Performance versus Computational Complexity Trade-off in Face Verification

Thirimachos Bourlai, Kieron Messer, and Josef Kittler

University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom,
{t.bourlai, k.messer, j.kittler}@surrey.ac.uk

Abstract. We discuss the implementation of a face verification system on smart card and study the trade-off between performance and computational complexity. In order to establish the system limitations, a study was performed on BANCA and XM2VTS databases. The results of the experiments performed show that the choice of a fixed precision data type does not affect system performance very much but can speed up the verification process. Moreover, the use of less than 8 bits per pixel gray-scale image resolution does not necessarily result in a degradation of system performance. We show that, for both databases, image resolution can be reduced without degrading system performance. However, the actual optimal settings will depend on application scenario.

1 Introduction

The design of automatic personal identity authentication systems based on facial images is a very challenging task [11][12] with many promising applications in the field of security. Such biometric systems [5] provide an effective and inherently more reliable way to carry out personal verification of individuals, seeking access to physical or virtual locations. A higher level of security is achieved by exploiting user specific biometric characteristics that it would be difficult to be appropriated by impostors. A secure personal ID system should address government and business policy issues and individual privacy concerns. Smart cards combined with biometric technology [10][8][9] is considered to be a challenging and promising solution in secure personal ID systems[6].

In a conventional face verification system, a database is used to store all biometric templates and users information and the biometric data captured by a camera is transmitted to a central computer where all the image processing and decision making takes place. Although the idea of a centralised online biometric system design is acceptable in many applications, it raises both privacy and security issues. For that reasons, a smart card based face verification system is proposed, demonstrated in *figure 1*. By combining a pin code with on-card storage of the biometric template, we can meet the privacy requirements. Security is also improved since information and processes are protected within the system. Other advantages are improved system performance and system evolvement.

However, with the use of smart cards there are restrictions to be considered like the available CPU computational power, storage capacity and bandwidth[4]. Therefore, a challenge in the implementation of a smart card face verification system is to optimise

the trade-off between the fundamental performance of face verification and the computational complexity of the experimental platform used (server or smart card). Issues to be considered are image quality, system delay, complexity, efficiency, error tolerance (accuracy), compatibility, scalability, memory management, data type dependence and finally, architecture of the system where the code is implemented. The evaluation of such a system is also dependent on the face database used to perform the experiments.

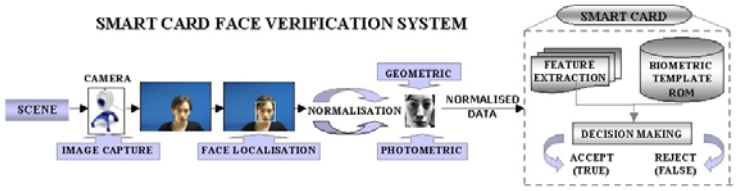

Fig. 1. Proposed smart card face verification system

In this paper, we discuss the implementation of a smart card face verification system and we study the trade-off between performance and computational complexity. In order to establish the limitations of such a system, a study was performed, selecting the performance and the computational cost of using different data types on the smart card and on the server, the bit precision of the matching function parameters, the gray-scale resolution of probe images (from 8bpp currently, down to 1bpp) and the image resolution that dominates system speed and memory management.

The rest of the paper is organised as follows. In the next section, the basic face verification process will be covered. In Section 3, the experimental setup and results will be presented. Finally, in Section 4 some conclusions are made.

2 Face Verification System

The face verification method adopted for the implementation on a smart card is the client specific linear discriminant analysis technique, which combines face representation and decision making into a singe step, requiring a template of the size of the input image (see *figure 1*). The overall face verification system involves face detection, photometric normalisation and finally the verification test. All but the last processing step are carried out in the host system. The photometrically normalised image is then transmitted to the smart card where the user biometric template is stored and the verification score computed as well as the final decision taken. The original resolution of the image data in both is 720x576. The experiments were performed with a relatively low resolution face images, namely 55x51. This resolution was used initially as a reference for our study. After detecting the face and localising the eye centres, a simple rotation and scaling of the eye positions onto two fixed points was used for geometric transformation. The photometric normalisation is achieved by a homomorphic filter and histogram equalisation. For the feature extraction stage, a *PCA* model is built to achieve a dimensionality

reduction and then an *LDA* model is produced to get the overall client *i* specific linear discriminant transformation $\mathbf{a}_i$, which defines the client specific fisher face for testing the claimed identity. The decision making stage produces a score, which defines how close the probe of the claimed identity is to the class of impostors. The thresholds in this stage have been determined based on the equal false rejection (FR) and acceptance (FA) error rates (EER).

