Misalignment-robust Face Recognition via Efficient Locality-constrained Representation

Yandong Wen, Weiyang Liu, Meng Yang, Member, IEEE, Yuli Fu, Zhifeng Li, Senior Member, IEEE

Abstract—Misaligned face recognition has been studied in the past decades, but still remains an open challenge. To address this problem, we propose a highly efficient misalignment-robust locality-constrained representation (MRLR) algorithm. Specifically, MRLR first aligns the query face via the \( \ell_2 \)-norm locality-constrained representation, and then recognizes it by a standard \( \ell_2 \)-norm collaborative representation algorithm. It achieves a high degree of robustness even with a small training set. Moreover, we take advantage of the block matrix inverse to develop an efficient solving algorithm whose efficiency and scalability are verified by computational complexity analysis. Experimental results on public data sets show that MRLR beats several state-of-the-art approaches in terms of efficiency and scalability with comparable performance.

Index Terms—Face recognition, face alignment, locality-constrained representation, efficient solving algorithm, computational complexity

I. INTRODUCTION

Over the past years, significant progresses have been made in face recognition [1]. Numerous techniques were applied to face recognition, most of which work extremely well in controlled circumstances [2, 3]. However, these algorithms usually fail in uncontrolled or loosely controlled face recognition [4]. Automatic recognition algorithms often suffer from varying illumination, occlusion and misalignment, failing to accurately recognize the query face. Among all these problems, misalignment may be the most challenging one [5], since slight misalignment globally transforms the whole image, while illumination and occlusion only locally deteriorate a part of region.

Wright et al proposed the sparse representation based classification (SRC) [13], which seeks to represent an aligned testing image by the linear sparse combination of training images. However, SRC performs poorly with misaligned faces. To overcome such shortcoming, Huang et al [9] propose the transform-invariant sparse representation (TSR). They add deformations in training set, simultaneously recovering the image transformation and sparse representation. However, TSR aligns testing image to global dictionary and thus easily gets trapped in local minima. To avoid that, robust alignment by sparse representation (RASR) [10] aligns the testing image to training samples of each subject, then warps training set and testing image to a unified transformation for recognition. The subject by subject strategy effectively finds the global optima, but the exhaustive search in every subject is extremely time-consuming, detrimental to efficiency and scalability. Besides, sufficient training samples with varieties of illuminations have to be guaranteed for satisfactory performance [11]. In [12], Yang et al proposed the efficient misalignment-robust representation (MRR) for face recognition. With the carefully controlled training set, they perform the singular value decomposition (SVD) to approximate the global dictionary with principal components, significantly enhancing its real-time ability. However, SVD operation therein is still time and space consuming, preventing MRR from being applied in large-scale datasets. In addition, using the principle component of dictionary instead of the original one inevitably reduces precision of alignment.

To address the above problems, we propose a misalignment-robust locality-constrained representation (MRLR) for misaligned face recognition [10]. Our contributions are summarized as follows. First, we propose to code a deformable query face via locality-constrained representation on carefully controlled gallery images. It effectively avoids substantial local minima and achieves accurate alignment by an analytical solution. Second, we further propose an efficient coding strategy via \( \ell_2 \)-norm, reducing the computational complexity. By making use of the block structure of the deformable dictionary, we accelerates the coding algorithm for the face alignment stage, significantly improving the efficiency and scalability of MRLR.

II. THE PROPOSED METHOD

We first review some necessary preliminaries. Then we present an efficient misalignment-robust locality-constrained representation algorithm for face recognition.

