Improving feature-domain super-resolution for iris recognition

Kien Nguyen, Clinton Fookes, Sridha Sridharan and Simon Denman
Image and Video Research Laboratory, Queensland University of Technology
2 George Street, GPO Box 2434, Brisbane, Queensland Australia 4001
Email: kien.nguyenthanh@student.qut.edu.au,{c.fookes,s.sridharan,s.denman}@qut.edu.au

***Abstract*—Super-resolution techniques have been considered for iris recognition to improve the recognition performance of the system. However, most existing super-resolution approaches proposed for the iris biometric super-resolve pixel intensity values, which become unstable when lighting conditions vary in the image sequence. Recently, a feature-domain super-resolution approach for iris video sequences using Principle Component Analysis (PCA) features has been proposed and has been shown to outperform pixel-domain super-resolution approaches in recognition performance. This paper investigates improvements to this approach using Linear Discriminant Analysis (LDA) and conjugate gradient optimisation method. LDA features are employed to improve classification of data classes when compared to PCA. Conjugate gradient optimisation is employed to improve the accuracy of the convergence of the estimated high-resolution features, when compared to steepest descent optimisation. Through experimentation, we demonstrate that that LDA features outperform PCA features for super-resolution, and we show the superior performance of conjugate gradient optimisation in the estimation of high-resolution iris features.**

I. Introduction

Super-resolution techniques have previously been employed to address the low resolution problems of imaging systems [1]. There are two variants of super-resolution approaches: reconstruction-based and learning-based [1]. Reconstruction-based approaches fuse the sub-pixel shifts among multiple low resolution images to obtain an improved resolution image. Alternatively, learning-based approaches model high-resolution training images and learn prior knowledge to constraint the super-resolution process [1].

Recently, super-resolution techniques have been investigated for biometric systems. A number of super-resolution techniques have been successfully developed for face [2], [3], [4] and iris [5], [6], [7]. However, one main concern raised by both Gunturk et al. [2] and Kien et al. [8] is how to apply super-resolution for a specific biometric modality effectively to improve recognition performance, not visual clarity. Two issues have been raised:

- The aim of applying super-resolution to biometrics is not for visual enhancement, but to improve recognition performance. Most existing super-resolution approaches are designed to produce a visual enhancement. *If recognition improvement is desired, why do we not focus on super-resolving only items essential for recognition?*
- Each biometric modality has its own characteristics. Most existing super-resolution approaches for biometrics are general-scene super-resolution approaches. *Can any specific information from biometric models be exploited to improve super-resolution performance?*

Based on this concern, Gunturk et al. [2] and Kien et al. [8] have proposed feature-domain super-resolution approaches for face and iris recognition respectively. These approaches no longer super-resolve images in the pixel-domain, but directly super-resolve in the feature-domain, and the super-resolution output (a super-resolved feature vector) is directly employed for recognition. In [2] and [8], PCA features are used to represent faces and irises respectively. These features are super-resolved to double resolution using a maximum a posteriori estimation approach. Specific knowledge of face and iris models is incorporated in the form of prior probabilities to constrain the super-resolution process, to make it robust to noise and segmentation errors. These approaches have been shown to outperform other pixel-domain super-resolution approaches for face and iris recognition.

To further improve the performance of feature-based super-resolution for iris recognition, we seek to investigate different features, and different approaches for obtaining a high resolution estimation. 2D Gabor phase encoding features, Linear Discriminant Analysis (LDA, also known as Fisheriris) features, and Principal Component Analysis (PCA, also known as Eigeniris) features are analysed and compared for a feature-domain super-resolution approach. Conjugate gradients and steepest descent optimisation methods are also investigated to improve the convergence of the high resolution estimation.

The remainder of this paper is organised as follows: iris feature encoding techniques are briefly summarised in Section 2; Section 3 describes the proposed feature-domain super-resolution approach for iris images; Section 4 explains our experiments and results; and the paper is concluded in Section 5.