The *Client Specific Fisherface* (CSLDA) representation adopted for smart card face verification contrasts with the conventional LDA method, which involves multiple fisherfaces. Apart from its good performance, the method is advantageous in the case of open-set scenarios when new clients are enrolled without system retraining, since it requires only a matrix multiplication of the client mean vector. Moreover, the client enrolment is insulated from the enrolment of other clients. Therefore, it becomes possible to use other than centralised architecture for the face verification system and the smart card processing becomes a reality without any need to restrict the representation framework, and as a result the representation capacity of the system. Finally, the speed of probe testing is more than two orders of magnitude faster than that achieved by the PCA and LDA methods[2], as the *CSLDA* method involves only a single fisher face $\mathbf{a}_i$ per client stored on the smart card.

In the verification stage we use the metric (distance) d_c:

$$d_c = \left| \mathbf{a}_i^T \mathbf{z} - \mathbf{a}_i^T \mu_i \right| \tag{1}$$

where $\mathbf{z}$ is the probe image and μ_i is the client mean.
Now, if $d_c \leq t_c$, then we accept the claim, where t_c is a pre-computed threshold.

3 Experimental Setup and Results

For the purpose of this study, BANCA (open-set protocol)[1] and XM2VTS (closed-set protocol)[7] face databases were used in the experiments. We are dealing with single-modality experiments. From the sets containing the face images, training, evaluation and test set is built. The training set is used to construct client models; the evaluation set produces client and impostor access scores (used to compute a client-specific or global threshold that determines acceptance or rejection of a person); and the test set is selected to simulate realistic authentication tests where impostor's identity is unknown to the system. The threshold is set to satisfy certain performance levels on the evaluation set. Finally, the performance measures of the verification system are the FA and FR rates on the test set.

The smart card used in these experiments was provided by Sharp Laboratories, Oxford, UK. It boasts a 13.5MHz processor, 1Mbyte of EEPROM, 8KBytes of RAM, a cryptographic co-processor, does not have a floating point co-processor and operates in both contact and contactless modes. In these experiments we used it in contact mode which has a data transfer rate of 76.8Kbits per second. The transmission rates of data between the server and card is fairly slow. Therefore, the amount of data being sent to and read from the smart card (e.g. a biometric template or a facial image) must be kept to a minimum. Finally, the amount of RAM available on the smart card is limited,

which means all the data can not be kept in memory for the calculations and the ROM must be used as a cache.

In the *first experiment*, the relationship between computational cost of verification on the smart card and the use of different data types was examined. We measured the time required to perform certain on-card extended data operations. These experiments demonstrated that the use of integers instead of floating point numbers can speed up the overall verification performance by over a factor of 6. The reason is the non-availability of a floating point co-processor on the card and the use of a simulated floating point unit instead. Fixed point numbers (FPN) proved to be a risky choice when used exclusively and the choice of double precision on the card was absolutely prohibiting in terms of computational costs. In order to perform the same experiment on the server, a specific fixed point library was built. Interestingly, the use of the generated n-bit precision data type instead of integers resulted in only 12% loss in the overall computational cost. Ideally, the use of integers in both the server and the smart card would be the best solution in terms of speed but not in terms of precision. However, the use of fixed n-bit precision data type on the server and the use of integers on the smart card is expected to significantly increase the overall speed when on-card verification is performed.