A. Preliminaries

Each grayscale \( a \times b \) frontal face image is stacked into a vector \( \mathbf{d}_j \in \mathbb{R}^n \) \((m = a \times b)\), called atom, which denotes the \( j \)th training sample. Combining all the atoms together constitutes a global dictionary \( \mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_n] \in \mathbb{R}^{m \times n} \). Assume that training samples belong to \( k \) distinct classes, and the gallery images of the \( i \)th class is denoted by a sub-dictionary \( \mathbf{D}_i \). Because frontal faces image usually span on the linear low-dimension space of corresponding subject [11], an aligned
testing sample $y \in \mathbb{R}^m$ can be well represented as a linear combination of atoms in $D$. The robust face recognition can be formulated as the following equation [13]:

$$\min_{x, e} \|x\|_1 + \|e\|_1 \quad \text{s.t.} \quad y = Dx + e$$

from which we can recover the sparse representation $x \in \mathbb{R}^n$, and classify it based on the minimal residual $\|y - D\delta_i(x)\|_2$, where $\delta_i(x)$ is a new vector whose entries are zeros except keeping those in $x$ associated with class $i$. $e \in \mathbb{R}^m$ is the reconstruction error between the recovered and testing image. However, Eq. 1 can only be applied to aligned face recognition. When it comes to misaligned faces, the model in Eq. 1 performs much worse.

Now we consider the misaligned face recognition problem. A warped testing image, denoted by $y_w = y \circ \tau^{-1}$, is subject to some misalignment. $\circ$ denotes a nonlinear operator, transforming the image $y$ by $\tau^{-1}$, where $\tau^{-1}$ belongs to a finite-dimensional group $T$ of transformation in image-plane, e.g., similarity transformation. The testing image $y$ is warped by $\tau^{-1}$ so it does not linearly span on the original subspace. In other words, the sparse representation $x_w$ with respect to $y_w$ may not reveal its label if the testing image is not aligned. However, $\tau^{-1}$ is invertible, which makes it possible to recover the aligned image by $y = y_w \circ \tau$ and remodel the recognition problem. Assume we can estimate $\tau$ first, the model could be reformulated as

$$\min_{x, e} \|x\|_1 + \|e\|_1 \quad \text{s.t.} \quad y_w \circ \tau = Dx + e$$

## B. Misalignment-robust Locality-constrained Representation

Because dealing with occlusion is beyond the scope of this paper, we only consider non-occluded scenario. In this case, we can replace the constraint on error $e$ in Eq. (2) with $\ell_2$ norm without sacrificing performance. Without occlusion, constraining $e$ with $\ell_2$ norm is equivalent to $\ell_1$ norm [10]. To avoid solving a time-consuming $\ell_1$-minimization problem on $x$, we further propose an alternative way to preserve the sparsity. In the light of the locality-constraint linear coding (LLC) [18], we constraining $x$ to be locally concentrated because locality must lead to smooth sparsity. Therefore, we incorporate efficient locality-constrained representation in our formulation, greatly enhancing the efficiency. The model of MRLR is written as

$$\min_{x, e} \|c \odot x\|_2^2 + \|e\|_2^2 \quad \text{s.t.} \quad y_w \circ \tau = Dx + e$$

where the notation $\odot$ means the element-wise multiplication between two vectors, and $c \in \mathbb{R}^{n \times 1}$ is the locality adapter that gives different penalties on the coefficients $x$. So far, the problem in Eq. (3) remains nonlinear. Few existing algorithms can directly deal with it. According to [10, 19], a small deformation in transform can be approximately linearized as $y_w \circ (\tau + \Delta \tau) = y_w \circ \tau + J \Delta \tau$, where $J = \frac{\partial}{\partial \tau} y_w \circ \tau$ is the Jacobian of $y_w \circ \tau$ with respect to $\tau$. If an initial $\tau$ is given, we can repeatedly search for an optimal $\Delta \tau$ to update $\tau$ and $J$. Eventually the final $\tau$ will be obtained to align the warped image $y_w$. The efficiency of the MRLR model lies in two folds. First, we use the $\ell_2$ norm constraints and derive an analytical solution for MRLR, which is much faster than solving $\ell_1$ norm minimization. Second, we take advantage of the block matrix inverse to design a highly efficient algorithm, which obtains exactly the same solution in shorter time. The MRLR algorithm is summarized as follows.

