

### Abstract

Low rank decomposition (LRD) is the state-of-the-art method for visual data reconstruction. However, it is challenged when the data contains significant occlusion, noise, illumination variation, and misalignment from rotation and/or viewpoint changing. We leverage the specific structure of data in order to improve the performance of LRD when the data are not ideal. To this end, we propose a new framework that allows to embed manifold priors into LRD. We show that the manifold constraints can be transferred from one variable to another. To implement the framework, a multipliers alternating direction method based on manifold is designed, which consistently integrates the manifold constraints during the optimization process. This is due to the assumption that we can recast the problem as the projection over the manifold via an embedding method. The proposed approach is successfully used to calculate low ranks from faces, digits, and window images, showing a consistent increase of performance when compared to the state of the art.

### Index Terms

Low rank, manifold, embedding, LRD, ADMM.

# Manifold Constraint for Rotation Invariant Low Rank Decomposition

Baochang Zhang\*, Linlin Yang\*, Jungong Han<sup>†</sup>, Vittorio Murino<sup>‡</sup>, Jianzhuang Liu<sup>¶</sup>, Linlin Shen<sup>§</sup>, Alessio Del Bue<sup>‡</sup>

\* School of Automation Science and Electrical Engineering, Beihang University, Beijing, China

<sup>†</sup> Department of Computer Science and Digital Technologies, Northumbria University, UK

<sup>‡</sup> Pattern Analysis and Computer Vision(PAVIS), Istituto Italiano di Tecnologia (IIT), Genova, Italy

<sup>§</sup> Computer School, Shenzhen University

<sup>¶</sup> Chinese University of Hong Kong, Hong Kong



# Manifold Constraint for Rotation Invariant Low Rank Decomposition

## 1 INTRODUCTION

With the increasing number of images and videos produced in our everyday life, it becomes more and more problematic when existing data analysis algorithms deal with wild data containing severe occlusion, rotation, noise, significant illumination variation, and viewpoint changing or misalignment [1], [2]. Therefore, a data analytical algorithm that is able to apply to high-volume real data is highly demanded.

Low-rank decomposition (LRD) technique has been an important tool in batch data analysis in the past decade, which effectively converts high dimensional raw data into a compact representation sample. This technique has been successfully used in a variety of applications such as subspace segmentation [5], [6], visual tracking, image clustering [7], [8] and video background/ foreground separation [9]. Usually, this technique works properly when the data is captured in an ideal situation or it is manually aligned. The performance of the algorithm, however, degrades significantly in case of rotation, corruption, occlusion and misalignment in the data. In such situations, the low-rank matrices cannot be accurately recovered from the data. Fortunately, researchers in this field are aware of this problem and reach a consensus that we have to solve it due to the fact that more and more wild data captured from our daily life need to be processed. Some attempts have been proposed. For instance, to treat the batch image alignment problem, the sophisticated measures of image similarity is used in [3], [27]. Alternatively, Learned-Miller's influential congealing algorithm [27] seeks an alignment that minimizes the sum of entropies of pixel values at each pixel location in the batch of aligned images. Instead of looking into the entropy, the least squares congealing procedure of [28] minimizes the sum of squared distances between pairs of images, and hence demands that the columns be nearly constant. In [29], Vedaldi *et al.* choose to minimize a log-determinant measure that can be viewed as a smooth surrogate for the rank function. The Robust Principal Component Analysis (RPCA) algorithm of [33] also fits a low-rank model, and uses a robust fitting function to reduce the influence of corruption and occlusion.

A nuclear-norm minimization method, called RASL [24], is presented based on the alternating direction method of multipliers (ADMM), which has shown the potential to solve realistic misalignments and corruptions. The core idea is to seek an optimal set of transformations such that the matrix of transformed images can be decomposed as the summation of a low-rank matrix of recovered aligned images and a sparse matrix of errors. The algorithm is subject to a set of linear equality constraints [24], which only consider the linear relationship among the input data. However, the fact that input data is generally of nonlinear structure, i.e., manifold, is not fully investigated in the optimization process.