In the *second experiment*, we investigated the trade-off between performance and n-bit precision for the verification function parameters when using fixed point arithmetic for authentication. These parameters are the client specific LDA transformation matrix **a**, the client mean vector μ_i and the global mean $\mu = \sum_{j=1}^{N} \mathbf{z}_i$, where N is the size of the training set and $\mathbf{z}_i$ are the training images. The basic idea behind that was to change the precision of the CSLDA transformation that is actually sent on the smart card for on-card verification based on the distance metric given in equation (1). In order to evaluate the system performance the following formula was computed:

$$PPD = \sum_{i=1}^{2} \frac{\left|\frac{FA\_Gi_{init} - FA\_Gi_{new}}{FA\_Gi_{init}}\right| + \left|\frac{FR\_Gi_{init} - FR\_Gi_{new}}{FR\_Gi_{init}}\right|}{4} * 100\% \qquad (2)$$

where *PPD* stands for *Percentage of Precision Degradation* and G_1, G_2 stand for the two groups we are using for training and testing alternatively. Within the formula, each percentage represents the absolute percentage difference between the original value of the (false acceptance)/(false rejection) rate and the new (false acceptance)/(false rejection) rate after applying the n-bit precision transformation. However, Half Total Error Rate (HTER) was also used as the basic metric of performance. In *table 1* we can see the best selected n-bit precision in terms of *PPD* when applying this to all BANCA protocols. Moreover, in *figure 2* we can see the overall results in terms of performance and n-bit precision in some selected BANCA protocols and the average results in all protocols. An interesting observation is that the use of a n-bit precision data type above 13 can result in an increased loss of performance due to quantisation errors. By trying to optimise the results for all cases and to find the best value of *n* for all BANCA protocols in terms of the minimum *average PPD*, we conclude that by using *13-bit precision* we achieve only a 0.32% loss in overall system performance . The overall average results are graphically represented in *figure 3*.

In the *third experiment*, we investigated the effect on performance by altering the gray-scale pixel resolution of the (55x51) normalised probe images in the training set.

Table 1. Best cases by using n-bit precision in BANCA

Protocol	*n-bit* precision	HTER init	HTER new	PPD
MC	12,13	0.07932	0.07932	0
MD	13	0.10737	0.10737	0
MA	13	0.09775	0.09775	0
UD	10	0.1915	0.19166	0.68
UA	13	0.26634	0.26682	0.2232
P	13	0.21927	0.2196	0.2246
G	10	0.07927	0.07879	1.03756

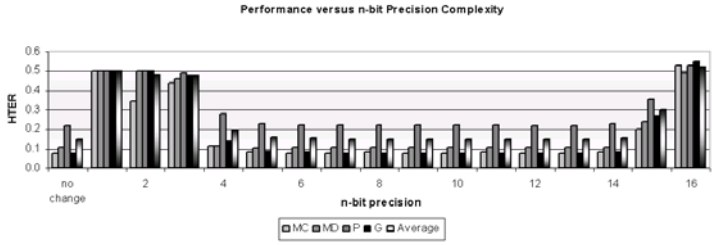

Fig. 2. Graphical representation of the effects of n-bit precision on the performance in MC, MD, P, G and average BANCA protocols

Since an 8 bit camera is used, the initial performance of our system was measured by using 8 bits per pixel for each face image. Then, the gray-scale resolution was reduced by a factor of 1bpp each time before building the PCA and LDA model. All seven cases of BANCA protocols were tested and the results in all protocols are given in *table 2*. One interesting observation is that the use of 4-bit gray-scale pixel resolution

Table 2. BANCA protocol performance results using different gray-scale pixel resolution

n-bit	MC	MD	MA	UD	UA	P	G	Average HTER
8bpp	0.07932	0.1073	0.09775	0.1915	0.2663	0.2192	0.0792	0.1487
7bpp	0.07932	0.1073	0.09935	0.1911	0.2665	0.2194	0.0792	0.1489
6bpp	0.08028	0.1078	0.09935	0.1915	0.2653	0.2204	0.0792	0.1491
5bpp	0.07884	0.1062	0.09759	0.1889	0.2639	0.2211	0.0785	0.1478
4bpp	0.07788	0.1041	0.09647	0.1915	0.2642	0.2205	0.0772	0.1474
3bpp	0.07628	0.1083	0.09727	0.1903	0.2663	0.2195	0.0808	0.1484
2bpp	0.07003	0.1169	0.10224	0.2139	0.2838	0.2322	0.0864	0.1579
1bpp	0.0915	0.2355	0.14033	0.3153	0.3685	0.3082	0.1637	0.2319

yields an improved overall performance by 0.95% compared to the initial reference images. This can be explained by the fact that pixel resolution normalisation introduces quantisation errors that can affect the computation of the error rates and thresholds.