### Algorithm 1: The MRLR Algorithm for Face Recognition

**Input:** A matrix of training samples $D$, a test sample $y_w$ and its initial transformation $\tau$ (it can be obtained by any off-the-shelf face detector, e.g. Viola-Jones detector), belonging to a deformation group $T$, a constant $\sigma$

**Output:** The label of $y_w$

**Step 1: Preprocessing: Offline alignment**
1. Align the training images with each other by RASL [15]

**Step 2: The Query Face Alignment**
2. **Outer Loop:** Align the testing image, obtain the final $\tau$
3. Calculate local adaptor
4. $c = D^2 y; c = \exp(c/\sigma); c = \max(c) - c; j \rightarrow 1$
5. **Inner Loop:** while not converged (or not reach max iteration number) do
   6. $\tilde{y}_w(\tau_{j-1}) = y_w(\tau_{j-1}) / \|y_w - \tau_{j-1}^i\|^2_2$
   7. $\Delta \tau = \arg \min_{\Delta \tau} \|c \odot x\|_2^2 + \|e\|_2^2 \quad \text{s.t.} \quad \tilde{y}_w(\tau_j) + J \Delta \tau = Dx + e$
   8. $\tau_j = \tau_{j-1} + \Delta \tau; j \rightarrow j + 1$
   9. **end**
10. $\tau = \tau_j; \tau_f = \tau_j$
11. **end**
12. Obtain Aligned testing image: $y = y_w \circ \tau$

**Step 3: Face Recognition**
13. $x = \min_{x, e} \|c \odot x\|_2^2 + \|e\|_2^2 \quad \text{s.t.} \quad y = Dx + e$
14. $\text{label}(y) = \arg \min_i \|y - D\delta_i(x)\|_2$

III. **EFFICIENT SOLVING ALGORITHM FOR MRLR VIA BLOCK MATRIX INVERSE**

This section presents a highly efficient solving algorithm to accelerate the MRLR algorithm.

The computational time is dominated by Step 7, with matrix inverse and matrix-matrix multiplication in the analytical solution. We reformulate the equation in Step 7

$$\Delta \tau = \min_{\Delta \tau} \|Cx\|_2^2 + \|e\|_2^2$$

$$s.t. \quad \tilde{y}_w + J \Delta \tau = Dx + e$$

where $C$ is a diagonal matrix, filled with zero entries except the diagonal elements equal to the locality adaptor vector $c$. We can further substitute $e = \tilde{y}_w - [D, -J] \left[ \begin{array}{c} x \\ \Delta \tau \end{array} \right]$ into Eq. (4), we have

$$\Delta \tau = \arg \min_{x, \Delta \tau} \|\tilde{y}_w - [D, -J] \left[ \begin{array}{c} x \\ \Delta \tau \end{array} \right]\|_2^2$$

$$= \arg \min_{x, \Delta \tau} \left[ \|\tilde{y}_w - D \left[ \begin{array}{c} x \\ \Delta \tau \end{array} \right]\|_2^2 - \|\tilde{y}_w - [D, -J] \left[ \begin{array}{c} x \\ \Delta \tau \end{array} \right]\|_2^2 \right]$$

$$= \arg \min_x \|u - Rz\|_2^2$$

(5)
where \( \mathbf{u} \), \( \mathbf{R} \) and \( \mathbf{z} \) denote \( \begin{bmatrix} \mathbf{y}_w \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{D} \\ \mathbf{C} \end{bmatrix} \) and \( \begin{bmatrix} \mathbf{x} \\ \Delta \mathbf{r} \end{bmatrix} \) respectively. It becomes a least square problem whose analytical solution is \( \mathbf{z} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{u} \). As one can see, the computational complexity is still high due to the large size of \( \mathbf{R} \).

A. Efficient Algorithm to Compute \( \Delta \mathbf{r} \)

Actually, the efficiency and the scalability can be greatly boosted if we make good use of the block structure of \( \mathbf{R} \). We introduce a lemma for block matrix inverse.