II. Iris features for super-resolution

Principal Component Analysis (PCA) features have been previously employed [8] for feature-based super-resolution for iris recognition. The benefit of using PCA features is that PCA is a linear encoding technique, which simplifies the estimation process of the high resolution features. The standard and widely employed encoding approach for iris is the 2D Gabor phase-quadrant encoding technique, which has been proved to effectively encode the unique and discriminant information in the iris image [9]. However, this encoding technique is non-linear, making it very challenging to formulate the relationship between the low-resolution features and the high-resolution features in the feature domain. Hence in this paper, we only investigate linear encoding techniques for our proposed feature-domain super-resolution.

The aim of PCA encoding is to project data from a higher dimensionality to a lower dimensional manifold such that the error incurred by reconstructing the data in the higher dimensionality is minimised. However, the dimension reduction to minimise reconstruction error may not be the optimal transform to best classify the data classes. When those transforms do not match, the classification performance of the PCA encoding approach is degraded. Linear Discriminant Analysis (LDA) is also a linear encoding technique. However, it seeks the projection that maximises the distance between the data classes. This transform is achieved by searching for the largest ratio of between-class variation and within-class variation when projecting the original data to a subspace. PCA and LDA encoding techniques, when applied to iris recognition, are named Eigeniris and Fisheriris respectively.

Here we briefly introduce the LDA encoding approach for iris recognition. Suppose we have K classes: $X_1, X_2, ..., X_K$. Let the ith observation vector from X_j be x_{ij}, where $j = 1, .., K; i = 1, .., N_j$ and N_j is the number of observations from class j. Then, the sample mean vector μ_j and the covariance matrix S_j of class j are given by,

$$\mu_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_{ij}, \tag{1}$$

$$S_j = \frac{1}{N_j} \sum_{i=1}^{N_j} (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T. \tag{2}$$

The within-class covariance matrix, S_W, and the between-class covariance matrix, S_B, are given by,

$$S_W = \sum_{j=1}^{K} S_j, \tag{3}$$

$$S_B = \sum_{j=1}^{K} N_j (\mu_j - \mu)(\mu_j - \mu)^T, \tag{4}$$

where $\mu = \frac{1}{N} \sum_{j=1}^{K} N_j \mu_j$ is the mean of all samples and $N = \sum_{j=1}^{K} N_j$.

Let T be the projection from a p-dimensional space to a m-dimensional space, accomplished by m discriminant functions as follows,

$$y = T^T x, \tag{5}$$

where x is a vector of size $p \times 1$, y is a projected vector of size $m \times 1$, and T is a transformation matrix of size $p \times m$. LDA seeks the projection (matrix T) to maximise the ratio of the projected between-class covariance matrix, $\tilde{S}_B$, and the projected within-class covariance matrix, $\tilde{S}_W$, as follows,

$$J(T) = \frac{\tilde{S}_B}{\tilde{S}_W} = \frac{T^T S_B T}{T^T S_W T}. \tag{6}$$

It can be shown that the solution to the above equation for T is in fact the matrix of the leading m eigenvectors of the Fisheriris space matrix, $S_W^{-1} S_B$ [10].

During the training phase, Fisheriris space is estimated from gallery images. In practice, the smallest eigenvalues of $S_W^{-1} S_B$ are disregarded, and a Fisheriris space matrix, ϕ, which contains the Fisheriris as its columns is formed. Thus the new transformed space has smaller dimensionality than the original space. Depending on the compromise between representation accuracy and compactness, the number of Fisheririses employed can be varied.

In the testing/recognition phase, a normalised iris image, x, can be projected onto the Fisheriris space as follows,

$$x = \phi a + e_x, \tag{7}$$

where ϕ is the Fisheriris space matrix, e_x is the representation error, and a is the eigenvalue which serves as a feature vector for representing the normalised iris image. The identity of the probe image will be estimated as the identity of the class whose feature vector has the smallest distance to the probe feature vector. The distance between two feature vectors can be determined using either Euclidean or Hamming distance. In this paper, we use the Euclidean distance.

III. Feature-domain super-resolution for iris recognition

Kien et al. [8] proposed a five-stage approach to estimate high-resolution features from low-resolution iris images. In [8], this framework was used to create super-resolved PCA features using steepest descent optimisation. We adapt this framework to use LDA, and conjugate gradient estimation. LDA, with its ability to maximise the classification capability among the data classes, is expected to improve recognition performance over PCA, and the conjugate gradients estimation method is expected to provide a more accurate convergence to the unconstrained optimisation problem. The improved approach is illustrated in Figure 1 and described in more detail in the remainder of this Section.

