EFFICIENT MISALIGNMENT-ROBUST FACE RECOGNITION VIA LOCALITY-CONSTRAINED REPRESENTATION

Yandong Wen[1], *Weiyang Liu*[2], *Meng Yang*[3*], *and Ming Li*[4,5]

[1]School of Electronic and Information Engineering, South China University of Technology, China
[2]School of Electronic and Computer Engineering, Peking University, China
[3]College of Computer Science & Software Engineering, Shenzhen University, China
[4]SYSU-CMU Joint Institute of Engineering, Sun Yat-sen University, China
[5]SYSU-CMU Shunde International Joint Research Institute, China

ABSTRACT

Current prevailing approaches for misaligned face recognition achieve satisfactory accuracy. However, the efficiency and scalability have not yet been well addressed, which limits their applications in practical systems. To address this problem, we propose a highly efficient algorithm for misaligned face recognition, namely misalignment-robust locality-constrained representation (MRLR). Specifically, MRLR aligns the query face by appropriately harnessing the locality constraint in representation. Since MRLR avoids the exhaustive subject-by-subject search in datasets and complex operation on large matrix, the efficiency is significantly boosted. Moreover, we take the advantage of the block structure in dictionary to accelerate the derived analytical solution, making the algorithm more scalable to the large-scale datasets. Experimental results on public datasets show that MRLR substantially improves the efficiency and scalability with even better accuracy.

Index Terms— Face recognition, misalignment-robust, efficiency, scalability

1. INTRODUCTION

This paper aims to propose a highly efficient and scalable misaligned face recognition algorithm. Since the success of Sparse Representation Classification (SRC) [1] and its variants [2, 3, 4, 5], the accuracy and robustness of face recognition have been greatly improved. However, the computational cost increases in unconstrained environment. For example, [5] and [6] use an exhaustive subject-by-subject search strategy to simultaneously align and recognize the query face. It is extremely time-consuming and impractical for a real-world systems. For this reason, it is crucial to accelerate the computation of misaligned face recognition.

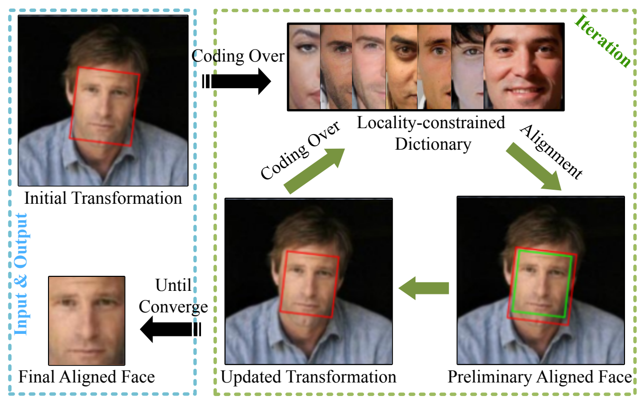

Fig. 1: An Illustration of MRLR. The red bounding box denotes the input estimate and the green one is the output estimate. MRLR works iteratively.

There are a few previous work along this research line. Transform-invariant sparse representation (TSR) [7] adds deformations in training set, simultaneously recovering the image transformation and representation coefficients. However, solving l_1 minimization problem on huge matrix limits its efficiency. Robust alignment by sparse representation (RASR) [5] and sparse illumination transfer (SIT) [6] aligns the testing image to training samples of each subject, then warps training set and testing image to a unified transformation for recognition. The exhaustive search effectively finds the global optima, but it is extremely time-consuming, especially when the subject number is large. [8] proposed the efficient misalignment-robust representation (MRR) for face recognition. By performing singular value decomposition (SVD), they use the principal components of training set for representation and reduce the computational complexity. However, SVD operation therein is still time and memory consuming, preventing MRR from being applied in large-scale datasets.

