Coupled Marginal Fisher Analysis for Low-Resolution Face Recognition

Stephen Siena, Vishnu Naresh Boddeti, and B.V.K. Vijaya Kumar

Carnegie Mellon University, Electrical and Computer Engineering,
5000 Forbes Avenue, Pittsburgh, Pennsylvania, USA 15213
ssiena@andrew.cmu.edu, naresh@cmu.edu, kumar@ece.cmu.edu

Abstract. Many scenarios require that face recognition be performed at conditions that are not optimal. Traditional face recognition algorithms are not best suited for matching images captured at a low-resolution to a set of high-resolution gallery images. To perform matching between images of different resolutions, this work proposes a method of learning two sets of projections, one for high-resolution images and one for low-resolution images, based on local relationships in the data. Subsequent matching is done in a common subspace. Experiments show that our algorithm yields higher recognition rates than other similar methods.

1 Introduction

Face recognition is a prevalent technology, with a wide range of applications in the public and private sectors. This is facilitated by the dramatic improvements to recognition systems in the past twenty years. Face recognition can be reliable when images are captured under controlled conditions.

Difficulties may arise in unconstrained environments where quality images are not easily collected. Oftentimes, these environments involve subjects that are uncooperative or unaware that an image is being captured. In some scenarios, it can be desirable to photograph subjects without their active participation or knowledge, such as in surveillance videos or when the subject is at a longer distance from the camera.

In these cases where the image acquisition environment is not ideal, captured faces can have a much lower resolution than faces captured in a controlled setting. Faces captured at a low resolution provide a challenge for recognition algorithms that rely on traditional feature extraction; such features can be impossible to compute on small images where facial features are represented by only a few pixels. The previously mentioned problems make it challenging to compare low-resolution (LR) images captured at testing to high-resolution (HR) images, as the LR and HR images do not share a common feature representation.

This problem has recently received increasing attention in the biometrics community. While naïve methods for matching LR images to HR images have been available for some time, algorithms tailored to LR face recognition have begun to emerge. This work introduces Coupled Marginal Fisher Analysis (CMFA), a

A. Fusiello et al. (Eds.): ECCV 2012 Ws/Demos, Part II, LNCS 7584, pp. 240–249, 2012.
© Springer-Verlag Berlin Heidelberg 2012

new algorithm that learns projections that map HR images and LR images to a common subspace. We find projections by extending the principles of Marginal Fisher Analysis (MFA) [1], an effective linear projection method for images of the same resolution. The formulation is closest to Simultaneous Discriminant Analysis (SDA) [2], but does not make the same assumptions about how the data is distributed in the subspace. While SDA is optimal for Gaussian distributions, CMFA learns projections based on the local neighborhoods of data samples, and makes no assumptions about the global data distribution in the subspace.

Most existing methods for LR face recognition operate in the scenario where the training and testing classes are distinct. SDA and other techniques have been demonstrated to operate well in such scenarios [3], [4]. These techniques are good at modeling the relation between HR and LR images that can be extended to unseen subjects. In this process, these methods try to model the blur function between HR and LR images, but do not account for class specific details. However, in certain supervised settings, such as identifying people on a watch list, we can afford to train on the class of images used in testing. Tuning the learned projections to model class specific features along with the image formation operation is expected to result in better classification performance. Further, in this scenario, learning these projections taking advantage of the local neighborhoods of the data samples is beneficial. This is the motivation for CMFA.

2 Related Work

This work addresses the matter of matching face images of different resolutions. When pixels are used as the feature representation of the face, images of different resolutions have feature vectors of different lengths. In this section, we discuss different approaches to transforming images so that all data samples have a feature vector of the same size.

A simple approach to matching LR probe images is to reduce the size of the gallery images. Once gallery images are downsampled, traditional linear projection methods for face recognition, including Principal Component Analysis (PCA) [5], [6], Linear Discriminant Analysis (LDA) [7], and MFA [1], can be used. While these approaches can be effective for comparing images of the same resolution, downsampling the gallery images needlessly discards information in the data.

