An Investigation of Predictive Profiling from Handwritten Signature Data

Michael Fairhurst and Márjory Abreu
Department of Electronics
University of Kent
Canterbury, Kent CT2 7NT, UK
{M.C.Fairhurst, mcda2}@kent.ac.uk

Abstract

Although it has long been recognised that non-biometric factors (for example, general demographic characteristics) can have an impact on the performance of automated person identification systems, such information is not routinely adopted in most practical biometric processing. In forensic applications, however, such additional information may be exploited most productively, since typical scenarios require the prediction of a wide range of individual characteristics from a range of available samples, biometric and non-biometric. In many such situations, it may be very useful to predict characteristics short of actual identity, since this type of prediction can significantly increase the amount of evidence available. In this paper we provide some benchmarking data to demonstrate the extent to which typical non-biometric information might be effective in practice in this context, and illustrate the differential effects depending on the classifier adopted. In particular, however, we investigate experimentally how traditionally measured handwritten signature data might be used to predict non-biometric categories in a way which can be exploited a variety of practical application scenarios.

1. Introduction

The use of biometrics-based processing to determine identity or to authenticate a claimed identity is now well established. Approaches to biometric identification, however, are highly diverse, with the choice of specific modality, classifier/verifier type, optimisation technique and overall structural configuration all very much dependent on task domain characteristics and specific application requirements in relation to different metrics of performance [6].

As practical deployment of systems becomes more widespread, it is therefore increasingly important to understand how best to specify and implement an optimal design to meet individual application requirements, and this has stimulated a significant body of both theoretical and experimental research across a broad spectrum of interrelated topics in biometrics. For example, the relative merits of the many competing options in choosing a modality are now well understood ([21], [19], [6], [41] and [35]), while approaches to the optimisation of individual pattern classifiers have a long history which even significantly predates their application in biometrics ([33], [3], [37], [29], [32] and [20]). More recently, there has been much attention paid to the design of multibiometric systems, both in order to improve levels of performance, by increasing and diversifying the amount and nature of identity information available, and to increase resistance to attack, but also to offer greater flexibility and choice to systems users ([14], [38], [36], [40] and [12]).

While the primary objective of biometric processing is to establish individual identity, it is increasingly recognised that there is a close link between identification of individuals in situations commonly targeted by biometric systems and the longer-established area of forensic data processing. For example, latent fingerprints at a crime scene provide important information which might contribute to identification of the perpetrator, while comparison of handwriting samples (including, very commonly, a signature sample) can be a very powerful instrument in supporting crime investigations involving documents (a discussion about the relation between biometrics and forensic analysis in this context can be found in [31] and [17].

Often, the use of such information is just one part of a more substantial and extensive body of evidence in the identification process. However, part of such an analysis of evidence might well include the need to estimate the age of the owner of the specific piece of evidence under consideration. For example, evidence of the age of a signer might be very valuable in evaluating and integrating evidence from multiple sources as part of a general situational analysis to eliminate individuals from an enquiry. However, predicting an individual's age has wider application, such as establishing whether that individual meets legal requirements for partic-

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.94

ular entitlements, and this might be especially important in the case of young individuals claiming to be older than they really are.

An important reason for this is also that it is well-known that the physical ageing process has, to a greater or lesser extent, a direct impact on the biometric measurements taken from all individuals, with very important practical consequences both for attainable performance and on the requirements for template updating. There have been a number of studies reporting research on the effects of age on biometric system performance, and results have been reported, for example, on the implications of ageing on fingerprint image quality, on the generation of the handwritten signature, and so on ([23], [16], [9], [10] and [1]).

The literature reports many studies related to the prediction of gender and age using modalities such as the face ([30] and [24]), voice [8] and iris [39]. However, the possibilities for using the handwritten signature for prediction of non-biometric information is not nearly so well developed.

In this paper we will both develop our understanding of how subject age affects the performance of a biometric system, using the handwritten signature for illustration but, importantly, as the principal focus of the paper we will report some work to give a significantly new perspective on the exploitation of non-biometric information, using our analysis to study the predictive capabilities of biometric systems in relation to individual characteristics. Specifically in this paper we study the relationship between subject age and performance in a biometrics-based identification task.

