Gait feature coupling for low-resolution face recognition

X.Y. Ben, M.Y. Jiang, Y.J. Wu and W.X. Meng

A novel low-resolution (LR) face recognition method has been developed by coupling gait features. The proposed kernel-based manifold method is able to couple the nonlinear features of an LR face with gait and map them a common space to minimise the distance between the two features extracted from the same individual. Experimental results demonstrate the effectiveness of the proposed method.

Introduction: Low resolution (LR) face recognition is one of the most challenging issues in biometric research because LR facial images provide limited information causing low accuracy in face recognition. A potential method to improve the accuracy is to reconstruct a high-resolution (HR) image through existing LR images. Owing to the high similarity between the HR image and its multiple LR counterparts, a manifold alignment-based hallucination method for super-resolution (SR) realisation can enhance the resolution of the facial image [1]. Huang and He [2] established the coherent subspaces by using canonical correlation analysis (CCA), and then built a radial basis function (RBF) mapping between the HR and the LR feature spaces. Then the SR image corresponding to the input LR one can be computed to fulfil face recognition using the trained RBF model. However, most existing SR algorithms might not be effective in LR face recognition because the face similarity metrics in the LR feature space cannot reflect the actual similarity in the HR feature space [3]. Li *et al.* [4] proposed an efficient method called coupled locality preserving mappings (CLPM) for face recognition without any SR preprocessing. To the best of our knowledge few research efforts have been made to couple gait features with LR facial images for LR face recognition. Inspired by Li's work [4], we propose a novel kernel-based manifold method to realise the gait feature coupling for LR face recognition. The nonlinear mapping process is to minimise the distance between the sample feature in the LR face dataset and the one in the gait dataset for an identical individual. We conducted experiments to demonstrate the convincing performance of the proposed method.

Gait feature coupling for LR face recognition: Coupling gait features for LR face recognition is based on the assumption that the human's gait is visible whereas the input face images suffer from degradation. As for the feature extraction of gait silhouette sequences, the robust dual-ellipse fitting approach [5] is used to detect the gait period. Each gait image is normalised to a fixed size 64 × 64 pixel during a complete walking cycle. A simple weighted average method is used to construct gait energy images (GEIs), as shown in Fig. 1*b*.

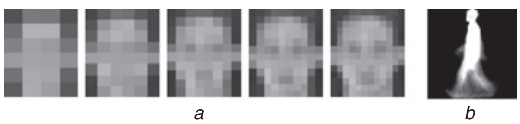

Fig. 1 *Examples of low-resolution face images (6 × 4, 8 × 6, 10 × 8, 12 × 10 and 14 × 12 pixels) and normal-resolution GEI (64 × 64 pixels)*

Fig. 2 shows an overview of the proposed method, which consists mainly of training and testing phases. In the training phase, first, the GEI set (the registration set) and LR face image set are projected into the kernel space. Here, let $x_1, \cdots, x_M$ and $y_1, \cdots, y_M$ be the LR face image set and their corresponding GEI set respectively, where M is the total number of samples. Denote ϕ for $\phi(x_i)$ or $\phi(y_j)$ for $i, j = 1, 2, \cdots, M$ a nonlinear function. To simplify the derivation process, we assume that $\phi(x_i)$ and $\phi(y_j)$ have their centre at the origin, i.e. $\sum_i \phi(x_i) = 0$ and $\sum_j \phi(y_j) = 0$. To find matrices P_1 and P_2 that minimise the distance between each sample in both LR face image and the GEI sets for each identical individual, the objective function is formulated as

$$f(P_1, P_2) = \min \sum\nolimits_{i,j} \|P_1^{\mathrm{T}}\phi(x_i) - P_2^{\mathrm{T}}\phi(y_j)\|^2 S_{ij} \qquad (1)$$

where S_{ij} is the element of the similarity matrix S.

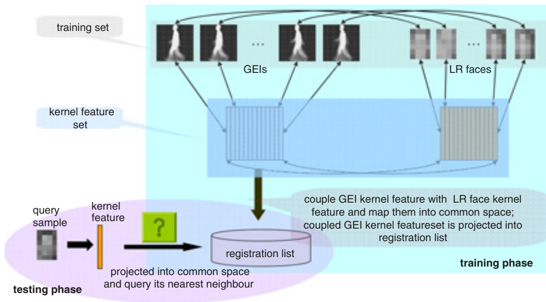

Fig. 2 *Overview of proposed low-resolution face recognition method*

To solve (1), we define $\phi(X) = [\phi(x_1), \phi(x_2), \cdots, \phi(x_M)]$, $\phi(Y) = [\phi(y_1), \phi(y_2), \cdots, \phi(y_M)]$, $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_M]^{\mathrm{T}}$ and $\beta = [\beta_1, \beta_2, \cdots, \beta_M]^{\mathrm{T}}$. Since there are M samples in total, P_1 and P_2 can be, respectively, spanned as $P_1 = \sum_i \alpha_i \phi(x_i)$ and $P_2 = \sum_j \beta_j \phi(y_j)$ for $i, j = 1, 2, \cdots, M$. With some algebraic deduction, (1) can be simplified as