The opposite approach is to increase the resolution of the probe images. Performing some form of super-resolution makes the probe images the same dimensionality as the gallery images, and standard single resolution methods can once again be used. Simple methods such as bilinear or bicubic interpolation do not require any training. Other super-resolution methods can be trained to learn face priors [8] or the relationship between high and low-resolution images [9]. These approaches can be effective for reconstructing HR images, and while they can produce visually appealing results, they often lack the high frequency components of true HR images to be very effective for recognition tasks.

Finally, methods that truly attempt to match images of different resolutions have emerged. Hennings-Yeoman et al. present a method of super-resolution that

simultaneously includes feature extraction to guide the process towards recognition instead of only reconstruction [10]. The methods closest to our proposed algorithm map images from both the high and low-resolutions to a common subspace. Li et al. learn a set of projections for each resolution using Coupled Mapping, and also use a penalty matrix to preserve local relationships [3]. Zhou et al. proposed Simultaneous Discriminant Analysis (SDA) to learn two sets of projections in a similar way while adding within-class scatter and between-class scatter to the objective function, much like LDA in the single resolution case [2]. Biswas et al. map images to a common subspace such that the distances between HR and LR images approximate the distances between two HR images [4].

3 Coupled Marginal Fisher Analysis

3.1 Problem Statement

Assume we have a gallery of HR images $\mathbf{H} = [h_1, h_2, \ldots, h_N]$, $h_i \in \Re^M$, where N is the number of gallery images and M is the feature dimension. Provided with each gallery image h_i is a class label π_i. Given a m-dimensional LR probe image x, where $m << M$, we want to assign it to class

$$\pi_i \text{ where } i = \arg\min_i d(f_h(h_i), f_l(x)), \tag{1}$$

where $d(\cdot,\cdot)$ is a distance metric and $f_h(\cdot)$ and $f_l(\cdot)$ are functions that map HR and LR images to a common subspace. If $f_h(\cdot)$ and $f_l(\cdot)$ are sets of linear projections, then Eq. 1 can be rewritten as

$$\pi_i \text{ where } i = \arg\min_i d(\mathbf{P}_H h_i, \mathbf{P}_L x), \tag{2}$$

where $\mathbf{P}_H$ and $\mathbf{P}_L$ are $\hat{m} \times M$ and $\hat{m} \times m$ matrices, respectively. The distance function $d(\cdot,\cdot)$ is therefore computed on $\hat{m}$-dimensional data, where $\hat{m} \leq m$. Euclidean distance is chosen for this work.

When learning $\mathbf{P}_H$ and $\mathbf{P}_L$, we recognize that we want $d(\mathbf{P}_H h_i, \mathbf{P}_L x)$ to be small when x belongs to the same class as h_i, and large otherwise. Therefore, our objective function is

$$J(\mathbf{P}_H, \mathbf{P}_L) = \min \frac{\sum_{\pi_i = \pi_j} \|\mathbf{P}_H h_i - \mathbf{P}_L l_j\|_2^2 w_{ij}}{\sum_{\pi_i \neq \pi_j} \|\mathbf{P}_H h_i - \mathbf{P}_L l_j\|_2^2 w_{ij}^P}, \tag{3}$$

where $\mathbf{L} = [l_1, l_2, \ldots, l_N]$ is the set of LR training images. In the absence of LR training images, HR training images can be downsampled and blurred instead. $\mathbf{W}$ and $\mathbf{W^P}$ represent intrinsic and penalty adjacency matrices, respectively, that weight image pairs to be emphasized when learning $\mathbf{P}_H$ and $\mathbf{P}_L$. For more discussion about these matrices, refer to [1]. The definitions for these matrices used in this work are provided in the next section.