Stage 1: Observation model in the spatial domain

Let x be the original HR iris image, and $y^{(i)}$ be the i^{th} observed LR iris image after being degraded by downsampling, $D^{(i)}$; blurring, $B^{(i)}$; and warping, $W^{(i)}$. The relation between $x, y^{(i)}$ is described as follows,

$$y^{(i)} = H^{(i)} x + n^{(i)} = D^{(i)} B^{(i)} W^{(i)} x + n^{(i)}, \tag{8}$$

where $n^{(i)}$ is the observation noise.

Stage 2: Observation model in the feature domain

We seek to transform the observation model from the spatial domain to the feature domain. In eigeniris recognition, HR

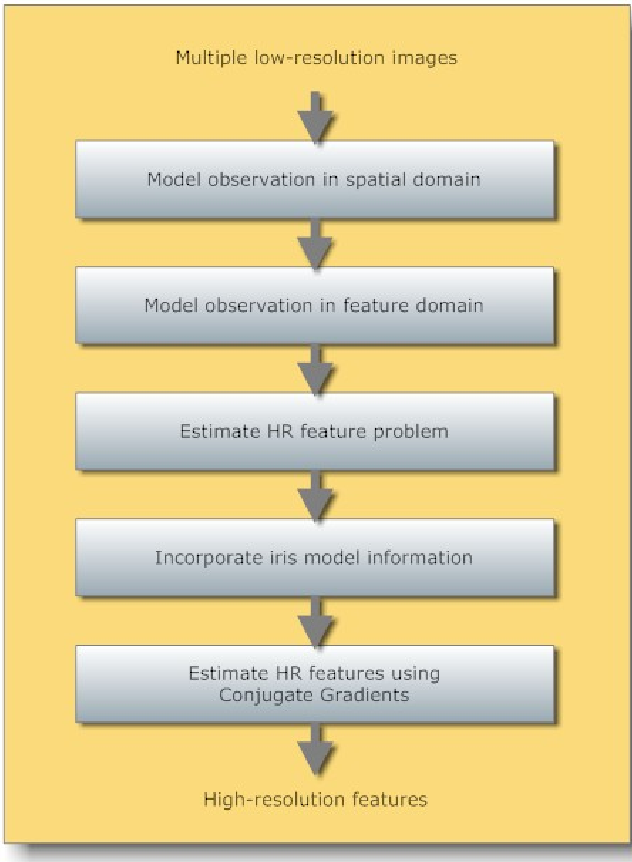

Fig. 1. The proposed feature-domain super-resolution approach for iris recognition.

irises, x, and observed LR irises, $y^{(i)}$, are represented as a superposition of eigenirises as follows,

$$x = \Phi a + e_x, \tag{9}$$

$$y^{(i)} = \Psi \widehat{a}^{(i)} + e_y, \tag{10}$$

where Φ, Ψ are Fisheriris matrixes, $a, \widehat{a}^{(i)}$ are feature vectors of HR and LR iris images, and e_x, e_y are representation errors.

Substituting the HR and LR feature representation of Equations (9) and (10) into the spatial observation model of Equation (8), we have,

$$\widehat{a}^{(i)} = \Psi^T H^{(i)} \Phi a + \Psi^T H^{(i)} e_x + \Psi^T n^{(i)}. \tag{11}$$

Equation (11) shows the relationship between the HR and observed LR iris features. The following sections will discuss a solution to estimate the HR iris features from this equation.

Stage 3: Estimating HR features

In Bayes statistics, a maximum a posteriori probability estimate can be used to estimate an unobserved quantity on the basis of empirical data. Using Bayes maximum a posteriori probability estimation, a HR feature can be estimated as,

$$\widetilde{a} = argmax_a p(\widehat{a}^{(1)}, \ldots, \widehat{a}^{(M)}|a)p(a). \tag{12}$$

The estimated HR feature, $\widetilde{a}$, is the value that maximises the product of the conditional probability $p(\widehat{a}^{(1)}, \ldots, \widehat{a}^{(M)}|a)$ and the priori probability $p(a)$.

