmented samples were segmented manually. Subsequently, the obtained iris and pupil parameters (which are specifically x-coordinate, y-coordinate and radius values) from the segmentation process are stored for each eye to be used subsequently.

- The Geometric Feature Extraction step extracts 12 geometric features (at the pixel level) from the iris (by using the detected iris (i_x: x-coordinate of the centre of the iris, i_y: y-coordinate of the centre of the iris, i_r: iris radius) and pupil (p_x: x-coordinate of the centre of the pupil, p_y: y-coordinate of the centre of the pupil, p_r: pupil radius) parameters in the segmentation stage. The extracted geometric features are shown in Table 1.

- The Feature Correlation step applies a correlation test to remove correlated geometric features to allow us to use only the more distinguishing and non-redundant features for the further processing in the proposed age estimation process of our experimental study. The inter-feature correlations were evaluated by using Spearman’s rank correlation [14] (a nonparametric-based estimate of correlation). For each of the features, the entire set of observations is ranked from smallest to largest and in the case of ties, average ranks are assigned. Spearman’s rank correlation (ρ) is determined for all of the possible combinations of two feature vectors (which results in 12*12 ρ values since there are 12 features) using Equation (1);

$$\rho = \frac{\Sigma(x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\Sigma(x_i - \bar{x}_i)^2 \Sigma(y_i - \bar{y}_i)^2}} \quad (1)$$

Where i=1,2,…n, and n is the number of features, x_i and y_i are the i^{th} components of a vector of ranks of observations of two features, and $\bar{x}_i$ and $\bar{y}_i$ are the mean of those features, x_i and y_i. Each obtained ρ value lies between -1 and 1, and a value of ρ close to zero indicates low correlation, while a value close to -1 or 1 indicates high correlation. For instance, a value close to 1 indicates positive correlation (y increases as x increases or y decreases as x decreases) and a value close to -1 indicates negative correlation (y increases as x decreases or y decreases as x increases).

Feature Number	**Feature Description**	**Feature Calculation**
GF1	Scalar distance between the x-coordinates of the centre of the iris and the pupil.	GF1=$\lvert p_x - i_x \rvert$
GF2	Scalar distance between the y-coordinates of the centre of the iris and the pupil.	GF2= $\lvert p_y - i_y \rvert$
GF3	Scalar distance between the centre of the iris and the pupil.	GF3=|GF1- GF2|
GF4	Total area of the iris	GF4=$\pi * i_r^2$
GF5	Total area of the pupil	GF5=$\pi * p_r^2$
GF6	True area of the iris	GF6= GF4- GF5
GF7	Area ratio	GF7= GF4 / GF5
GF8	Dilation ratio	GF8= i_r / p_r
GF9	Iris circumference	GF9 = p_i *2 * i_r
GF10	Pupil circumference	GF10= p_i*2*p_r
GF11	Circumference ratio	GF11=GF9/GF10
GF12	Circumference difference	GF12= GF9-GF10

Table 1: Extracted geometric features

- The Prediction step uses the data generated at the output of the previous step and performs the age prediction task according to the three age categories, <25, 25-60 and >60. For the sake of diversity, we use a range of common individual classifiers;
  - SVM [14]: This classifier is based on an induction method which minimises the upper limit of the generalisation error related to uniform convergence, dividing the problem space using hyperplanes or surfaces, splitting the training samples into positive and negative groups and selecting the surface which keeps more samples.
  - MLP [15]: This is a Perceptron neural network with multiple layers [16]. The output layer receives stimuli from the intermediate layer and generates a classification output. The intermediate layer extracts the features, their weights being a codification of the features presented in the input samples, and the intermediate layer allows the network to build its own representation of the problem. Here, the MLP is trained using the standard backpropagation algorithm [17] to determine the weight values.
  - KNN [18]: This is a version of the extremely simple K nearest neighbour classifier. A modified Hamming distance [19] is used which takes account of possible rotational inconsistencies in the original iris image, when the Hamming distance between two templates is calculated, by shifting left and right bit-wise one of the templates with a Hamming distance value calculated from successive shifts, and the smallest difference taken for the subsequent matching process. After calculation of Hamming distances, classification is performed by using a simple nearest neighbour classifier with k = 1.
  - Decision Tree [20]: This classifier uses a generalized “divide and conquer” strategy, splitting a complex problem into a succession of smaller sub-problems, and forming a hierarchy of connected internal and external nodes. An internal node is a decision point determining, according to a logical test, the next node reached. When this is an external node, the test sample is assigned to the class associated with that node.