In this paper, we provide new insights into the nuclear-norm minimization method, in particular a relevant intuition that was neglected in previous work. That is, data often lies on specific manifolds [20], [37], especially when the data comes from a well-

defined object from a given set of samples (e.g. faces, digits, etc.). The methods taking advantage of the implicit structure of the data can generally harvest the better performance [31]. From the optimization perspective, assuming that the solution of the optimization problem is always data related, the constraints derived from the data structure can make the algorithm immune to the variations existed in the testing data [10], [17], [31]. Consequently, it is important to incorporate the structure prior of data in the learning procedure. To implement it, it is required to embed manifold constraints during the optimization procedure. Here we show that there exists a solution with high practicability that can include manifold constraints in ADMM. In LRD, we provide a simple way to calculate a projective matrix to constrain the optimized variable. To sum up, the contributions of this paper are twofold:

(i) We proposed to incorporate the manifold constraints into LRD method, as the manifold constraints are in accordance with the nonlinear structure, which fits better to the realistic data. Our approach achieves much better low rank results than the prior art.

(ii) We present a manifold based MADMM framework, where manifold constraint is done based on a projection matrix calculated by a neighbor-preserving embedding process, which simplifies the optimization procedure when learning models.

For ease of explanation, we summarize all the variables in Table 1, which are briefly described here.  $V_r$  is the low rank result calculated from the input sample  $V_d$ . The geometric transformation  $\tau$  and  $\Delta\tau$  are used to calculate  $V_m$  from  $V_r$ . And  $V'_m$  is a manifold embedding of  $V_m$ .

The rest of the paper is organized as follows. We introduce the related work in Section 2. Section 3 details how manifold constraints can be efficiently embedded in a ADMM optimization framework. The experiments over several datasets are discussed in Section 4. Finally, conclusions are drawn in Section 5.

## 2 RELATED WORK

The foundations of our work are about the RASL and manifold methods. Therefore, in this section, the literature overview focuses on RASL methodologies, as well as relevant manifold approaches.

### 2.1 RASL

The misalignment problem is one of the most difficult problems in the field of computer vision. By formulating the batch image alignment as the search for a set of transformations that minimize the rank of the transformed images, RASL investigates the linearly correlated relationship among the input images, which is designated as Problem 1:

$$\begin{aligned} \hat{V}_r, \hat{\tau} &= \arg \min \quad \text{rank}(V_r) \\ &\text{subject to} \\ V_d \circ \tau &= V_r \\ \tau &\in \mathcal{G} \end{aligned} \tag{P1}$$

TABLE 1  
A brief description of variables used in the paper

variable	description	variable	description
$V_r$	low rank	$V_d$	input samples
$\tau$	geometric transformation	$\Delta\tau$	to calculate new $\tau$
$V_m$	calculated by $\Delta\tau$ and $V_d$	$V'_m$	Embedding of $V_m$
$E$	the error matrix	$S_\alpha[x]$	soft-thresholding function

Given a B-frames video sequence  $V_d$  with the resolution of  $M \times N$ , we intend to extract a compact video representation  $V_r$  that contains the low rank elements from  $V_d$ . In particular,  $V_d$  can be considered as a collection of an image instance and its variations including rotation, illumination variation, occlusion, and geometric transformations  $\tau \in \mathcal{G}$ .  $\mathcal{G}$  is defined as a  $3 \times 3$  matrix [24]. To efficiently solve the problem, we describe a linearisation process such that:

$$V_d \circ (\tau + \Delta\tau) = V_d \circ \tau + \sum_i^B J_i \Delta\tau_i \epsilon_i \quad (1)$$

$(\tau + \Delta\tau)$  generates the new  $\tau$  in the iteration process.  $\Delta\tau$  is calculated as suggested in [24].  $J_i$  is the Jacobian of the  $i^{th}$  image with respect to the transformation parameters, and  $\epsilon_i$  denotes the standard basis for  $\Re_n$ . The above linearisation process only holds locally, and we repeatedly linearise current estimates by solving a sequence of convex programming. After the linearisation, the problem solved for our framework is semidefinite programming in thousands or millions of variables, so scalable solutions are essential for its practical use. Fortunately, a recent flurry of works on high-dimensional nuclear norm minimization has shown that such problems are well within the capabilities of a standard PC [24].