The aforementioned approaches ignore the high similarity among faces of different identities. Instead of using the training samples of every subject, we propose to use the locality-constrained training set for alignment. With appropriate locality-constrained coefficients regularization, the

*corresponding author. This work is partially supported by the National Natural Science Foundation for Young Scientists of China (Grant no. 61402289) and Shenzhen Scientific Research and Development Funding Program (Grant no. JCYJ20140509172609171).

training set becomes more adaptive to the testing image. In the course of alignment, we update the locality-constrained coefficients according to the latest transformation. After several iterations, the algorithm eventually converges to the accurate transformation, as demonstrated by experiments. Our major contributions are summarized as follows:

- We propose a misalignment-robust locality-constrained representation (MRLR) for robust face recognition, as illustrated in Fig. 1. MRLR avoids the exhaustive search in every subject of the training set, greatly reducing the computational time.
- An analytical solution for the optimization could be derived. Moreover, we could further simplify the solution by making use of the block structure of the deformable matrix. The simplified solution is not sensitive to the scale of training set and make this approach scalable.
- The proposed algorithm do not sacrifice any accuracy performance, which is demonstrated by experiments.

2. PRELIMINARIES

Each grayscale frontal face image is stacked into a vector $\boldsymbol{d} \in \mathbb{R}^m$, called atom. Combining n atoms together constitutes a global dictionary $D = [\boldsymbol{d}_1, \boldsymbol{d}_2, ..., \boldsymbol{d}_n] \in \mathbb{R}^{m\times n}$. Assume that training samples belong to k distinct classes, and the gallery images of the ith class is denoted by a sub-dictionary D_i . According to [9], an aligned testing sample $\boldsymbol{y} \in \mathbb{R}^m$ can be well represented as a linear combination of atoms in D . The robust face recognition can be formulated as

$$\min_{\boldsymbol{x},\boldsymbol{e}} \|\boldsymbol{x}\|_1 + \|\boldsymbol{e}\|_2 \quad \text{s.t. } \boldsymbol{y} = D\boldsymbol{x} + \boldsymbol{e} \tag{1}$$

from which we can recover the sparse representation $\boldsymbol{x} \in \mathbb{R}^n$, and classify it based on the minimal residual $\|\boldsymbol{y} - D_i\delta_i(\boldsymbol{x})\|_2$, where $\delta_i(\boldsymbol{x})$ is a new vector whose entries are zeros expect keeping those in $\boldsymbol{x}$ associated with class i . $\boldsymbol{e} \in \mathbb{R}^m$ is the reconstruction error between the recovered and testing image.

Now we consider the misaligned face recognition problem. A warped testing image, denoted by $\boldsymbol{y}_w = \boldsymbol{y} \circ \boldsymbol{\tau}^{-1}$, is subject to some misalignment. $\circ$ denotes a nonlinear operator, transforming the image $\boldsymbol{y}$ by $\boldsymbol{\tau}^{-1}$, where $\boldsymbol{\tau}^{-1}$ belongs to a finite-dimensional group T of transformation in image-plane, e.g. similarity transformation. The warped testing image $\boldsymbol{y}_w$ does not linearly span on the original subspace. In other words, the sparse representation $\boldsymbol{x}_w$ with respect to $\boldsymbol{y}_w$ may not reveal its label. [5] proposed to exhaustive search the transformation $\boldsymbol{\tau}$ in sub-dictionary of each subject for alignment, as formulated in Eq. (2).

$$\min_{\boldsymbol{x},\boldsymbol{e}} \|\boldsymbol{x}\|_1 + \|\boldsymbol{e}\|_2 \quad \text{s.t. } \boldsymbol{y}_w \circ \boldsymbol{\tau} = D_i\boldsymbol{x} + \boldsymbol{e} \tag{2}$$

Although RASR performs well in misaligned face recognition, it is extremely time- and computation-consuming. We can easily see that Eq. (2) need to repeatedly and iteratively solve ℓ_1 minimization problem in k subjects. In this recognition framework, face alignment dominates the time complexity. Therefore an efficient alignment method for face recognition is the key to reduce computational time.