2. Experimental Infrastructure

In the experimental work to be reported here we use handwritten signature samples collected as part of a multimodal database, where each of 79 users provided their information in two sessions. The data used were collected in the Department of Electronics at the University of Kent in the UK as part of a Europe-wide project undertaken by the BioSecure Network of Excellence [11].

In the handwritten signature, the samples were collected using an A4-sized graphics tablet with a density of 500 lines per inch. Each user donated 15 genuine samples of a handwritten signature. From each sample, 16 commonly used features (a mix of static and dynamic measurements) were extracted [16] and can be seen in Table 1.

For simplicity we have partitioned the writer population into just three age groups taking into account the age bands most commonly used in the literature ([16], [2], [8], [26] and [27]). The age groups adopted are < 25, 25-60 and $>$ 60, providing the opportunity to explore age-related effects while maintaining a good representation of signers in each sub-group.

In order to investigate and evaluate performance in our experimental study, a range of different individual classifiers were selected in order to cover as many different classification approaches as possible, which would also form the base components of a multiclassifier system (MCS) used in the later part of our experimentation. These are as follows:

Feature	Type
Execution Time	Dynamic
Pen Lift	Dynamic
Signature Width	Static
Signature Height	Static
Height to Width Ratio	Static
Average Horizontal Pen Velocity in X	Dynamic
Average Horizontal Pen Velocity in Y	Dynamic
Vertical Midpoint Pen Crossings	Dynamic
Azimuth	Dynamic
Altitude	Dynamic
Pressure	Dynamic
Number of points comprising the image	Static
Sum of horizontal coordinate values	Static
Sum of vertical coordinate values	Static
Horizontal centralness	Static
Vertical centralness	Static

Table 1. Signature features

- Multi-Layer Perceptron (MLP) [18].
- Fuzzy Multi-Layer Perceptron (FMLP) [5].
- Radial Basis Function Neural Network (RBF) [4].
- Optimised IREP (Incremental Reduced Error Pruning) (JRip) [15].
- Support Vector Machines (SVM) [28].
- Decision Trees (DT) [34].
- K-Nearest Neighbours (KNN) [13].

The multiclassifier systems (MCS) adopted in the present study are organised as ensembles (each individual base classifier will predict the sample class and an appropriate combination method is invoked to determine the overall output of the system). The combination methods considered include a range of different possible strategies. In the present work we have chosen one selection-based method and three fusion-based methods, each of which can be briefly described as follows:

- Dynamic Classifier Selection based on local accuracy class (DCS-LA) [42]: This selection-based method uses local analysis of competence to nominate a classifier to label an input.
- FuzzyMLP [5]: This fusion-based method is a neural classifier which incorporates fuzzy set theory into a multi-layer Perceptron framework, and results from the direct fuzzyfication in the network level in the MLP, in the learning level, or in both.

- Naive Bayesian Learning [7]: This fusion-based method is a simple probabilistic classifier based on the application of Bayes theorem with the assumption of strong independence.
- Majority Voting [22]: This non-linear fusion-based method takes into account only the top outputs of the component experts. The outputs of the classifiers are represented in a winner-take-all form (for each classifier, the output of the winner is 1 and the remaining outputs are 0) and the weights for all the experts are equal to 1.

In order to evaluate the robustness of the classification system, the cross validation methodology was chosen. Tenfold cross validation has been shown to be statistically sound in evaluating the performance of a typical classifier [25].

3. Predicting Age from Biometric Data

Against this background, our current study has therefore explored specifically the predictive capability of biometric data in establishing the age of an individual. Taking the classifiers in our initial pool, we have investigated the accuracy with which the age grouping of individuals can be predicted within our handwritten signature database.

Table 2 shows the experimental results obtained. A consideration of the error rates returned by our classifiers is rather instructive here. First, we note that the capacity for correctly predicting the age grouping of the sample donors is encouragingly high, at least within the constraints of these experiments, when using the handwritten signature as the source data. We note also that the largest error rates occur in predicting the age of subjects whose true age falls within the highest age band (in this case > 60).

Next, it is useful to analyse performance in greater detail, and to do this we will focus on just one of the classifiers, in this case the FMLP (italicised rows of Table 2) which performed relatively well. However, let us now ask a question about the predictive capability of the system in relation to a particular likely application area. For example, how does the system perform in predicting whether an individual signature susbstantiates an expectation that a suspected signer is over 25 years of age?