**Lemma 1.** Let \( \mathbf{A}_{ij} \) denote the \( ij \)th block of the matrix \( \mathbf{A} \). The inverse can be expressed by the use of

\[
\begin{bmatrix}
\mathbf{A}_{11} & \mathbf{A}_{12} \\
\mathbf{A}_{21} & \mathbf{A}_{22}
\end{bmatrix}^{-1} =
\begin{bmatrix}
\mathbf{I}^{\mathbf{A}_{11}} & -\mathbf{A}_{12} \mathbf{A}_{21}^{-1} \\
-\mathbf{A}_{22} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22}^{-1}
\end{bmatrix}
\]

as \( \mathbf{Z}_1 = \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \) and \( \mathbf{Z}_2 = \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \). Then we rewrite the analytical solution as

\[
\mathbf{z} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{u}
\]

\[
= \begin{bmatrix}
\mathbf{D}^T \\ -\mathbf{J}^T \\
\end{bmatrix} \begin{bmatrix}
\mathbf{D} & \mathbf{C} \\
\mathbf{C} & 0
\end{bmatrix}^{-1} \begin{bmatrix}
\mathbf{D}^T \\ -\mathbf{J}^T \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\mathbf{D}^T \mathbf{D} + \mathbf{C}^T \mathbf{C} & -\mathbf{D}^T \mathbf{J} \\
-\mathbf{J}^T \mathbf{D} & \mathbf{J}^T \mathbf{J}
\end{bmatrix}^{-1} \begin{bmatrix}
\mathbf{D}^T \\ -\mathbf{J}^T \\
\end{bmatrix}
\]

\[
\mathbf{z} = \mathbf{Z}_1^{-1} \begin{bmatrix}
\mathbf{D}^T \mathbf{D} + \mathbf{C}^T \mathbf{C} \\
\mathbf{J}^T \mathbf{D} \times \mathbf{Z}_2^{-1}
\end{bmatrix} \begin{bmatrix}
\mathbf{D}^T \\ -\mathbf{J}^T \\
\end{bmatrix}
\]

We denote \( \mathbf{D}^T \mathbf{D} + \mathbf{C}^T \mathbf{C}, \mathbf{D}^T \mathbf{J} \) and \( \mathbf{J}^T \mathbf{J} \) as \( \mathbf{T}_1, \mathbf{T}_2 \) and \( \mathbf{T}_3 \) respectively. In particular, \( \mathbf{T}_1 = \mathbf{R}^\top \mathbf{R} \) can be pre-calculated before the inner iteration from Step 5 to Step 9. The other variables can be represented as \( \mathbf{Z}_1^{-1} = (\mathbf{T}_1 - \mathbf{T}_2 \mathbf{R}^{-\top} \mathbf{T}_2^{-1})^{-1} \) and \( \mathbf{Z}_2^{-1} = (\mathbf{T}_3 - \mathbf{T}_2 \mathbf{R}^{-\top} \mathbf{T}_2^{-1})^{-1} \). Eq. (6) can be represented as

\[
\mathbf{z} = \begin{bmatrix}
\mathbf{x} \\
\Delta \mathbf{r}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\mathbf{Z}_1^{-1} \mathbf{D}^T \mathbf{y}_w - \mathbf{T}_1^{-1} \mathbf{T}_2 \mathbf{Z}_2^{-1} \mathbf{J}^T \mathbf{y}_w \\
\mathbf{Z}_2^{-1} \mathbf{Z}_1^{-1} \mathbf{D}^T \mathbf{y}_w - \mathbf{Z}_1^{-1} \mathbf{T}_2 \mathbf{Z}_2^{-1} \mathbf{J}^T \mathbf{y}_w
\end{bmatrix}
\]