If we assume that $w_{ij} = 0$ if $\pi_i \neq \pi_j$ and $w^P_{ij} = 0$ if $\pi_i = \pi_j$, then Eq. 3 can be rewritten as

$$J(\mathbf{P}_H, \mathbf{P}_L) = \min \frac{Tr(\mathbf{P}_H^T \mathbf{H} \mathbf{D}^H \mathbf{H}^T \mathbf{P}_H + \mathbf{P}_L^T \mathbf{L} \mathbf{D}^L \mathbf{L}^T \mathbf{P}_L - \mathbf{P}_H^T \mathbf{H} \mathbf{W}^T \mathbf{L}^T \mathbf{P}_L - \mathbf{P}_L^T \mathbf{L} \mathbf{W} \mathbf{H}^T \mathbf{P}_H)}{Tr(\mathbf{P}_H^T \mathbf{H} \mathbf{D}^{\mathbf{P}H} \mathbf{H}^T \mathbf{P}_H + \mathbf{P}_L^T \mathbf{L} \mathbf{D}^{\mathbf{P}L} \mathbf{L}^T \mathbf{P}_L - \mathbf{P}_H^T \mathbf{H} \mathbf{W}^{\mathbf{P}T} \mathbf{L}^T \mathbf{P}_L - \mathbf{P}_L^T \mathbf{L} \mathbf{W}^{\mathbf{P}} \mathbf{H}^T \mathbf{P}_H)} \tag{4}$$

where $\mathbf{D}^L$, $\mathbf{D}^H$, $\mathbf{D}^{\mathbf{P}L}$, and $\mathbf{D}^{\mathbf{P}H}$ are diagonal matrices defined as $d^L_{ii} = \sum_j w_{ij}$, $d^H_{jj} = \sum_i w_{ij}$, $d^{PL}_{ii} = \sum_j w^P_{ij}$, and $d^{PH}_{jj} = \sum_i w^P_{ij}$. $Tr(\cdot)$ represents the trace operator. Eq. 4 can be rewritten as

$$J(\mathbf{P}_H, \mathbf{P}_L) = \min \frac{Tr(\mathbf{P}^T \mathbf{X} \mathbf{A} \mathbf{X}^T \mathbf{P})}{Tr(\mathbf{P}^T \mathbf{X} \mathbf{B} \mathbf{X}^T \mathbf{P})}, \tag{5}$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_L \\ \mathbf{P}_H \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} \mathbf{D}^L & -\mathbf{W} \\ -\mathbf{W}^T & \mathbf{D}^H \end{bmatrix}, \text{and } \mathbf{B} = \begin{bmatrix} \mathbf{D}^{\mathbf{P}L} & -\mathbf{W}^{\mathbf{P}} \\ -\mathbf{W}^{\mathbf{P}T} & \mathbf{D}^{\mathbf{P}H} \end{bmatrix}. \tag{6}$$

We can compute $\tilde{\mathbf{A}} = \mathbf{X}\mathbf{A}\mathbf{X}^T$ and $\tilde{\mathbf{B}} = \mathbf{X}\mathbf{B}\mathbf{X}^T$, and substituting these into Eq. 5 yields

$$J(\mathbf{P}_H, \mathbf{P}_L) = \min \frac{Tr(\mathbf{P}^T \tilde{\mathbf{A}} \mathbf{P})}{Tr(\mathbf{P}^T \tilde{\mathbf{B}} \mathbf{P})}, \tag{7}$$

which can be turned into the generalized eigenvalue problem,

$$\tilde{\mathbf{A}}\mathbf{P} = \lambda \tilde{\mathbf{B}}\mathbf{P} \tag{8}$$

3.2 Computing W and $\mathbf{W}^{\mathbf{P}}$

MFA. MFA is a discriminative projection learning algorithm designed to be effective on data irrespective of the distribution of each class [1]. On the other hand, LDA is designed to be optimal when each individual class follows a Gaussian distribution, but will not be as effective on non-Gaussian data classes. As discussed in [1], MFA does not encode global relationships between classes into the intrinsic and penalty adjacency matrices. Instead, MFA encodes the intra-class compactness and interclass separability with local relationships. The adjacency matrices are encoded using nearest neighbors as follows:

$$w_{ij} = \begin{cases} 1, \text{ if } i \in N^+_{k_1}(j) \text{ or } j \in N^+_{k_1}(i) \\ 0, \mathit{else}. \end{cases} \tag{9}$$

$$w^P_{ij} = \begin{cases} 1, \text{ if } (i,j) \in P^+_{k_2}(\pi_i) \text{ or } (i,j) \in P^+_{k_2}(\pi_j) \\ 0, \mathit{else}. \end{cases} \tag{10}$$

$N_{k_1}^{+}(i)$ represents the set of the k_1 nearest neighbors of x_i from the same class. c_i represents the class label of x_i, and $P_{k_2}^{+}(\pi_i)$ represents the set of the k_2 nearest pairs among the set $\{(i,j), i \in \pi_i, j \notin \pi_i\}$. In essence, MFA seeks to project data into a subspace so that a particular data sample is projected close to its local neighborhood from the same class. At the same time, it attempts to maximize the distance between the pairs of data from different classes that were originally the closest together. Note that with this formulation, MFA has two parameters which must be learned, k_1 and k_2.

CMFA. CMFA uses similar rules for defining $\mathbf{W}$ and $\mathbf{W^P}$. Initially, the same rules were used, and nearest neighbors were computed using the Euclidean distance of HR images. Tests showed that a linear combination of the Euclidean distance of HR images and LR images had a negligible effect on the nearest neighbors when the LR images were generated from corresponding HR images.

One change that was tested and retained after stronger performance was the rule for defining $\mathbf{W^P}$. We tried forming $\mathbf{W^P}$ by finding the nearest neighbors between classes of each data sample instead of the closest pairs between classes, so the penalty adjacency graph is computed as follows,

$$w_{ij}^{P} = \begin{cases} 1, \text{ if } i \in N_{k_2}^{-}(j) \text{ or } j \in N_{k_2}^{-}(i) \\ 0, else. \end{cases} \tag{11}$$

where $N_{k_2}^{-}(i)$ represent the of the k_2 nearest neighbors of x_i from a different class.

3.3 Regularization

Often is it desirable to add in a regularization factor when learning projections. If we consider Eq. 5, we consider that regularization can be added as follows,

$$J(\mathbf{P}_H, \mathbf{P}_L) = \min \frac{Tr(\mathbf{P}^T\mathbf{X}\mathbf{A}\mathbf{X}^T\mathbf{P})}{Tr(\mathbf{P}^T\mathbf{X}(\mathbf{B}+\alpha\mathbf{I})\mathbf{X}^T\mathbf{P})}, \tag{12}$$

where $\mathbf{I}$ is the identity matrix, and α is the regularization constant. If regularization is included, the derivation of the eigenvalue problem is no different, since the sum of $\mathbf{B}+\alpha\mathbf{I}$ is known.

4 Experiments

4.1 General Testing Protocols

We compare CMFA to SDA and MFA. SDA is a coupled mapping and learns projections for HR and LR probe images at the same time. On the other hand, MFA must be learned at a single dimension. Results labeled 'MFA' is a baseline using HR gallery and probe images. 'MFA-LR' uses LR gallery and probe images. Results for 'MFA-BL' use HR gallery images and learns HR projections, and

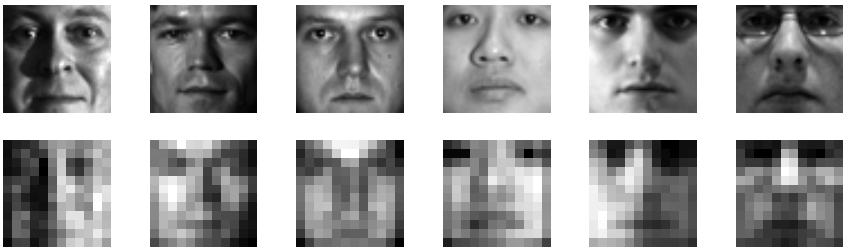

Fig. 1. HR and corresponding LR images from PIE (left 3 columns) and Multi-PIE (right 3 columns)

then uses LR probe images that are upsampled to the HR dimensionality using bi-linear interpolation.