The results show that the system returns a prediction of being over 25 years of age in the case of those with a real age less than 25 years with an error of only 1.134% (= 0.964+0.170). Thus, in a typical practical scenario, the predictive capability of the biometric data appears to be particularly promising. It is also interesting to note that the most common incorrect prediction for an individual who is actually younger than 25 years would place him/her in the age grouping at the high end of the age range (i.e. in the > 60 group) with approximately the same probability, errors towards the middle age range accounting for the overall performance difference.

Looking at the inverse of this situation, we might consider the case where we wish to ascertain whether an individual signer is likely to be older than 60 years. In this situation the error rates are slightly higher (for example, 2.325%). Here, however, the error categories are more equally dispersed, with a slightly higher probability of being categorised as belonging to the middle (25-60) age group than the youngest (< 25) group.

Thus, our results tend to confirm findings reported from other, differently focused, studies in that it is clear that the elderly are the most likely group to exhibit problems in relation to identification based on biometric data processing. However, in predicting age from biometric data, while the signature appears to offer a strong predictive option in an unrestricted scenario, the evidence here suggests that if checking age in a scenario where meeting a minimum age requirement at the lowest end of the age range is the principal aim, then there is less to choose in performance between the classifiers than is the case if checking is occurring against a minimum age at the other end of the age scale. The error categories are also more distinct, and perhaps therefore more easily corroborated by other means when checking in the lower age group.

Finally, we have carried out experiments with a multiclassifier configuration operating on our handwritten signature database. Here we use a combination of all seven individual base classifier algorithms and explore combination algorithms based on FuzzyMLP, Majority Vote, DCS-LA and Naive Bayes. The results can also be seen in Table 2. We can see that this multiclassifier approach has marginally improved performance in the predictive capacity of the signature in forensics-related tasks (0.71% = 5.67%-4.96%). Again, in the scenario where an individual over age 60 is predicted as being younger than 60, the errors are marginally lower than with the single classifier solutions, at 2.04% (= 0.099%+1.949%).

Of course, these results are relatively preliminary, and indicate only general trends without claims of statistical significance. It is also the case that we have considered only three age bands, though we have related the results to practical scenarios potentially of real-world significance. The multiclassifier results are perhaps unexpected although caution should again be exercised in interpreting these results without some consideration of optimisation strategies.

4. Conclusions

We have reported some recent results from our current research programme which is exploring a number of important issues concerning the relationships between single

Individual Classifiers	Err Mean$\pm$ Stan Dev	<25 (predicted)		26-60 (predicted)		>60 (predicted)	
		25-60 (real)	>60 (real)	<25 (real)	>60 (real)	<25 (real)	25-60 (real)
MLP	6.39$\pm$2.10	0.058	1.220	1.086	1.406	0.192	2.428
FMLP	*5.67$\pm$2.08*	*0.051*	*1.083*	*0.964*	*1.247*	*0.170*	*2.155*
RBF	6.87$\pm$2.12	0.062	1.312	1.168	1.511	0.206	2.611
Jrip	7.22$\pm$2.42	0.065	1.379	1.227	1.588	0.217	2.744
SVM	7.99$\pm$2.31	0.072	1.526	1.358	1.758	0.240	3.036
DT	9.37$\pm$2.43	0.084	1.790	1.593	2.061	0.281	3.561
KNN	9.98$\pm$2.37	0.090	1.906	1.697	2.196	0.299	3.792
Multiclassifier Systems	Err Mean$\pm$ Stan Dev	<25 (predicted)		26-60 (predicted)		>60 (predicted)	
		25-60 (real)	>60 (real)	<25 (real)	>60 (real)	<25 (real)	25-60 (real)
FMLP	*4.96$\pm$2.01*	*0.035*	*0.992*	*0.843*	*1.042*	*0.099*	*1.949*
Majority Vote	5.19$\pm$1.97	0.036	1.038	0.882	1.090	0.104	2.040
DCS-LA	5.01$\pm$2.14	0.035	1.002	0.852	1.052	0.100	1.969
Naive Bayes	5.52$\pm$2.29	0.039	1.104	0.938	1.159	0.110	2.169

Table 2. Error Mean, Standard Deviation, False Positive and False Negative Rates of the Individual Classifiers and Multiclassifier systems

classifier, multiclassifier and multibiometric solutions, and system implementations which may involve biometric data alone or both biometric and non-biometric data, and where many optional configurations can be considered. We have also demonstrated the potential for such work to provide new opportunities for evidence gathering in forensic applications.