However, we also investigate the use of classifier combination techniques to enhance performance. We deploy conventional Sum-based and Majority Voting fusion techniques but also, importantly, we evaluate the use of a novel agent-based intelligent classifier structure in two configurations based respectively on Sensitivity and Game Theory-based Negotiation (as described in [21]).

- The Game Theory-based Negotiation Method has been used as a cooperation tool in multi-agent systems. In game theory, the systematic description of the results can be carried out through the use of strategic games. A strategic game is a game in which a player chooses a plan of action only once and at the same time as his opponent. In order to help the players to make their decisions, a payoff matrix is used, in which each cell represents the payoff values which the players will have in a situation where these actions are chosen. The cell with the highest value is chosen.

- In the Sensitivity Negotiation Method, we have implemented an adaptation of the game theory negotiation method. The basic idea underpinning our proposed method is that a decrease in the confidence level of the agents is considered through the use of a sensitivity analysis during the testing phase. This analysis can be achieved by excluding and/or varying the values of an input feature and analysing the variation in the performance of the classifier method. The main aim of this analysis is to investigate the sensitivity of a classifier to a certain feature and to use this information in the negotiation process.

The training and the testing sets were formed to be person-disjoint sets. Approximately 72% of the subjects in each age group are used for training and the remaining subjects used as a testing set. The available number of subjects in the testing and the training sets for each age group is shown in Table 2.

Sets	**Age groups**		
	<25	25-60	>60
All	70	115	15
Training Set	50	82	11
Testing Set	20	33	4

Table 2: Number of subjects

4 Experiments and results

This section describes several experiments and presents results based on the iris modality for age prediction using geometrical features.

For the experimental study, all iris samples in each of the datasets are processed to form the biometric templates used by passing through the steps of segmentation and feature extraction as described in Section 3. Subsequently, in order to use non-redundant features for our experimental study, for each dataset, a correlation test is applied to the extracted 12 features as explained in Section 3, and highly correlated features are discarded (their ρ values lie between -1 and 1 and close to zero, but are not absolutely zero. Hence, we define a threshold value for classifying features as 'highly correlated'. Our choice is 0.4 which allows us to select a sufficient set of features while excluding the most highly correlated ones). Five features remain in our implementation (GF1-GF4 and GF8), and these form a feature vector for each iris sample in the dataset. Our empirical study then evaluates the accuracy of the proposed age prediction approach by using the defined feature vectors and classifiers presented in Section 3 with respect to the defined dataset.

The results are shown in Table 3. It should be noted that, using a Core 2 Quad PC with 2.33 GHz processor, 8 GB RAM and MATLAB R2011a software, the feature extraction time for the simple features adopted is only around 0.81 seconds.

The results obtained show that predictive accuracy across the five different individual classifiers ranges from approximately 50% to somewhat over 60% in the best case. This compares with an accuracy range of 50%-60% reported in [3] when predicting age from speech.

It is seen that adopting a multiclassifier approach with a traditional fusion technique such as the sum or vote approaches improves this somewhat as might be expected.

However, adopting our proposed intelligent agent-based configuration increases the prediction accuracy substantially, returning an accuracy greater than 75% using the negotiation configuration. Obtaining this high level of performance with only five geometric features can be explained by the fact that geometric features of the iris are apparently highly correlated with age. For example, the relationship between the dilation of the pupil (which is one of the geometric features used in this study – GF8) and subject age has been investigated in some detail in [10].

It should be noted too that Sgroi *et. al.* [6] reported experimental results with age classification model with an accuracy of around 64.68%, but they used 630 texture features which are much more computationally intensive to extract and, of course, their age prediction model is a two class problem only ("young" and "old"), rather than the three-class ("young", "middle" and "old") problem considered in our age-prediction scenario. Our prediction results are thus particularly encouraging.

Classifier		**Accuracy (%)**
SVM		62.06
MLP		61.80
Jrip		62.50
KNN		52.41
Decision Tree (J48)		51.09
Fusion	Sum	64.11
	Vote	62.94
Negotiation	Game theory	72.65
	Sensitivity	75.09

Table 3: Accuracy of the proposed approach

5 Conclusion

In this paper we have investigated experimentally an approach to age prediction from iris images which uses a combination of a small number of very simple (and therefore easily and efficiently computable) geometric features (ignoring texture-based information which is less likely to carry significant age-related information).