3. THE PROPOSED METHOD

3.1. Formulation of MRLR

The problem in Eq. (2) is nonlinear. According to [10, 19], a small deformation in transform can be approximately linearized as $\boldsymbol{y}_w \circ (\boldsymbol{\tau} + \Delta\boldsymbol{\tau}) = \boldsymbol{y}_w \circ \boldsymbol{\tau} + J\Delta\boldsymbol{\tau}$, where $J = \frac{\partial}{\partial r}\boldsymbol{y}_w \circ \boldsymbol{\tau}$ is the Jacobian of $y_w \circ \boldsymbol{\tau}$ with respect to $\boldsymbol{\tau}$. If an initial $\boldsymbol{\tau}$ is given, we can repeatedly search for an optimal $\Delta\boldsymbol{\tau}$ to update $\boldsymbol{\tau}$ and J. Eventually the final $\boldsymbol{\tau}$ will be obtained to align the warped image $\boldsymbol{y}_w$. We replace ℓ_2 norm constraint by local constraint and obtain the formulation of MRLR.

$$\min_{\boldsymbol{x},\boldsymbol{e}} \|\boldsymbol{c} \odot \boldsymbol{x}\|_2^2 + \|\boldsymbol{e}\|_2^2 \quad s.t. \quad \boldsymbol{y}_w + J\Delta\boldsymbol{\tau} = D\boldsymbol{x} + \boldsymbol{e} \tag{3}$$

where the notation $\odot$ means the element-wise multiplication between two vectors, and $\boldsymbol{c} \in \mathbb{R}^{n\times 1}$ is the locality adaptor that gives different penalties on the coefficients $\boldsymbol{x}$.

Different from [10, 11], which utilized local constraint for recognition, MRLR explicitly incorporate the locality in the dictionary construction for alignment. The efficiency of the MRLR model lies in two folds. First, An analytical solution for MRLR can be derived , which is much faster than solving ℓ_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 Alignment

Require: The dictionary of training samples $\boldsymbol{D}$, the warped testing image $\boldsymbol{y}_w$, the initial transformation $\boldsymbol{\tau}$ (it can be obtained by any off-the-shelf face detector, e.g. Viola-Jones detector), a constant σ.
Ensure: The aligned face $\boldsymbol{y}$

1: **while** not converge or reach maximal iteration **do**
2: Compute the locality adaptor: $\boldsymbol{c} \leftarrow \exp(\frac{\boldsymbol{D}^T\boldsymbol{y}}{\sigma})$, for all i, $\boldsymbol{c}_i \leftarrow max(\boldsymbol{c}) - \boldsymbol{c}_i$.
3: $j \leftarrow 1$.
4: **while** not converge or reach maximal iteration **do**
5: $\hat{\boldsymbol{y}}_w(\boldsymbol{\tau}_{j-1}) \leftarrow \frac{\boldsymbol{y}_w \circ \boldsymbol{\tau}_{j-1}}{\|\boldsymbol{y}_w \circ \boldsymbol{\tau}_{j-1}\|_2}$, $J \leftarrow \frac{\partial}{\partial \boldsymbol{\tau}_{j-1}}\hat{\boldsymbol{y}}_w(\boldsymbol{\tau}_{j-1})|_{\boldsymbol{\tau}_{j-1}}$.
6: $\Delta\boldsymbol{\tau} = \arg\min_{\Delta\boldsymbol{\tau},\boldsymbol{x},\boldsymbol{e}} \|\boldsymbol{c} \odot \boldsymbol{x}\|_2^2 + \|\boldsymbol{e}\|_2^2$ s.t. $\hat{\boldsymbol{y}}_w(\boldsymbol{\tau}_j) + J\Delta\boldsymbol{\tau} = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{e}$.
7: $\boldsymbol{\tau}_j \leftarrow \boldsymbol{\tau}_{j-1} + \Delta\boldsymbol{\tau}$.
8: $j \leftarrow j + 1$.
9: **end while**
10: $\boldsymbol{\tau} \leftarrow \boldsymbol{\tau}_j$, $\boldsymbol{\tau}_0 \leftarrow \boldsymbol{\tau}_j$.
11: **end while**
12: Output the final aligned face $\boldsymbol{y} = \boldsymbol{y}_w \circ \boldsymbol{e}$.