We have focused specifically on the relationship between subject age and the nature of the biometric data generated. While we have illustrated how subject age can affect differentially the performance of a biometric system adopting handwritten signature and differing processing structures, we have also investigated the complementary issue of real significance in the present context, that of seeking some insight into the degree to which biometric data can be a predictor of subject age.

Our experimental results provide some valuable indicators of how to optimise a processing system configuration in this type of practical scenario, and provide a useful basis on which to develop our study, both with respect to more detailed analysis of the principles which have emerged and how to increase the resolution of prediction, but also in relation to their extension to multibiometric structures.

5. Acknowledgement

The authors gratefully acknowledge the financial support given to Mrs Abreu from CAPES (Brazilian Funding Agency) under grant BEX4903-06-4.

References

[1] M. Abreu and M. Fairhurst. Analyzing the impact of non-biometric information on multiclassifier processing for signature recognition applications. In *IEEE Second International Conference on Biometrics: Theory, Applications and Systems (BTAS'08)*, Washington DC., 2008.

[2] J. Ajmera. Effect of age and gender on lp smoothed spectral envelope. pages 1–4, June 2006.

[3] M. Allah. Artificial neural networks based fingerprint authentication with clusters algorithm. *Informatica (Slovenia)*, 29(3):303–308, 2005.

[4] M. Buhmann and M. Buhmann. *Radial Basis Functions*. Cambridge University Press, New York, NY, USA, 2003.

[5] A. Canuto. *Combining Neural Networks and Fuzzy Logic for Aplications in Character Recognition.* PhD thesis, Department of Electronics, University of Kent, Canteburry, UK, May 2001. Adviser-Michael C. Fairhurst.

[6] S. Dass, Y. Zhu, and A. Jain. Validating a biometric authentication system: Sample size requirements. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):1902–1319, 2006.

[7] C. Elkan. Boosting and naive bayesian learning. Technical report, 1997.

[8] F. M. et al. Comparison of four approaches to age and gender recognition for telephone applications. *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 4:IV–1089–IV–1092, April 2007.

[9] J. A.-F. et al. On the relation between biometric quality and user-dependent score distributions in fingerprint verification. In *Workshop on Multimodal User Authentication, MMUA*, May 2006.

[10] J. F.-A. et al. Exploiting general knowledge in user-dependent fusion strategies for multimodal biometric verification. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and*