Note that the purpose of the face alignment part is to search for a deformation step \( \Delta \mathbf{r} \), so computing \( \mathbf{x} \) is unnecessary. When not computing \( \mathbf{x} \), much computation can be saved. Moreover, as mentioned in [18], since \( \mathbf{c} \) usually impose weak constraint on only a few atoms, suppressing most of the atoms. We can simply keep the smallest \( s \) (\( s \ll n \)) entries and force others to be positive infinity. This strategy further accelerates the coding, as we present in complexity analysis and experiments (This strategy is termed as MRLR2, while the former proposed one is termed as MRLR1). For MRLR1, Eq. (6) can be further accelerated via \( (\mathbf{M} + \mathbf{c} \mathbf{c}^\top)^{-1} = \mathbf{M}^{-1} - (\mathbf{M}^{-1} \mathbf{c} - \mathbf{c} \mathbf{M}^{-1} \mathbf{c}^\top)^{-1} \mathbf{c} \mathbf{M}^{-1} \mathbf{c} \) where \( \mathbf{M} \) equals \( \mathbf{D}^T \mathbf{D} \) and its inverse can be pre-calculated.

B. Computational Complexity Analysis

This section analyzes the computational complexity before and after optimized for clear comparison. Obviously, the computation of inner iteration dominate the whole complexity of the proposed algorithm. Solving Eq. (5) via \( \mathbf{z} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{u} \) needs a matrix-matrix multiplication, matrix inverse and two matrix-vector multiplications. The complexity adds up to \( O(n^2 + mn^2 + pm^2 + p^2n + pm + p^3) \). The constant \( p \) is the dimension of \( \mathbf{r} \), which is determined by the transformation group. (e.g. 4 for similarity transformation, 6 for affine transformation, so \( p = 4 \) in this paper). In the proposed efficient algorithm, we only compute \( \Delta \mathbf{r} \) each inner iteration. Noted that \( \mathbf{T}_1^{-1} \) can be pre-calculated in every outer iteration, we merely need to update \( \mathbf{T}_2, \mathbf{T}_3 \) and \( \mathbf{Z}_2^{-1} \) to compute \( \Delta \mathbf{r} \) with the complexity of \( O(3n^2 + mn^2) \) to \( O(n^2 + mn) \) for MRLR1. Furthermore, the MRLR2 only remains \( s \) atoms in locality-constrained dictionary, its complexity can be accordingly inferred as \( O(ps^2 + pm*) \), which is much less than the original one.

IV. Experiments

Three experiments on benchmark database are presented. The running time of alignment via MRLR is greatly reduced, e.g. MRLR2 is roughly 2 to 10 times faster than MRR. In addition, MRLR2 scales well in large-scale database.

A. Experimental details

In Algorithm 1, the \( \sigma \) is fixed to 0.1 and the parameter \( s \) in MRLR2 is 20. The maximum iteration of outer and inner loop are 3 and 30, respectively. For MRR, the two parameters \( \eta_1 \) and \( \eta_2 \), denoting the length of principle vector are fixed to 25. Besides, the number of remained class in MRR and RASR are set to 10. It is also crucial to state that we use only one project matrix of 500 rows in TSR, which controls the tradeoff between accuracy and speed [10]. The \( \ell_1 \)-minimization is based on Augmented Lagrange Multiplier (ALM) [21]. Matlab code is available at [https://sites.google.com/site/weiyangliu92/](https://sites.google.com/site/weiyangliu92/).

B. The region of attraction

The region of attraction evaluates the robustness of an approach to 2D deformations. We compare MRLR with MRR [12], RASR [10] and TSR [9] on Extended Yale B database [22], which includes 2414 images of 38 subjects. We use the uncropped images of 28 subjects in experiments. 32 training images of each subject are randomly selected, and the rest are using for testing. All the training images are resized to \( 80 \times 70 \). We get access to the group truth of eyes and artificially add perturbation to them. Then we calculate the corresponding recognition rates under various initial transformations.
One can see that all the approaches perform well within a certain range of misalignment, e.g. 20 percent translation in x direction (7–8 pixels), or 20 percent scale variation. All of them could work fine in practical scenario if the detector falls safely inside the region of attraction. However, TSR performs relatively poor even in small perturbations. It is mainly because aligning testing image to all training images is more prone to local minima, resulting in inaccurate alignment. Compared to MRLR1, MRLR2 and RASR, MRR performs slightly worse on robustness to deformation. Because the performance of MRR largely depends on reasonably adjusting \( \eta_x \) and \( \eta_y \). With a fixed strategy, it could not adaptively deal with various query images. MRLR1 and MRLR2 perform almost the same as RASR, demonstrating that locality-constrained representation effectively avoids local minima. In other word, we could recover the accurate transformations by locality-constrained representation.