## 2.2 Manifolds

Manifolds are popular topic in machine learning, because they allow to describe more complicated structures in Euclidean space. Most of existing works related to manifolds focus on modeling data. To compact high-dimensional data, manifold learning [38] projects the original data onto two or three dimensions such that its inherent structure can be preserved. As another application of manifold learning, an embedding of a sample can be obtained by projecting onto a well-designed sub-manifold [20]. A sub-manifold is a subset of a manifold that is itself a manifold. To exploit the geometry of the marginal distribution, a semi-supervised framework based on manifold regularization is used to learn from both labeled and unlabeled data [35] in the form of a multiple kernel learning. In [36], by representing the covariance matrix as a point on a manifold a new metric is learned for that manifold. Differently, in [14], we deploy the manifold constraints in an ALM strategy by using a matrix projection as a constraint of the optimized variable, which efficiently computes the solution over several given manifolds (e.g. Stiefel, unit sphere). The work is remarkable and leads to a new framework using manifolds to solve the optimization problem. However, it fails to explain why the variable should stop on a manifold.

Unlike the existed works, we present here a new method that exploits the manifold constraints in the ADMM framework. Instead of empirically adding manifold constraints on a variable, we introduce a manifold based ADMM approach to intuitively regulate the optimization problem, which is based on our previous

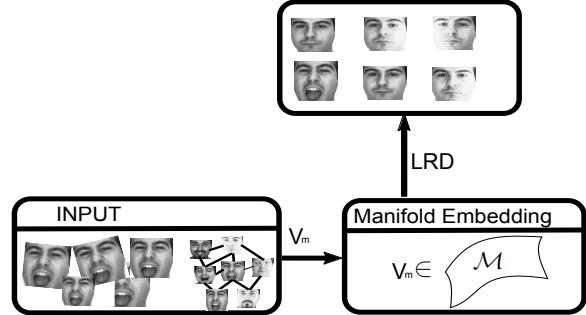


Fig. 1. The framework of MADMM

work [31]. Rather than concentrating on sparse representation only, this work provides a more comprehensive study into manifold for the LRD problem. Additionally, our idea is also different from that in [2], in which a non-negative sparse hyper-Laplacian regularized low-rank representation model for image clustering and classification is proposed.

## 3 LOW-RANK DECOMPOSITION BASED ON MANIFOLD CONSTRAINTS

A constrained learning model allows for incorporating domain-specific knowledge to balance the learned model based on the implicit structure of the data [11]. From machine learning perspective, it is of high significance to simplify the learning stage while improving the accuracy of the solution. In this section, we present how manifold constraints can be embedded into an optimization. We show a proposal to formulate the LRD optimization problem in terms of manifold based ADMM, thus resulting in a relaxed and more efficient solution to the new problem defined in P2. Our idea is intuitively illustrated in Fig. 1, where the input images are first embedded into a manifold and the low rank results are obtained afterwards by LRD.

### 3.1 LRD reformulation based on manifold constraint

To efficiently calculate low rank from data with a nonlinear structure, manifold based ADMM (MADMM) deploys the manifold structure in the optimization process. We first introduce a new variable  $V_m$  such that:

$$V_m = V_d \circ (\tau + \Delta\tau) \quad (2)$$

$V_m$  is introduced to replace  $V_d \circ \tau$ , which is a new variable and linearly correlated to  $V_r$  in the new problem. The MADMM based

LRD method is formulated as:

$$\begin{aligned} \hat{V}_r, \hat{E}, \hat{\tau} &= \arg \min \quad \text{rank}(V_r) + \lambda * \|E\|_1 \\ &\text{subject to} \\ V_m &= V_r + E, \\ V_{r,i} &\in \mathcal{M}, \\ \tau &\in \mathcal{G}, \end{aligned} \tag{P2}$$

Normally, only a small fraction of pixels will be affected by partial occlusions or corruptions, thus  $E$  is considered to be a sparse one. Supposed that the input data is generally of nonlinear structure,  $V_r$  is reasonably considered to be from a manifold  $\mathcal{M}$ . That is,  $V_{r,i} \in \mathcal{M}$ . This is hard to be solved, however, in case  $\mathcal{M}$  is undefined.