Stage 4: Incorporating iris model information

To solve the above estimation problem, specific information relating to iris models can be incorporated in the form of two assumptions:

1. Prior probability is jointly Gaussian,

$$p(a) = \frac{1}{Z} exp(-(a-\mu_a)^T \Lambda^{-1}(a-\mu_a)). \tag{13}$$

2. Total observation noise and representation error is an Independent Identically Distributed (IID) Gaussian with a diagonal covariance matrix,

$$v^{(i)} = H^{(i)} e_x + n^{(i)},$$
$$\widehat{a}^{(i)} = \Psi^T H^{(i)} \Phi a + \Psi^T v^{(i)}, \tag{14}$$
$$p(v^{(i)}) = \frac{1}{Z} exp(-(v^{(i)} - \mu_v^{(i)})^T K^{-1}(v^{(i)} - \mu_v^{(i)})).$$

With these two assumptions, the conditional probability $p(\widehat{a}^{(1)}, \ldots, \widehat{a}^{(M)}|a)$ can be re-calculated as follows. According to multivariate random variable theory, $\Psi^T v^{(i)}$ is also a Gaussian distribution,

$$p(\Psi^T v^{(i)}) =$$
$$\frac{1}{Z} exp(-(\Psi^T v^{(i)} - \Psi^T \mu_v^{(i)})^T Q^{-1}(\Psi^T v^{(i)} - \Psi^T \mu_v^{(i)})),$$

where $Q = \Psi^T K \Psi$. Since $\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a = \Psi^T v^{(i)}$, so the individual conditional probability can be shown as,

$$p(a^{(i)}|a) = \frac{1}{Z} exp(-(\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a -$$
$$\Psi^T \mu_v^{(i)})^T Q^{-1}(\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \Psi^T \mu_v^{(i)})).$$

From (14), $\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a$ is IID as a consequence of the fact that $\Psi^T v^{(i)}$ is IID, thus,

$$p(\widehat{a}^{(1)}, \ldots, \widehat{a}^{(M)}|a) = \prod_i p(\widehat{a}^{(i)}|a) = \frac{1}{Z} exp(-\sum_{i=1}^{M}(\widehat{a}^{(i)} -$$
$$\Psi^T H^{(i)} \Phi a - \eta)^T Q^{-1}(\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \eta)).$$

The estimation problem can then be rewritten as,

$$\widetilde{a} = argmax_a p(\widehat{a}^{(1)}, \ldots, \widehat{a}^{(M)}|a)p(a)$$
$$= argmax_a \frac{1}{Z} exp(-\sum_{i=1}^{M}(\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \eta)^T Q^{-1}(\widehat{a}^{(i)} -$$
$$\Psi^T H^{(i)} \Phi a - \eta)) \times \frac{1}{Z} exp(-(a-\mu_a)^T \Lambda^{-1}(a-\mu_a))$$
$$= argmin_a (\sum_{i=1}^{M}(\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \eta)^T Q^{-1}(\widehat{a}^{(i)} -$$
$$\Psi^T H^{(i)} \Phi a - \eta) + (a-\mu_a)^T \Lambda^{-1}(a-\mu_a)).$$

Stage 5: Estimating the solution

In [8], the unconstrained optimisation problem in Stage 4 was solved by iterative steepest descent. With a proper choice of the step size and the maximum number of steps, the iterative steepest descent method is capable of converging to the local minimum sharply. However, iterative steepest descent may never reach the true minimum [11]. Instead of employing