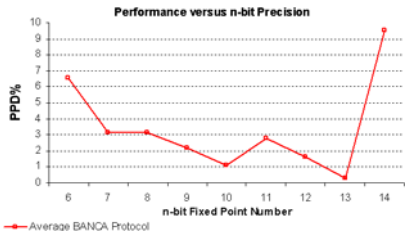

Fig. 3. Average BANCA protocol performance when using n-bit fixed point numbers ($n \in [6,..,14]$)

Since we know that in verification systems there is always a trade-off between FA and FR rates, there are cases where these rates can change in such a way that, on average, the overall verification performance is positively affected (that is, for example, the FA rate degrades while FR improves even faster than FA degrades).

In the *last experiment*, we investigated the trade-off between performance and image resolution that can be exploited to minimise the computational costs in both BANCA and XM2VTS databases. In this study, the initial raw face images were geometrically and photometrically normalised to a predefined size. As a reference, image resolution 55x51 was selected. In *figure 4* the experimental results are presented. In the top row, we can see the detailed results in both databases, clearly for all protocols and configurations. In XM2VTS (*figure 4a*), better overall results are achieved and particularly configuration II seems to achieve a more desirable demonstration of system performance. This is probably because the training set in configuration II is larger than that in configuration I. In the case of BANCA (*figure 4b*), the overall trend of the results does not change, although the overall results are worse than those of XM2VTS. In the bottom row, a representation of the average results obtained on each database are shown (see *table 3*). In (*figure 4c*) the average results on XM2VTS are presented. Note that going from 55x51 to 20x18 we achieve an overall 0.65% improvement in performance while at the same time speed is increased 87,6%. However, the best average performance is achieved by using 30x28 resolution where we achieve 5.8% improvement in performance while at the same time the speed is increased 71.1%. In comparison, in *figure 4d* the average results on BANCA are presented. However, in this case, in order to improve the system performance at 17.8%, we have to tolerate that the speed will be halved. This major difference in the two databases is due to their different protocol configuration; Basically, in XM2VTS we have more controlled conditions and fully frontal faces, whereas in BANCA different condition scenarios are presented. It seems that closed-set protocols perform better and a redesign of the verification system when new clients are added in an open-set protocol will result in better results.

4 Discussion and Conclusions

Optimisation of a smart card face verification system design in terms of performance and computational complexity is a very complex process, with many key factors to consider. The specification includes gray-level and spatial image resolution, the overall

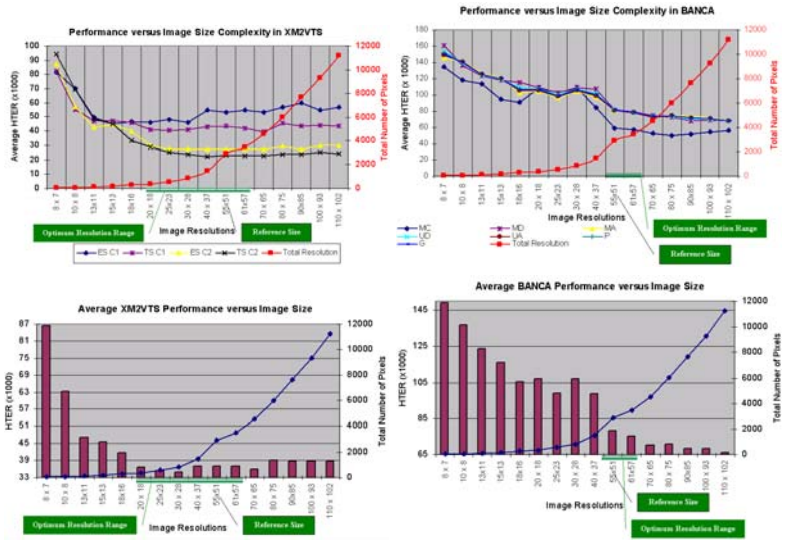

Fig. 4. The effect on performance in terms of *HTER* by using different image resolution for probe images in the evaluation/test set (in both configurations) of XM2VTS (a) and in all protocols in BANCA (b). The same effect is presented when we average the results in XM2VTS (c) and in BANCA (d).

system speed and performance requirements. The number of bits to be transmitted from the server to the smart card may have to be further restricted when fusion methods need to be incorporated onto the smart card [3] and therefore an increased number of biometric templates have to be stored on the card.