Both MFA and CMFA require two parameters, k_1 and k_2. There is no clear method for setting these values, so a grid search is performed to find which values of k_1 and k_2 to use for each test. Similarly, different values for the regularization constant α are tested for all tested algorithms. We tested the algorithms with values of $\alpha = [10^{-6}, 10^{-5}, 10^{-4}]$, and found that in nearly all cases, $\alpha = 10^{-5}$ was optimal for the algorithms used.

In many cases, using as many projections as is possible to learn does not provide the best results. Because the 12×12 LR images used have 144 feature dimensions, we report the single best result from using $[10, 20, 30, ..., 140]$ projections.

Following projection learning, matching of probe images is done by finding the Euclidean distance nearest neighbor. Performance is measured by rank-1 identification rate, which equals the percentage of probe images that find a true match using the nearest neighbor classification.

All reported results are the average of 20 trials.

4.2 PIE / Multi-PIE

For testing we use the CMU Pose, Illumination, and Expression (PIE) database [11], and the CMU Multi-PIE database [12]. We take a part of the PIE dataset containing frontal images of 66 subjects with varying illumination, such that each subject has 21 images taken in one session. Multi-PIE addresses some concerns of the PIE dataset, and includes 337 subjects, with multiple sessions (up to 5) for most subjects. Again, frontal images with a neutral expression are used. Each of the 337 subjects has between 20 and 100 images (20 per session). Images in both datasets are registered based on eye locations. For all tests, HR images are 48×48, and in lieu of genuine LR images, they are generated by downsampling and blurring the HR images to 12×12 (see Fig. 1).

The data is split such that each subject has an equal number of images in the gallery, and the remaining images become the probe set. The gallery images are then used for training. For tests on PIE, the best parameters for CMFA

Table 1. Rank 1 identification rates when training on a set of gallery images, and testing on LR and HR probe images. The number in parentheses for Data is the number of gallery images per subject. The number in parentheses for each reported performance level is the standard deviation over 20 runs.

	LR Probe				HR Probe		
Data	SDA	CMFA	MFA-LR	MFA-BL	SDA	CMFA	MFA
PIE (2)	75.07 (2.26)	78.34 (2.56)	**78.58** (2.85)	45.13 (6.11)	78.68 (2.28)	**84.33** (2.39)	83.82 (2.42)
PIE (4)	89.87 (2.14)	92.66 (1.48)	**93.61** (1.41)	38.50 (4.93)	92.26 (2.03)	**94.44** (1.28)	93.30 (1.53)
PIE (6)	95.30 (1.29)	97.83 (0.77)	**98.17** (0.69)	35.10 (5.37)	97.45 (0.87)	**98.47** (0.65)	98.07 (0.72)
Multi-PIE (2)	69.71 (0.93)	**71.61** (0.81)	65.29 (1.10)	9.53 (1.45)	74.18 (1.00)	**75.48** (0.99)	72.08 (1.25)
Multi-PIE (6)	94.00 (0.44)	**94.46** (0.46)	91.80 (0.45)	3.58 (1.03)	**96.32** (0.31)	96.03 (0.46)	87.54 (0.61)
Multi-PIE (10)	96.46 (1.32)	**96.80** (1.16)	95.16 (1.32)	2.90 (0.66)	**98.19** (0.77)	98.13 (0.77)	93.85 (1.18)

was $k_1 = 2$ and k_2 set to a value between 5 and 25. For Multi-PIE, $k_1 = C$ where C equals to the number of training images per class, and k_2 was equal to roughly C^3. Small changes in the value of k_2 do not have a significant impact on performance.