$$\begin{aligned} f(\alpha, \beta) = \mathrm{tr}\Bigg(& \begin{bmatrix} \alpha \\ \beta \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \phi(X)^{\mathrm{T}} & \\ & \phi(Y)^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \phi(X) & \\ & \phi(Y) \end{bmatrix} \\ & \times \begin{bmatrix} D_h & -S \\ -S^{\mathrm{T}} & D_v \end{bmatrix} \begin{bmatrix} \phi(X) & \\ & \phi(Y) \end{bmatrix}^{\mathrm{T}} \\ & \times \begin{bmatrix} \phi(X) & \\ & \phi(Y) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \Bigg) \\ = \mathrm{tr}\Bigg(& \begin{bmatrix} \alpha \\ \beta \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} K_x & \\ & K_y \end{bmatrix} \begin{bmatrix} D_h & -S \\ -S^{\mathrm{T}} & D_v \end{bmatrix} \\ & \times \begin{bmatrix} K_x & \\ & K_y \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \Bigg) \end{aligned} \qquad (2)$$

where $K_x = \phi(X)^{\mathrm{T}}\phi(X)$, and $K_y = \phi(Y)^{\mathrm{T}}\phi(\mathrm{Y})$ can be Gaussian kernel functions. Both D_h and D_v are diagonal matrices with $D_{hii} = \sum_j S_{ij}$ and $D_{vjj} = \sum_i S_{ij}$.

To derive the projection matrices α and β, (2) is reformulated as

$$f(W) = \mathrm{tr}(W^{\mathrm{T}} Z \Theta Z^{\mathrm{T}} W) \qquad (3)$$

where $W = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, $Z = \begin{bmatrix} K_x & \\ & K_y \end{bmatrix}$ and $\Theta = \begin{bmatrix} D_h & -S \\ -S^{\mathrm{T}} & D_v \end{bmatrix}$. Then, the projection W can be solved by the following generalised eigen-decomposition:

$$(Z\Theta Z^{\mathrm{T}})W = \lambda(ZZ^{\mathrm{T}})W \qquad (4)$$

Let $w_1, \cdots, w_{d'}$ be the eigenvectors corresponding to the d' smallest eigenvalues ordered according to $\lambda_1 \leq \cdots \leq \lambda_{d'}$. $W = [w_1, \cdots, w_{d'}]$ can be divided into the two matrices $\alpha \in \mathrm{R}^{M \times d'}$ and $\beta \in \mathrm{R}^{M \times d'}$. Note that the matrix ZZ^{T} may be singular, and can be adjusted to $ZZ^{\mathrm{T}} + \tau I$, where τ is a small positive value (say $\tau = 10^{-6}$), and I represents an identity matrix.

The features of the LR face set $[f_{x1}, f_{x2}, \cdots, f_{xM}] = P_1^{\mathrm{T}}\phi(X) = \alpha^{\mathrm{T}} K_x$ and the GEI set $[f_{y1}, f_{y2}, \cdots, f_{yM}] = P_2^{\mathrm{T}}\phi(Y) = \beta^{\mathrm{T}} K_y$ are all coupled and mapped into a common space.

In the testing phase, the kernel feature of a LR face probe sample x' is expressed as $\phi(x')$. Then it is projected into the common space, i.e. $f_x' = P_1^{\mathrm{T}}\phi(x') = \alpha^{\mathrm{T}}[\phi(x_1)\phi(x'), \cdots, \phi(x_M)\phi(x')]^{\mathrm{T}}$. Finally, we apply the nearest neighbour classifier to query the identity of x' in the gait registration list $[f_{y1}, f_{y2}, \cdots, f_{yM}]$.