C. Practical Recognition and Running Time

We also perform practical recognition experiments on both Extended Yale B and CAS-PEAL database [23]. For Extended Yale B database, we adopt the same settings in section B. For CAS-PEAL, 20 subjects were chosen, each of them including more than 32 images. We randomly selected 20 images from each subject and resized them to 80 × 70 for training, then test on the left 12 images. The initial \( \tau \) are automatically given by Viola-Jones detector [24]. Table 2 gives the recognition rates and average running time.

### Table I: Recognition Rates (%) and Consuming Time (s)

|                | Extended Yale B | CAS-PEAL |
|----------------|-----------------|----------|
|                | Recognition Rate | Average Time | Recognition Rate | Average Time |
| MRLR1          | 91.53           | 0.6207    | 89.57           | 0.3307      |
| MRLR2          | 91.68           | 0.1783    | 90.43           | 0.1462      |
| MRR            | 89.34           | 0.7773    | 90.05           | 0.5684      |
| RASR           | 93.72           | 9.7587    | 87.83           | 5.4466      |
| TSR            | 81.61           | 7.396     | 86.96           | 4.2695      |

With sufficient training samples (32 images each subject), RASR achieves highest recognition rate, 93.72% in Extended Yale B database. However, such subject by subject searching is time-consuming, it averagely costs 9.76 second on each testing image when the amount of subjects is 28. The recognition rate of MRLR2 is 91.68%, but it takes only 0.18 second to deal with a testing image, roughly 4, 55 and 41 times faster than MRR, RASR and TSR respectively. In CAS-PEAL, the dictionary is constituted by 20 training images each subject. RASR performs not so well (87.83%), because RASR requires varieties of training samples in each subject for representation. So its performance gets worse with less gallery images. However, Table 1 shows that MRLR1 and MRLR2 are not as sensitive as RASR on small samples issue, mainly due to the collaborative representation in MRLR. Specifically, MRLR2 achieves the best performance 90.43%, with the least average running time 0.15 seconds.

D. Scalability

We vary the number of subject from 10 to 100 and resize the images from 40 × 35 to 160 × 140, for the purpose of evaluating the scalability of our algorithm. Table 1 and table 2 show the experimental results.

### Table II: Average Running Time (s) Under Different Dimension

|                | 40 × 35 | 64 × 56 | 80 × 70 | 120 × 105 | 160 × 140 |
|----------------|---------|---------|---------|-----------|-----------|
| MRLR1          | 0.0851  | 0.1947  | 0.3307  | 0.5689    | 0.9401    |
| MRLR2          | 0.0656  | 0.1175  | 0.1462  | 0.3034    | 0.5048    |
| MRR            | 0.1325  | 0.3417  | 0.5933  | 2.2585    | 5.9970    |
| RASR           | 3.0698  | 4.5312  | 5.4466  | 10.4468   | 17.1737   |
| TSR            | 3.6445  | 3.8614  | 4.2695  | 4.6718    | 5.4682    |

### Table III: Average Running Time (s) Under Different Amount of Subjects

|                | 10 | 20 | 40 | 70 | 100 |
|----------------|----|----|----|----|-----|
| MRLR1          | 0.1623 | 0.2554 | 0.4244 | 0.6856 | 0.9403 |
| MRLR2          | 0.1318 | 0.1373 | 0.1559 | 0.197 | 0.2616 |
| MRR            | 0.5776 | 0.5928 | 0.6082 | 0.6394 | 0.6994 |
| RASR           | 2.7377 | 5.2358 | 10.255 | 16.943 | 24.074 |
| TSR            | 2.1533 | 3.2825 | 5.5280 | 8.4034 | 11.532 |

It is clear that RASR and TSR cost too much time, far from being applicable in real-time systems. The running time of TSR remains relatively stable as the dimension increases, but rises linearly with more subjects. MRR maintains excellent real-time capability with the growth of the subject number. However, its running time rises dramatically when the resolution of image increasing. Unlike the abovementioned approaches, MRLR1 and MRLR2 are not very sensitive to the dimension or number of subjects, preserving competitive performance. Specially, MRLR2 costs the least running time and the lowest increasing rate as we enlarge the dimension or number of subjects, showing the best scalability among state-of-the-art approaches.