3.2. Efficient Solving Algorithm

This section presents a highly efficient solution for the MRLR algorithm. By analyzing Algorithm 1, we find the optimization in step 6 dominates the overall computational time. Although it has an analytical solution, it contains the inversion operation of a large-size matrix. We aim to take advantage of the block structure of the matrix to decompose the inversion. We first reformulate the optimization in Step 6 as

$$\Delta\boldsymbol{\tau} = \min_{\Delta\tau} \|\boldsymbol{C}\boldsymbol{x}\|_2^2 + \|\boldsymbol{e}\|_2^2 \quad s.t.\ \hat{\boldsymbol{y}}_w + J\Delta\tau = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{e} \tag{4}$$

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

$$\begin{aligned}
\Delta\tau &= \arg\min_{\boldsymbol{x},\Delta\tau} \|\boldsymbol{C}\boldsymbol{x}\|_2^2 + \left\|\hat{\boldsymbol{y}}_w - [\boldsymbol{D}, -J]\left[\begin{array}{c} \boldsymbol{x} \\ \Delta\tau \end{array}\right]\right\|_2^2 \\
&= \arg\min_{\boldsymbol{x},\Delta\tau} \left\|\left[\begin{array}{c} \hat{\boldsymbol{y}}_w \\ \boldsymbol{0} \end{array}\right] - \left[\begin{array}{cc} \boldsymbol{D} & -J \\ \boldsymbol{C} & \boldsymbol{0} \end{array}\right]\left[\begin{array}{c} \boldsymbol{x} \\ \Delta\tau \end{array}\right]\right\|_2^2 \\
&= \arg\min_{\boldsymbol{z}} \|\boldsymbol{u} - R\boldsymbol{z}\|_2^2
\end{aligned} \tag{5}$$

where $\boldsymbol{u}$, $\boldsymbol{R}$ and $\boldsymbol{z}$ denote $\left[\begin{array}{c} \hat{\boldsymbol{y}}_w \\ \boldsymbol{0} \end{array}\right]$, $\left[\begin{array}{cc} \boldsymbol{D} & -J \\ \boldsymbol{C} & \boldsymbol{0} \end{array}\right]$ and $\left[\begin{array}{c} \boldsymbol{x} \\ \Delta\tau \end{array}\right]$ respectively. It becomes a least square problem whose analytical solution is $\boldsymbol{z} = (\boldsymbol{R}^T\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{u}$. As one can see, the computational complexity is still high due to the large size of $\boldsymbol{R}$. Actually, the efficiency and the scalability can be greatly boosted if we make good use of the block structure of the matrix $\boldsymbol{R}$.

Using the block matrix inversion, we can rewrite the analytical solution as

$$\begin{aligned}
\boldsymbol{z} &= (\boldsymbol{R}^T\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{u} \\
&= \left(\left[\begin{array}{cc} \boldsymbol{D}^T & \boldsymbol{C}^T \\ -J^T & \boldsymbol{0} \end{array}\right]\left[\begin{array}{cc} \boldsymbol{D} & -J \\ \boldsymbol{C} & \boldsymbol{0} \end{array}\right]\right)^{-1}\left[\begin{array}{cc} \boldsymbol{D}^T & \boldsymbol{C} \\ -J^T & \boldsymbol{0} \end{array}\right]\left[\begin{array}{c} \hat{\boldsymbol{y}}_w \\ \boldsymbol{0} \end{array}\right] \\
&= \left[\begin{array}{cc} \boldsymbol{D}^T\boldsymbol{D} + \boldsymbol{C}^T\boldsymbol{C} & -\boldsymbol{D}^T J \\ -J^T\boldsymbol{D} & J^T J \end{array}\right]^{-1}\left[\begin{array}{c} \boldsymbol{D}^T \\ -J^T \end{array}\right]\hat{\boldsymbol{y}}_w \\
&= \left[\begin{array}{cc} \boldsymbol{Z}_1^{-1} & \begin{array}{c}(\boldsymbol{D}^T\boldsymbol{D} + \boldsymbol{C}^T\boldsymbol{C})^{-1} \\ \times(\boldsymbol{D}^T J)\boldsymbol{Z}_2^{-1}\end{array} \\ \begin{array}{c}\boldsymbol{Z}_2^{-1}(J^T\boldsymbol{D})\times \\ (\boldsymbol{D}^T\boldsymbol{D} + \boldsymbol{C}^T\boldsymbol{C})^{-1}\end{array} & \boldsymbol{Z}_2^{-1} \end{array}\right]\left[\begin{array}{c} \boldsymbol{D}^T \\ -J^T \end{array}\right]\hat{\boldsymbol{y}}_w
\end{aligned} \tag{6}$$