Experimental results: To verify the proposed approach, we constructed an experimental database containing data from the ORL face database [6] and the CASIA(B) gait database [7]. This experimental database includes 40 subjects, randomly selected from these two databases. Each subject consists of three gait sequences and seven face images (down sampled to LR, such as 6 × 4, 8 × 6, 10 × 8, 12 × 10, 14 × 12 pixels, as shown in Fig. 1*a*). In each LR face recognition experiment, we randomly selected three LR face images coupled with three GEIs for each identical individual as the training set and the rest of the face images were used for testing. We conducted each experiment 30 times and calculated the average correct classification rate (CCR). We adopted two options for constructing S_{ij}. The first option considers the supervised class information in the training phase. If x_i is the intra-class of x_j, then $S_{ij} = \exp(-\|x_i - x_j\|^2/t)$ (t is a scale factor); otherwise $S_{ij} = 0$. The second is cosine similarity. If x_i is the k-nearest neighbour of x_j, or x_j is the k-nearest neighbour of x_i, then $S_{ij} = (x_i \cdot x_j)/(\|x_i\|\|x_j\|)$, otherwise $S_{ij} = 0$.

We compared the CCR and the cumulative match score (CMS) of the proposed method with state-of-the-art methods such as CLPM [4], Huang and He's [2] and the principal component analysis (PCA) combined with RBF (PCA-RBF) [2]. The comparisons of recognition performance with different resolutions are shown in Fig. 3. According to the experimental results, the proposed kernelised method based on both S_{ij} construction options is fairly robust to the degraded resolution of facial images. The proposed kernelised method can improve the performance of their original linear algorithm (i.e. CLPM), and also outperform Huang's and PCA-RBF methods.

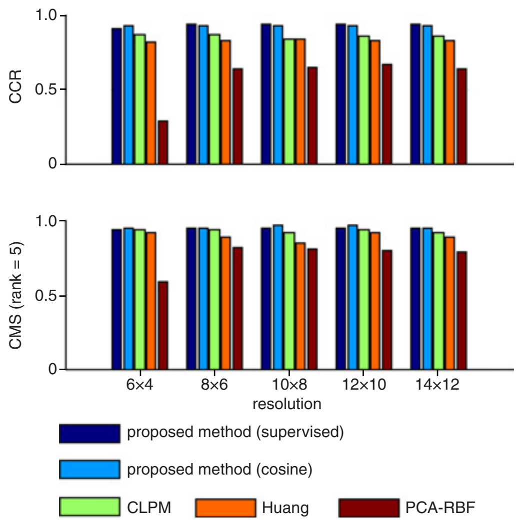

Fig. 3 *Recognition performance comparisons of CCR and CMS among various methods*

Conclusion: Presented is an innovative LR face recognition method which couples GEIs with LR face images and maps them into the common space to improve the accuracy of LR face recognition. The experimental results show the effectiveness and advantage of the proposed method.

Acknowledgements: The authors thank the Institute of Automation Chinese Academy of Sciences for granting us permission to use the CASIA(B) gait database, and also thank H. Huang for providing the code of his method. This project has been supported the National Science Foundation for Post-doctoral Scientists of China (Grant 20110491087)

25 December 2011
doi: 10.1049/el.2011.4041
One or more of the Figures in this Letter are available in colour online.

X.Y. Ben and M.Y. Jiang (*School of Information Science and Engineering, Shandong University, No.27, Shanda South Road, Jinan City 250100, Shandong Province, People's Republic of China*)

E-mail: benxianye@gmail.com

Y.J. Wu (*Parks College of Engineering, Aviation and Technology, Saint Louis University, St. Louis, MO 63103, USA*)

W.X. Meng (*School of Electronics Information Engineering, Harbin Institute of Technology, No.92, Western Dazhi Street, Nangang District, Harbin, Heilongjiang Province, People's Republic of China*)

References

1 Li, B., Chang, H., Shan, S.G., and Chen, X.L.: 'Aligning coupled manifolds for face hallucination', *IEEE Signal Process. Lett.*, 2009, **16**, (11), pp. 957–960
2 Huang, H., and He, H.T.: 'Super-resolution method for face recognition using nonlinear mappings on coherent features', *IEEE Trans. Neural Netw.*, 2011, **22**, (1), pp. 121–130
3 Zou, W., and Yuen, P.: 'Very low resolution face recognition problem', *IEEE Trans. Image Process.*, doi: 10.1109/TIP.2011.2162423
4 Li, B., Chang, H., Shan, S.G., and Chen, X.L.: 'Low-resolution face recognition via coupled locality preserving mappings', *IEEE Signal Process. Lett.*, 2010, **17**, (1), pp. 20–23
5 Ben, X.Y., Meng, W.X., and Yan, R.: 'Dual-ellipse fitting approach for robust gait periodicity detection'. *Neurocomput.*, doi:10.1016/j.neucom.2011.10.009
6 Gross, R.: 'Face databases, handbook of face recognition', *in* Li, S.Z., and Jain, A.K. (Ed.) (Springer-Verlag, 2005)
7 Yu, S., Tan, D., and Tan, T.N.: 'A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition'. Proc. 18th Int. Conf. on Pattern Recognition, Hong Kong, China, 2006, pp. 441–444