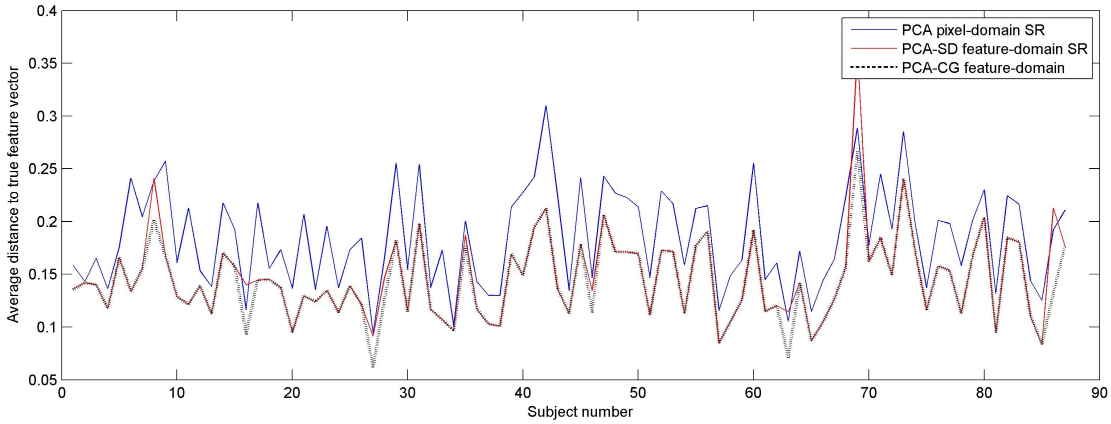

Fig. 2. Average distance to true feature vector of each subject using PCA features in pixel-domain SR, feature-domain SR with iterative steepest descent estimation and feature-domain SR with iterative conjugate gradients estimation.

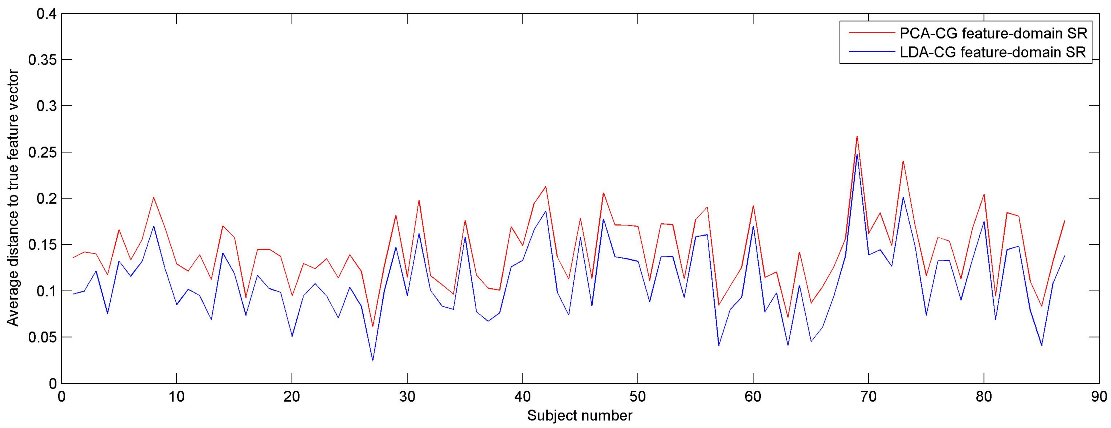

Fig. 3. Average distance to true feature vector of each subject using PCA and LDA features for feature-domain SR.

steepest gradient directions for iterative updating, a conjugate gradients method utilises conjugate directions, which enables the method to converge more accurately in at most n steps, where n is the size of the matrix of the system [11]. Given this, we propose to solve the estimation problem in Stage 4 by iterative conjugate gradients. Let the cost function E(a) be defined as,

$$E(a) = \sum_{i=1}^{M} (\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \eta)^T Q^{-1} (\widehat{a}^{(i)} - \Psi^T H^{(i)} \Phi a - \eta) + (a - \mu_a)^T \Lambda^{-1} (a - \mu_a).$$

The solution for optimisation can be estimated iteratively as follows,

$$a_{n+1} = a_n + \alpha_n \Gamma a_n, \tag{15}$$

where Γa_n is defined as,

$$\Gamma a_n = \triangle a_n + \beta_n \Gamma a_{n-1},$$

where $\triangle a_n = -\nabla_a E(a_n)$ and $\beta_n = max(0, \beta_n^{PR})$, and

$$\beta_n^{PR} = \frac{\triangle a_n^T (\triangle a_n - \triangle a_{n-1})}{\triangle a_{n-1}^T \triangle a_{n-1}}.$$

α_n is the parameter to minimise $E(a_n + \alpha_n \Gamma a_n)$ through a line search. Hence, with an initial estimation a_0, the iterative conjugate gradients estimation method will converge to the true high-resolution $\widetilde{a}$ which minimises the cost function $E(a)$.