We denote $\boldsymbol{D}^T\boldsymbol{D} + \boldsymbol{C}^T\boldsymbol{C}$, $\boldsymbol{D}^T J$ and $J^T J$ as $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$ respectively. In particular, $\boldsymbol{T}_1$ and $\boldsymbol{T}_1^{-1}$ can be pre-calculated before the inner iteration from Step 4 to Step 9. The other variables $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ can be represented as $\boldsymbol{Z}_1^{-1} = (\boldsymbol{T}_1 - \boldsymbol{T}_2\boldsymbol{T}_3^{-1}\boldsymbol{T}_2^T)^{-1}$ and $\boldsymbol{Z}_2^{-1} = (\boldsymbol{T}_3 - \boldsymbol{T}_2^T\boldsymbol{T}_1^{-1}\boldsymbol{T}_2)^{-1}$. Eq. (6) can be represented as

$$\boldsymbol{z} = \left[\begin{array}{c} \boldsymbol{x} \\ \Delta\tau \end{array}\right] = \left[\begin{array}{c} \boldsymbol{Z}_1^{-1}(\boldsymbol{D}^T\hat{\boldsymbol{y}}_w) - \boldsymbol{T}_1^{-1}\boldsymbol{T}_2\boldsymbol{Z}_2^{-1}(J^T\hat{\boldsymbol{y}}_w) \\ \boldsymbol{Z}_2^{-1}\boldsymbol{T}_2^T\boldsymbol{T}_1^{-1}(\boldsymbol{D}^T\hat{\boldsymbol{y}}_w) - \boldsymbol{Z}_2^{-1}(J^T\hat{\boldsymbol{y}}_w) \end{array}\right] \tag{7}$$

Note that the purpose of the face alignment is to search a deformation step $\Delta\boldsymbol{\tau}$, so computing $\boldsymbol{x}$ is unnecessary. Without computing $\boldsymbol{x}$, we can save greatly reduce the computation. Moreover, as mentioned in [12], since $\boldsymbol{c}$ usually imposes weak constraint on only a few atoms, suppressing most of the atoms. We can simply keep the smallest $s, (s \ll n)$ entries in $\boldsymbol{c}$ and force other entries to be positive infinity. It further accelerates the coding (This strategy is termed as M-RLR2, while the former proposed one is termed as MRLR1). We can summarize that the computational complexity is decreased from $\mathcal{O}\left(n^3 + mn^2\right)$ to $\mathcal{O}\left(pn^2 + pmn\right)$ for MRLR1. The constant p is the dimension of $\boldsymbol{\tau}$, which is determined by the transformation group. (e.g. 4 for similarity transformation, 6 for affine transformation, so $p = 4$ in this paper). Furthermore, the MRLR2 only remains s atoms in locality-constrained dictionary, its complexity can be accordingly inferred as $\mathcal{O}\left(ps^2 + pms\right)$, much less than the original one.

4. EXPERIMENTS

We conduct experiments on face database in constrained and unconstrained conditions, including Extended Yale B [13], CAS-PEAL [14] and Labeled Faces in the Wild (LFW) dataset [15]. Comprehensive evaluations on MRLR in terms of region of attraction, recognition rate, running time and scalability are presented. The experimental results show that MRLR achieves competitive performance with much less running time, and scales better in large datasets.

4.1. Implementation details

In MRLR2 and MRR, the length (the number of atoms) of dictionary for alignment is fixed to 20 for fair comparison. We basically follow the same setting in [5], 10 classes after first stage are remained in MRR and RASR, and one project matrix of 500 rows is used in TSR. The illumination dictionary in [6] follows its original setup, and the amount of illumination atoms is 30 in all experiments. The maximum iteration of outer and inner loop for MRLR in these methods are consistently set to 3 and 30. The l_1-minimization algorithm uses the Augmented Lagrange Multiplier [16]. Matlab code is available at https://github.com/ydwen/MRLR/.