V. CONCLUSION

In this paper, we propose an effective misalignment-robust algorithm, MRLR, for misaligned face recognition. By taking advantage of the pre-aligning and locality-constrained representation, a loosely controlled image can be accurately aligned for the subsequent recognition. Compared to the existing approaches, locality-constrained representation is not so sensitive to the small sample number per subject. Moreover, motivated by the block structure of dictionary, we propose an efficient solving algorithm to speed up the alignment. Theoretical computational complexity analysis and extensive experiments show that MRLR considerably reduce the running time without sacrificing the performance.
REFERENCES

[1] Jafri R, Arabnia H R. A survey of face recognition techniques[J]. journal of information processing systems, 2009, 5(2): 41-68.
[2] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: Application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037-2041.
[3] S. Li, X. Hou, H. Zhang. Learning spatially localized, parts-based representation, in: Conference on Computer Vision and Pattern Recognition, IEEE, 2001, pp. 207-212.
[4] Shan S, Chang Y, Gao W, et al. Curse of mis-alignment in face recognition: problem and a novel mis-alignment learning solution[C]//Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on. IEEE, 2004; 314-320.
[5] Waheed K, Staszewski R B, Rezeq S. Curse of digital polar transmission: Precise delay alignment in amplitude and phase modulation paths[C]//Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on. IEEE, 2008: 3142-3145.
[6] Amberg B, Blake A, Fitzgibbon A, et al. Reconstructing high quality face-surfaces using model based stereo[C]//Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on. IEEE, 2007: 1-8.
[7] Blanz V, Vetter T. Face recognition based on fitting a 3D morphable model[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2003, 25(9): 1063-1074.
[8] Cootes T F, Edwards G J, Taylor C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2001 (6): 681-685.
[9] Huang J, Huang X, Metaxas D. Simultaneous image transformation and sparse representation recovery[C]//Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008: 1-8.
[10] Wagner A, Wright J, Ganesh A, et al. Toward a practical face recognition system: Robust alignment and illumination by sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(2): 372-386.
[11] Basri R, Jacobs D W. Lambertian reflectance and linear subspaces[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2003, 25(2): 218-233.
[12] Yang M, Zhang L, Zhang D. Efficient misalignment-robust representation for real-time face recognition[M]//Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012: 850-863.
[13] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2009, 31(2): 210-227.
[14] 234.
[15] Peng Y, Ganesh A, Wright J, et al. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(11): 2233-2246.
[16] Wen Y, Liu W, Yang M, et al. Structured Occlusion Coding for Robust Face Recognition[J]. arXiv preprint arXiv:1502.00478, 2015.
[17] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: Which helps face recognition?[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 471-478.
[18] Wang J, Yang J, Yu K, et al. Locality-constrained linear coding for image classification[C]//Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010: 3360-3367.
[19] Baker S, Matthews I. Lucas-kanade 20 years on: A unifying framework[J]. International journal of computer vision, 2004, 56(3): 221-255.
[20] Xiong X, De la Torre F. Supervised descent method and its applications to face alignment[C]//Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013: 532-539.
[21] Yang A Y, Sastry S S, Ganesh A, et al. Fast ℓ 1-minimization algorithms and an application in robust face recognition: A review[C]//Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, 2010: 1849-1852.
[22] Georghiades A S, Belhumeur P N, Kriegman D J. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2001, 23(6): 643-660.
[23] Gao W, Cao B, Shan S, et al. The CAS-PEAL large-scale Chinese face database and baseline evaluations[J]. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2008, 38(1): 149-161.
[24] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001, 1: 1-511-I-518 vol. 1.