IV. Experiments

In this paper, we use a subset of high quality still iris images from 87 subjects in the Multiple Biometric Grand Challenge dataset [12]. Two high quality iris images are selected from the dataset for each subject for experiments. One iris image serves as the gallery image, while the other is degraded by Gaussian blurring, random warping and downsampling by a factor of four to create a series of 16 low resolution images.

The proposed feature-domain super-resolution approach extracts PCA and LDA features from all LR iris images in the LR sequence, and then super-resolves these features directly to create a HR feature for recognition. Both steepest descent and conjugate gradients estimation methods are implemented in Stage 5 of the proposed approach for comparison. Different features and different estimation methods are compared based on the average distance of the reconstructed features to the true features (Section IV-A) and the recognition performance (Section IV-B).

To demonstrate the performance of the proposed feature-domain super-resolution approach against pixel-domain approaches, the pixel-domain super-resolution approach described in [6] is reproduced for comparison.

A. Average Distance to true feature vectors

The distance between a reconstructed feature vector, $\widetilde{a}$, and the true HR feature vector, a, is calculated for each subject as follows,

$$D(a, \widetilde{a}) = \frac{\|a - \widetilde{a}\|}{\|a\|} \times \frac{1}{Length(a)},$$

where $\|.\|$ is the Euclidean distance; and $Length(a)$, the length of vector a, is used for normalisation.

The distance between a reconstructed HR feature vector using the proposed feature-domain super-resolution approach and the true HR feature vector is compared with the distance between a reconstructed HR feature vector using the pixel-domain super-resolution approach and the true HR feature vector, and the distance between the LR feature and the true HR feature. As the distances between LR features and the true HR features are much higher than the distances between reconstructed HR features and the true HR features (as shown in Table I), we do not plot those distances in the figures to enable better visual comparison of the other distances. We show two comparisons: in Figure 2 we illustrate the performance improvement obtained by using conjugate gradient estimation when compared to iterative steepest descent using PCA features; and in Figure 3 we compare LDA and PCA features using conjugate gradient estimation.

TABLE I
AVERAGE DISTANCE TO TRUE FEATURE VECTORS OF ALL SUBJECTS.

Approaches	Average distance
LR	1.576
PCA-SD Pixel-based SR	0.185
PCA-SD Feature-based SR	0.165
PCA-CG Pixel-based SR	0.163
LDA-CG Feature-based SR	0.136

From the average distance to true HR vectors in Figure 2 and Figure 3, we witness that HR features estimated using LDA features and the conjugate gradients estimation are closer to the true values. Looking closely at Figure 2, we see that there are a number of subjects (e.g. 8,16,63,69,86) where HR features estimated by feature-based super-resolution using iterative steepest descent are further from the true values than those estimated by pixel-based super-resolution. In these cases, our experiments show that the steepest descent estimation approach fails to converge to a stable value. Conjugate gradients estimation is capable of converging in a more stable manner, which has been shown in Figure 3. This is also demonstrated in Table I.

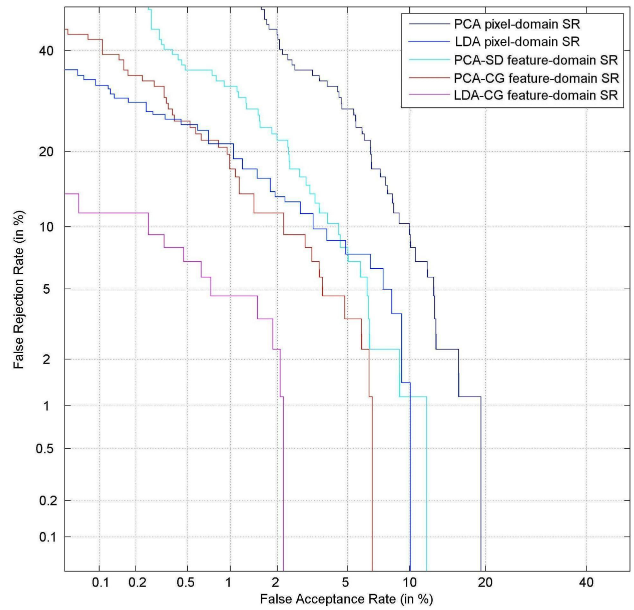

Fig. 4. Recognition performance.