4.2. The region of attraction

The region of attraction evaluates the robustness against 2D deformations. We compare MRLR with TSR [7], RASR [5], MRR [8] and SIT [6] on Extended Yale B database, which includes 2414 images of 38 subjects. We use the uncropped images of 28 subjects in experiments. 32 training images per category are randomly selected, and the rest are used for testing. All the training images are resized to 80×70. We get access to the ground truth of eyes and add perturbation to them. Then we calculate the corresponding recognition accuracy under various initial transformations. One can see that MRLR performs well and stably within a certain range of misalignment,

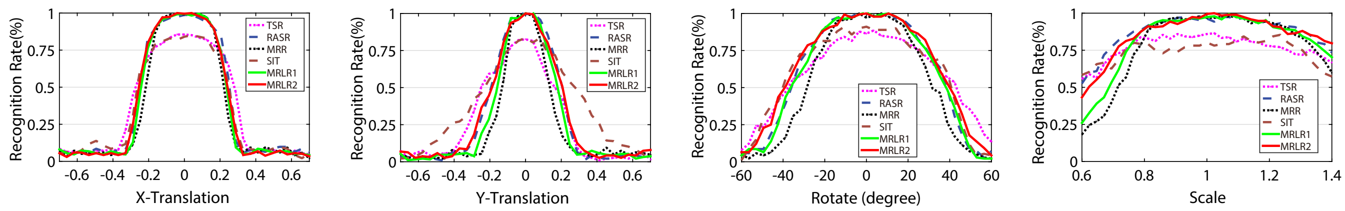

Fig. 2: The region of attraction. (The amount of translation is given as a fraction of the distance between eyes) (a) Translation in the Y-direction only. (b) Translation in the X-direction only. (c) In-plane rotation only. (d) Scale variation only.

e.g. 20 percent translation in x direction (14-16 pixels), 20 degree rotation or 30 percent scale variation. It significantly enhances the robustness of practical recognition, because the average misalignment of a face detector safely falls within 10 percent translation and 8 percent scale variation. MRLR1 and MRLR2 perform almost the same as RASR, demonstrating that locality-constrained representation effectively avoids local minima.

4.3. Recognition performance and running time

We conduct the recognition experiments on Extended Yale B, CAS-PEAL and LFW datasets. For Extended Yale B, we adopt the same settings in the previous section. For CAS-PEAL and LFW, 20 subjects were chosen, each of them including more than 32 images. We randomly selected 20 images per subject and resized them to 80×70 for training, then test on the remaining 12 images. The initial τ_0 are automatically given by Viola-Jones detector [17]. Table 1 gives the recognition rates and average running time.

Table 1: The recognition accuracy (%) and running time (s) on Extended Yale B, CAS-PEAL and LFW datasets.

Method	Extended Yale B		CAS-PEAL		LFW	
	Acc.	Time	Acc.	Time	Acc.	Time
TSR	81.61	7.396	86.96	4.2695	73.63	4.3477
RASR	92.42	9.7587	89.92	5.4466	81.43	5.9201
MRR	90.95	0.7773	90.00	0.5684	78.84	0.6339
SIT	84.53	9.9823	86.76	6.0329	55.91	6.3751
MRLR1	92.31	0.6207	89.76	0.3307	**82.98**	0.3403
MRLR2	**92.53**	**0.1783**	**90.43**	**0.1462**	81.75	**0.1840**

From table 1 we can observe that 1) MRLR achieve the competitive performance (92.53%, 90.43% and 82.98%) in three datasets, slightly better than RASR (92.42%, 89.92% and 81.43%). 2) MRLR takes only 0.18 and 0.15 seconds to deal with a testing image, roughly accelerate $4\times$, $55\times$, $41{\times}and59{\times}$ compared to MRR, RASR, TSR and SIT respectively. Because the dictionary for alignment consists of illumination dictionary (outside samples) and single training sample per class, it shares the same scale with RASR, resulting in similar running time. The experimental results demonstrate the effectiveness and efficiency of MRLR.

4.4. Scalability

We vary the number of subject from 10 to 100 and resize the images from 40×35 to 160×140, to evaluate the scalability of our algorithm on CAS-PEAL. Table 2 and Table 3 show the experimental results. One can observe that TSR, RASR and SIT 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. 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.

Table 2: Running time (s) under different dimensions (image size).

Method	40×35	64×56	80×70	120×105	160×140
TSR	3.645	3.861	4.270	4.672	5.468
RASR	3.499	4.452	6.110	10.324	17.111
MRR	0.133	0.342	0.593	2.259	5.997
SIT	3.564	4.637	6.565	11.035	19.215
MRLR1	0.085	0.195	0.331	0.569	0.940
MRLR2	**0.066**	**0.118**	**0.146**	**0.303**	**0.505**

Table 3: Running time (s) under different amount of classes.