Signal Processing, ICASSP, volume 5, pages 617–620, May 2004.
[11] J. O.-G. et al. Software tool and acquisition equipment recommendations for the three scenarios considered. Technical Report Report No.: D6.2.1. Contract No.: IST-2002-507634, Jun 2006.
[12] K. K. et al. Reliability-based decision fusion in multimodal biometric verification systems. *EURASIP Journal on Advances in Signal Processing*, 2007(1):74–74, 2007.
[13] S. A. et al. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM*, 45(6):891–923, 1998.
[14] J. Fierrez-Aguilar, J. Ortega-Garcia, and J. Gonzalez-Rodriguez. Fusion strategies in multimodal biometric verification. In *ICME '03: Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 3 (ICME '03)*, pages 5–8, Washington, DC, USA, 2003. IEEE Computer Society.
[15] J. Furnkranz and G. Widmer. Incremental reduced error pruning. In *ICML*, pages 70–77, New Brunswick, NJ, 1994.
[16] R. Guest. Age dependency in handwritten dynamic signature verification systems. *Pattern Recognition Letter*, 27(10):1098–1104, 2006.
[17] J. Haggerty, D. Llewellyn-Jones, and M. Taylor. Forweb: file fingerprinting for automated network forensics investigations. In *e-Forensics '08: Proceedings of the 1st international conference on Forensic applications and techniques in telecommunications, information, and multimedia and workshop*, pages 1–6, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
[18] S. Haykin. Neural networks: a comprehensive foundation. *Knowl. Eng. Rev.*, 13(4):409–412, 1999.
[19] K. Kroeker. Graphics and security: Exploring visual biometrics. *IEEE Comput. Graph. Appl.*, 22(4):16–21, 2002.
[20] Y. Kuan, A. Teoh, and D. Ngo. Secure hashing of dynamic hand signatures using wavelet-fourier compression with biophasor mixing and 2n discretization. *EURASIP J. Appl. Signal Process.*, 2007(1):32–32, 2007.
[21] A. Kumar, D. Wong, H. Shen, and A. Jain. Personal verification using palmprint and hand geometry biometric. In *Proc. of 4th International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA)*, pages 668–678, Guildford, UK, 2003.
[22] L. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.
[23] M. Liwicki, A. Schlapbach, P. Loretan, and H. Bunke. Automatic detection of gender and handedness from on-line handwriting. In *Proc. 13th Conf. of the Graphonomics Society*, pages 179–183, 2007.
[24] F. Matta, U. Saeed, C. Mallauran, and J. Dugelay. Facial gender recognition using multiple sources of visual information. *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, pages 785–790, Oct. 2008.
[25] T. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.
[26] T. Miyoshi and M. Hyodo. Aging effects on face images by varying vertical feature placement and transforming face shape. volume 2, pages 1548–1553, Oct. 2006.
[27] S. Modi, S. Elliott, J. Whetsone, and H. Kim. Impact of age groups on fingerprint recognition performance. pages 19–23, June 2007.
[28] C. Nello and S. John. An introduction to support vector machines and other kernel-based learning methods. *Robotic*, 18(6):687–689, 2000.
[29] M. Ong, T. Connie, A. Jin, and D. Ling. A single-sensor hand geometry and palmprint verification system. In *WBMA '03: Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 100–106, New York, NY, USA, 2003. ACM.
[30] E. Patterson, A. Sethuram, M. Albert, K. Ricanek, and M. King. Aspects of age variation in facial morphology affecting biometrics. *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6, Sept. 2007.
[31] V. Pervouchine and G. Leedham. Extraction and analysis of forensic document examiner features used for writer identification. *Pattern Recogn.*, 40(3):1004–1013, 2007.
[32] O. Polat. Hand geometry identification without feature extraction by general regression neural network. *Expert Syst. Appl.*, 34(2):845–849, 2008.
[33] S. Prabhakar. *Fingerprint classification and matching using a filter bank*. PhD thesis, East Lansing, MI, USA, 2001. Adviser-Anil K. Jain.
[34] J. Quinlan. *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
[35] C. Riley, H. McCracken, and K. Buckner. Fingers, veins and the grey pound: accessibility of biometric technology. In *ECCE '07: Proceedings of the 14th European conference on Cognitive ergonomics*, pages 149–152, New York, NY, USA, 2007. ACM.
[36] A. Ross and A. Jain. Multimodal biometrics: an overview. In *Proceedings of the 12th European Signal Processing Conference*, pages 1221–1224, 2004.
[37] A. Ross, A. Jain, and J. Reisman. A hybrid fingerprint matcher. In *ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 3*, page 30795, Washington, DC, USA, 2002. IEEE Computer Society.
[38] R. Snelick, M. Indovina, J. Yen, and A. Mink. Multimodal biometrics: issues in design and testing. In *ICMI '03: Proceedings of the 5th international conference on Multimodal interfaces*, pages 68–72, New York, NY, USA, 2003. ACM.
[39] V. Thomas, N. Chawla, K. Bowyer, and P. Flynn. Learning to predict gender from iris images. *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–5, Sept. 2007.
[40] D. Toledano, R. Pozo, A. Trapote, and L. Gómez. Usability evaluation of multi-modal biometric verification systems. *Interact. Comput.*, 18(5):1101–1122, 2006.
[41] P. Wang and S. Yanushkevich. Biometric technologies and applications. In *AIAP'07: Proceedings of the 25th conference on Proceedings of the 25th IASTED International Multi-Conference*, pages 226–231, Anaheim, CA, USA, 2007. ACTA Press.
[42] K. Woods, W. Kegelmeyer Jr., and K. Bowyer. Combination of multiple classifiers using local accuracy estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):405–410, 1997.