Table 1 shows that CMFA and MFA both outperform SDA on the PIE dataset, and CMFA is the best performing algorithm on the Multi-PIE dataset. The table shows similar trends when testing HR probe images, but it is interesting to note that CMFA and MFA-LR both outperform the baseline MFA in some cases. Fig. 2 shows that on the Multi-PIE dataset it is not necessary to learn many projections to achieve the highest possible performance on LR probe images.

CMFA's better performance compared to SDA indicates that CMFA does a better job discriminating between classes it trains on. This is consistent with the relative performance between MFA and LDA [1].

While the difference between CMFA and SDA is approximately one to two standard deviations, given a particular amount of training, Fig. 3 shows that while the performance does vary between the 20 trials, the relative performance is much more consistent. This indicates that while the selected gallery has an impact on the performance of each algorithm, the advantage of CMFA is not tied to the particular images in the gallery. A paired t-test comparing CMFA and SDA showed that $p < .0001$ for each experiment setting reported in Table 1, indicating a significant difference in their respective performances.

4.3 LR Feature Representation

One of the benefits of coupled mapping is that it can be used for effective LR face matching when images are too small for most feature extraction techniques

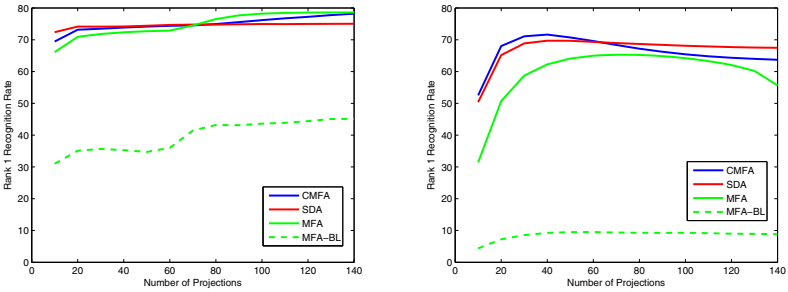

Fig. 2. Performance on LR probe images, with 2 gallery/training images per subject on the PIE dataset (left) and Multi-PIE dataset (right)

used in face recognition. Still, it is possible that some feature representations can be effective for lower resolution images.

While the previous results are based on using the pixels in the original images, we also consider two feature extraction techniques which are designed with the goal of illumination invariance. Simplified Local Binary Pattern (SLBP) considers the number of pixels in the 8-connected neighborhood that have a higher intensity than the pixel being considered [13]. The result for each pixel is an integer value between 0 and 8. We exclude pixels that on the edge of images that do not have an 8-connected neighborhood.

We also consider color ratio (CR) as a means of illumination invariance [14]. CR is defined as

$$I(x,y) = \frac{I(x,y)}{\mu(h(x,y))} \tag{13}$$

where $I(x,y)$ is the pixel in an image, and $h(x,y)$ is a window around $I(x,y)$. We only tested a window size of 3×3. As with SLBP, we do not compute the CR for the pixels on the edge of the image. Thus, for an $K \times K$ image, the corresponding SLBP and CR images used are $K - 2 \times K - 2$. An example of SLBP and CR can be seen in Fig. 5.

We tested these methods on the PIE dataset, using 2 training/gallery images per subject as in Section 4.2. We tested CMFA and SDA over a range of resolutions for the LR probe images, applying SLBP and CR to both the HR and LR images when using them. When testing with CMFA, we used the best parameters learned in earlier test, and did not relearn the values for the new feature representations. Results can be seen in Fig. 5. It is observed that at lower resolutions, just using the pixels is the most effective. However, as the resolution of the probe images become larger, the recognition rate when using pixels does not increase much at all, whereas both CR and SLBP features improve dramatically. Fig. 5 also shows that the as the resolution changes, the performance difference between CMFA and SDA stays nearly constant when using pixels. When using SLBP and CR, the difference between the two algorithms is much smaller.