Experiments showed that the use of fixed point arithmetic can speed up the template matching on the card. However, the ideal solution is to use integers both on the card and the server. Using less than 8bpp grey-scale image resolution for the normalised face images does not necessarily result in a degradation of the system performance. This allows for fewer bytes of data to be sent to the smart card. There is a trade-off between performance and image resolution range that dominates system speed and memory management. In both databases we can improve the system performance when decreasing the gray-scale resolution, but not when decreasing image size. Different results were achieved on the BANCA in contrast to XM2VTS database suggesting that the use of different operating scenarios for system evaluation may call for different optimum operating point.

Acknowledgements

The authors would like to acknowledge the support received from OmniPerception Ltd, Sharp Laboratories, U.K and EPSRC under grant GR/S46543/01(P).

Table 3. Overall average results in both BANCA and XM2VTS

Image Resolution	Total no of Pixels	BANCA Average HTER	XM2VTS Average HTER
8 x 7	56	0.14933	0.08649
10 x 8	80	0.13667	0.06309
13x11	143	0.12357	0.04714
15x13	195	0.11602	0.04546
18x16	288	0.10531	0.04168
20 x 18	360	0.10692	0.03664
25x23	575	0.09896	0.03544
30 x 28	840	0.10692	0.03475
40 x 37	1480	0.09868	0.03695
55x51	2907	0.07824	0.03688
61x57	3477	0.07533	0.03687
70 x 65	4550	0.07048	0.03584
80 x 75	6000	0.07075	0.03893
90x85	7650	0.06847	0.03873
100 x 93	9300	0.06838	0.03870
110 x 102	11220	0.06640	0.03864

References

1. E. Bailly-Baillire, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz, and J.-Ph Thiran, 'The banca database and evaluation protocol', *AVBRA*, (2003).
2. P.N. Belhumeur, J. Hespanha, and D. J. Kriegman, 'Eigenfaces vs. fisherfaces: Recognition using class specific linear projection', *IEEE PAMI*, **19**, 45–58, (1996).
3. J. Czyz, S. Bengio, C. Marchel, and L. Vanderdorpe, 'Scalability analysis of audio-visual person identity verification', *AVBRA*, 752–760, (June 2003).
4. Jacek Czyz and Luc Vandendorpe, 'Evaluation of lda-based face verification with respects to available computational resources', in *PRIS*, (April 2002).
5. J. L. Dugelay, J. C. Junqua, C. Kotropoulos, R. Kuhn, F. Perronnin, and I. Pitas, 'Recent advances in biometric person authentication', *ICASSP (special session on biometrics), Orlando, Florida*, (May 2002).
6. Copyright Smart Card Alliance Inc., 'Smart card and biometrics in privacy-sensitive secure personal identification systems', *A Smart Card Alliance white paper, http://www.datakey.com/resource/whitePapers/Smart_Card_Biometric_paper.pdf*, (May 2002).
7. K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, 'Xm2vtsdb: The extended m2vts database', *AVBRA*, 72–77, (March 1999).
8. M. Osborne and N. K. Ratha, 'A jc-bioapi compliant smart card with biometrics for secure access control', *AVBRA*, 903–910, (June 2003).
9. W. Rankl and W. Effing, 'Smart card handbook', *John Wiley & Sons*, (2000).
10. R. Sanchez-Reillo and C. Sanchez-Avila, 'Fingerprint verification using smart cards for access control systems', *IEEE AESM*, **17**(9), 12–15, (2002).
11. M. Turk and A. Pentland, 'Eigenfaces for recognition', *Cognitive Neuroscience IEEE PAMI*, **3**(1), 71–86, (1991).
12. W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips, 'Face recognition: A literature survey', *UMD CfAR Technical Report CAR-TR-948*, (2000).