Method	10	20	40	70	100
TSR	2.1533	3.2825	5.5280	8.4034	11.5327
RASR	2.7377	4.6596	8.8647	15.4644	22.1281
MRR	0.5776	0.5928	0.6082	0.6394	0.6994
SIT	2.86	5.1996	9.9817	17.6875	27.1734
MRLR1	0.1977	0.2819	0.513	0.8552	1.4096
MRLR2	**0.1318**	**0.1373**	**0.1559**	**0.197**	**0.2616**

5. CONCLUDING REMARKS

In this paper, we proposed an efficient misalignment-robust locality representation algorithm, namely MRLR, for face alignment. The locality constraint therein avoids the exhaustive search in every subject, greatly reducing running time while still preserving accurate alignment. Moreover, motivated by the block structure of dictionary, we proposed an efficient solving algorithm to speed up the alignment. Computational complexity analysis and extensive experiments showed that MRLR considerably reduced the running time with even better performance.

6. REFERENCES

[1] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma, "Robust face recognition via sparse representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 2, pp. 210–227, 2009. 1

[2] Zihan Zhou, Andrew Wagner, Hossein Mobahi, John Wright, and Yi Ma, "Face recognition with contiguous occlusion using markov random fields," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1050–1057. 1

[3] Meng Yang, Lei Zhang, Jian Yang, and David Zhang, "Robust sparse coding for face recognition," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 625–632. 1

[4] Lei Zhang, Meng Yang, and Xiangchu Feng, "Sparse representation or collaborative representation: Which helps face recognition?," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 471–478. 1

[5] Andrew Wagner, John Wright, Arvind Ganesh, Zihan Zhou, Hossein Mobahi, and Yi Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 2, pp. 372–386, 2012. 1, 2, 4.1, 4.2

[6] Liansheng Zhuang, Allen Y Yang, Zihan Zhou, S Shankar Sastry, and Yi Ma, "Single-sample face recognition with image corruption and misalignment via a sparse illumination transfer," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 3546–3553. 1, 4.1, 4.2

[7] Junzhou Huang, Xiaolei Huang, and Dimitris Metaxas, "Simultaneous image transformation and sparse representation recovery," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8. 1, 4.2

[8] Meng Yang, Lei Zhang, and David Zhang, "Efficient misalignment-robust representation for real-time face recognition," in *Computer Vision–ECCV 2012*, pp. 850–863. Springer, 2012. 1, 4.2

[9] Ronen Basri and David W Jacobs, "Lambertian reflectance and linear subspaces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 2, pp. 218–233, 2003. 2

[10] Yu-Wei Chao, Yi-Ren Yeh, Yu-Wen Chen, Yuh-Jye Lee, and Yu-Chiang Frank Wang, "Locality-constrained group sparse representation for robust face recognition," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 761–764. 3.1

[11] Xi Peng, Lei Zhang, Zhang Yi, and Kok Kiong Tan, "Learning locality-constrained collaborative representation for robust face recognition," *Pattern Recognition*, vol. 47, no. 9, pp. 2794–2806, 2014. 3.1

[12] Jinjun Wang, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas Huang, and Yihong Gong, "Locality-constrained linear coding for image classification," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 3360–3367. 3.2

[13] Athinodoros S Georghiades, Peter N Belhumeur, and David J Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 6, pp. 643–660, 2001. 4

[14] Wen Gao, Bo Cao, Shiguang Shan, Xilin Chen, Delong Zhou, Xiaohua Zhang, and Debin Zhao, "The cas-peal large-scale chinese face database and baseline evaluations," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 38, no. 1, pp. 149–161, 2008. 4

[15] Gary B Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Tech. Rep., Technical Report 07-49, University of Massachusetts, Amherst, 2007. 4

[16] Allen Y Yang, S Shankar Sastry, Arvind Ganesh, and Yi Ma, "Fast l1-minimization algorithms and an application in robust face recognition: A review," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1849–1852. 4.1

[17] Paul Viola and Michael Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. IEEE, 2001, vol. 1, pp. I–511. 4.3