B. Recognition performance

Figure 4 illustrates recognition performance of feature-domain super-resolution against pixel-domain super-resolution. For both cases of PCA and LDA features, feature-domain super-resolution achieves superior performance by directly super-resolving only information essential for recognition, and incorporating iris model constraints. Figure 4 also demonstrates that using LDA features and conjugate gradients estimation approach improves recognition performance over PCA features and steepest descent estimation.

V. Conclusion

Feature-domain super-resolution has been shown to improve the performance of biometrics such as face and iris recognition. This technique focuses on recognition performance improvement, not visual enhancement, by super-resolving directly in the feature domain. The incorporation of the biometric model specific information into the high-resolution feature estimation strengthens the accuracy of the estimation. In this paper we have investigated a method to improve feature-domain super-resolution for iris recognition by using Fisheriris (LDA) as the feature. The LDA encoding technique boosts recognition performance of a feature-based super-resolution approach since it seeks a projection that maximises the separation between classes. In addition, conjugate gradients optimisation has also been applied instead of steepest descent, and has been shown to enhance the accuracy of convergence in the estimation problem.

Future work will perform further evaluations on the proposed approach using more challenging data such as the

MBGC portal dataset. We are also focusing on investigating other non-linear features for the feature-domain super-resolution approach.

ACKNOWLEDGMENT

Our thanks to University of Notre Dame du Lac, National Institute of Standards and Technology (NIST) for allowing us to use the MBGC dataset in this research.

REFERENCES

[1] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21 – 36, 2003.

[2] B. Gunturk, A. Batur, Y. Altunbasak, M. Hayes, and R. Mersereau, "Eigenface-domain super-resolution for face recognition," *Image Processing, IEEE Transactions on*, vol. 12, no. 5, pp. 597 – 606, 2003.

[3] K. Jia and S. Gong, "Generalized face super-resolution," *Image Processing, IEEE Transactions on*, vol. 17, no. 6, pp. 873 –886, 2008.

[4] F. Lin, C. Fookes, V. Chandran, and S. Sridharan, "Super-resolved faces for improved face recognition from surveillance video," in *Lecture Notes in Computer Science*, vol. 4642 LNCS, Seoul, Korea, Republic of, 2007, pp. 1 – 10.

[5] G. Fahmy, "Super-resolution construction of iris images from a visual low resolution face video," in *Signal Processing and Its Applications, ISSPA 2007. 9th International Symposium on*, 12-15 2007, pp. 1 –4.

[6] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Focus-score weighted super-resolution for uncooperative iris recognition at a distance and on the move," *25th International Conference of Image and Vision Computing New Zealand, IVCNZ*, 2010.

[7] K. Y.S., K. R.P., B. J.K., and S. J.P., "Super-resolution method based on multiple multi-layer perceptrons for iris recognition," in *Ubiquitous Information Technologies Applications, ICUT '09. International Conference on*, 20-22 2009, pp. 1 –5.

[8] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Feature-domain super-resolution for iris recognition," *18th IEEE International Conference on Image Processing*, 2011.

[9] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 21 – 30, 2004.

[10] D. G. S. Richard O. Duda (Author), Peter E. Hart (Author), *Pattern Classification*. John Wiley and Sons Inc, 2001.

[11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes : the art of scientific computing*. Cambridge University Press, 2007.

[12] P. Phillips, P. Flynn, J. Bevcridge, W. Scruggs, A. O'Toole, D. Bolme, K. Bowyer, B. Draper, G. Givens, Y. M. Lui, H. Sahibzada, I. Scallan, J.A., and S. Weimer, "Overview of the multiple biometrics grand challenge," in *Advances in Biometrics. Proceedings Third International Conference, ICB 2009*, Berlin, Germany, 2009, pp. 705 – 14.