The performance we have been able to achieve - assigning each tested subject to one of three age groups (corresponding to “younger”, “middle-aged” and “older” categories) in relation to prediction accuracy, even with such a small feature set, is seen to be comparable to that reported elsewhere for the prediction of only a two-class age determination problem, which also used a very much larger and more diverse feature set.

Two principal novel points emerge here. First, these are the first reported results to show the reliable prediction of subject age from the iris patterning of an individual on the basis of more than two possible age-related categories. Second, we have shown how adopting an appropriate combination structure to support a multiclassifier configuration offers substantial benefits in this type of scenario, providing potential solutions to a range of important practical problems.

While we have demonstrated here some valuable principles on which to base the implementation of effective prediction mechanisms, future work will develop these ideas to determine how to optimise feature definition and selection so as to support increased resolution in the age prediction process.

References

[1] G. Guodong, F. Yun, C. R. Dyer, and T. S. Huang, "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," *Image Processing, IEEE Transactions on*, vol. 17, pp. 1178-1188, 2008.

[2] G. Xin, Z. Zhi-Hua, and K. Smith-Miles, "Automatic Age Estimation Based on Facial Aging Patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 2234-2240, 2007.

[3] F. Metze, J. Ajmera, R. Englert, U. Bub, F. Burkhardt, J. Stegmann, C. Muller, R. Huber, B. Andrassy, J. G. Bauer, and B. Littel, "Comparison of Four Approaches to Age and Gender Recognition for Telephone Applications," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 2007, pp. IV-1089-IV-1092.

[4] M. Nishimoto, Y. Azuma, N. Miyamoto, T. X. Fujisawa, and N. Nagata, "Subjective age estimation using speech sounds: Comparison with facial images," in *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, 2008, pp. 1900-1904.

[5] M. Fairhurst and M. Abreu, "An Investigation of Predictive Profiling from Handwritten Signature Data," in *Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on*, 2009, pp. 1305-1309.

[6] A. Sgroi, K. W. Bowyer, and P. J. Flynn, "The Prediction of Young and Old Subjects from Iris Texture," in *IAPR International Conference on Biometrics*, June 2013.

[7] V. Thomas, N. V. Chawla, K. W. Bowyer, and P. J. Flynn, "Learning to predict gender from iris images," in *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, 2007, pp. 1-5.

[8] S. Lagree and K. W. Bowyer, "Predicting ethnicity and gender from iris texture," in *Technologies for Homeland Security (HST), 2011 IEEE International Conference on*, 2011, pp. 440-445.

[9] X. Qiu, Z. Sun, and T. Tan, "Global Texture Analysis of Iris Images for Ethnic Classification " in *Advances in Biometrics*. vol. 3832, D. Zhang and A. Jain, Eds., ed: Springer Berlin / Heidelberg, 2005, pp. 411-418.

[10]M. Fairhurst and M. Erbilek, "Analysis of physical ageing effects in iris biometrics," *Computer Vision, IET,* vol. 5, pp. 358-366, 2011.

[11]M. Erbilek and M. Fairhurst, "A Methodological Framework for Investigating Age Factors on the Performance of Biometric Systems," in *The 14th ACM Workshop on Multimedia and Security*, Coventry, UK, 2012.

[12]M. Erbilek and M. Fairhurst, "Framework for managing ageing effects in signature biometrics," *Biometrics, IET,* vol. 1, pp. 136-147, 2012.

[13]Biosecure, "Biometrics for Secure Authentication, FP6 NoE, IST-2002-507634," http:\\www.biosecure.info, 2004.

[14]N. Cristianini. and J. Shawe-Taylor., "An introduction to support vector machines and other kernel-based learning methods," *Robotica*, vol. 18, pp. 687–689, 2000.

[15]S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed.: Prentice Hall PTR, 1998.

[16]F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review,* vol. 65, pp. 386-408, 1958.

[17]Y. Chauvin and D. E. Rumelhart, *Backpropagation: theory, architectures, and applications*: L. Erlbaum Associates Inc., 1995.

[18]S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *J. ACM*, vol. 45, pp. 891-923, 1998.

[19]L. Masek and P. Kovesi, "Recognition of Human Iris Patterns for Biometric Identification," Bachelor of Engineering The School of Computer Science and Software Engineering, The University of Western Australia, 2003.

[20]J. Quinlan, *C4.5: programs for machine learning*: Morgan Kaufmann Publishers Inc., 1993.

[21]M. C. Da Costa Abreu and M. Fairhurst, "Enhancing Identity Prediction Using a Novel Approach to Combining Hard- and Soft-Biometric Information," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* vol. 41, pp. 599-607, 2011.