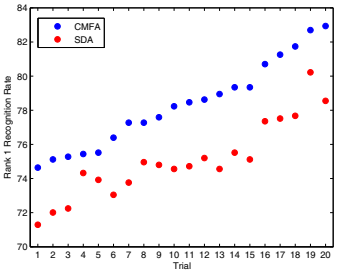

Fig. 3. Performance for each of 20 trials on PIE, using 2 training images per subject. Trials are sorted by CMFA performance.

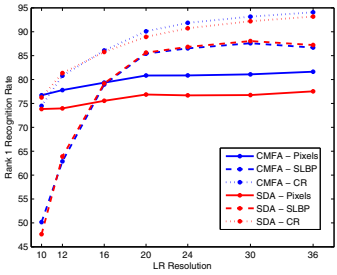

Fig. 4. Performance of different feature representations for coupled mapping. Results are using the PIE dataset, with 2 gallery/training images per subject.

Fig. 5. Sample from PIE dataset showing the original image (left), image using Simplified LBP (center), and image using color ratio (right)

Overall, it seems that using pixels is the best of the three options at very low resolution, while color ratio helps improve performance as the probe image is a higher resolution. Finding the best feature representation for different resolutions can help systems that capture images over a range of different resolutions.

5 Conclusions

The work presents CMFA, an algorithm for performing coupled mapping to match images of different resolutions. CMFA appears to work well when it can learn from the local relationships that will be encountered in testing. CMFA thus provides an alternative solution to LR face matching that can be more desirable depending on the application. In addition, we show that some basic image processing can improve LR face recognition results. In future work we would like to more thoroughly explore different LR feature representations, as well as consider other options for learning projections with coupled mapping.

References

1. Yan, S., Xu, D., Zhang, B., Zhang, H.J., Yang, Q., Lin, S.: Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 40–51 (2007)

2. Zhou, C., Zhang, Z., Yi, D., Lei, Z., Li, S.Z.: Low-resolution face recognition via Simultaneous Discriminant Analysis. In: International Joint Conference on Biometrics (2011)
3. Li, B., Shan, S., Chen, X.: Low-resolution face recognition via coupled locality preserving mappings. IEEE Signal Processing Letters 17, 20–23 (2010)
4. Biswas, S., Bowyer, K.W., Flynn, P.J.: Multidimensional scaling for matching low-resolution face images. IEEE Transactions on Pattern Analysis and Machine Intelligence (forthcoming)
5. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A 4, 519–524 (1987)
6. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience 3, 71–86 (1991)
7. Belhumeur, P.N., Hespanha, J.: Eigenfaces vs. fisherfaces: Recognition using specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 711–720 (1997)
8. Baker, S., Kanade, T.: Hallucinating faces. In: IEEE International Conference on Automatic Face and Gesture Recognition (2000)
9. Zou, W.W.W., Yuen, P.C.: Very low resolution face recognition problem. In: IEEE International Conference on Biometrics: Theory Applications and Systems (2010)
10. Hennings-Yeoman, P.H., Baker, S., Kumar, B.V.: Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In: IEEE Conference on Computer Vision and Pattern Recognition (2008)
11. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression (PIE) database. In: IEEE International Conference on Automatic Face and Gesture Recognition (2002)
12. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. In: IEEE International Conference on Automatic Face and Gesture Recognition (2008)
13. Tao, Q., Veldhuis, R.: Illumination normalization based on simplified local binary patterns for a face verification system. In: Biometrics Symposium at The Biometrics Consortium Conference (2007)
14. Adjeroh, D.A.: On ratio-based color indexing. IEEE Transactions on Image Processing 10, 36–48 (2001)