We propose to solve the problem in three steps. We first exploit the ALM framework in this subsection to solve the problem without taking manifold constraints into account. In the second step, we introduce the manifold constraint into the objective function in the subsection 3.2. Finally, we solve all variables in the algorithm 1, 2, and Equ. ?? in the subsection 3.4.

The basic idea of the ALM method is to search for a saddle point of the augmented Lagrangian function instead of directly solving the original constrained optimization problem. For the original problem without considering the manifold constraint, we define  $f(V_r, E, \Delta\tau) = f(V_r, V_m, E, \Delta\tau) = (V_r + E) - (V_d \circ \tau + \sum J_i \Delta\tau_i \epsilon_i) = (V_r + E) - V_m$ . Then we have:

$$\begin{aligned} L_\mu(V_r, E, \Delta\tau, Y) &= \|V_r\|_* + \lambda * \|E\|_1 - \\ &\langle Y, f(V_r, E, \Delta\tau) \rangle + \frac{\mu}{2} \|f(V_r, E, \Delta\tau)\|^2, \end{aligned} \tag{3}$$

where  $Y \in \Re^{M \times N}$  is a Lagrange multiplier matrix, and  $\mu$  is a positive scalar  $\langle \cdot, \cdot \rangle$  denotes the matrix inner product. For an appropriate choice of the Lagrange multiplier matrix  $Y$  and sufficiently large constant  $\mu$ , it can be shown that ALM has the same minimizer as the original constrained optimization problem.

## 3.2 MADMM

The details of manifold based low rank method are described in the this subsections. In particular, we solve a new Problem (P2) whose data lies over a manifold. More specifically, we propose to consider  $V_r$  as an unknown variable of the optimization by performing variable cloning i.e.,  $V_r \rightarrow V'_r \in \mathcal{M}$  and to explicitly enforce manifold constraints over the cloned variables  $V'_r$ . This aims to introduce explicitly the manifold constraints at the expenses of replicating a set of variables. To solve the problem mentioned above, we incorporate a new variable  $V'_r$  defined to be from a manifold.  $V'_{r,i} = V_{r,i}$  is alternatively used to add the manifold constraint to replace  $V_{r,i} \in \mathcal{M}$ . Now, the problem (P2) can be rewritten as:

$$\begin{aligned} \hat{V}_r, \hat{E}, \hat{\tau} &= \arg \min \quad L_\mu(V_r, E, \Delta\tau, Y) \\ &\text{subject to} \\ V'_{r,i} &= V_{r,i} \\ V'_{r,i} &\in \mathcal{M} \end{aligned} \tag{P3}$$

Based on [31], we get the following objective as:

$$\begin{aligned} L_{\mu,1}(V_r, E, \Delta\tau, Y) &= \|V_r\|_* + \lambda * \|E\|_1 - \\ &\langle Y, f(V_r, E, \Delta\tau) \rangle + \frac{\mu}{2} \|f(V_r, E, \Delta\tau)\|^2 \\ &+ \sum_i^B \frac{\sigma_i}{2} \|V_{r,i} - V'_{r,i}\|^2 \end{aligned} \tag{4}$$

where  $\sigma_i$  is positive. However, the above objective is still too complicated to be solved in a real application, as  $L_{\mu,1}$  is the combination of  $L_\mu$  and another function related to  $V_r$  as shown in Equ. 5. In this case, the calculation of  $V_r$  is more complicated than that of RASL, which is based on  $L_\mu$ . Assume the linear constraint on  $V_m$  and  $V_r$ , i.e.,  $V_m = V_r + E$ , we neglect the error matrix  $E$ , and get the following objective:

$$\begin{aligned} L_{\mu,2}(V_r, V_m, E, \Delta\tau, Y) &= \|V_r\|_* + \lambda * \|E\|_1 - \\ &\langle Y, f(V_r, V_m, E, \Delta\tau) \rangle + \frac{\mu}{2} \|f(V_r, V_m, E, \Delta\tau)\|^2 \\ &+ \sum_i^B \frac{\sigma_i}{2} \|V_{m,i} - V'_{m,i}\|^2 \end{aligned} \tag{5}$$

Again, we have  $V'_{m,i} = V_{m,i}$  and  $V'_{m,i} \in \mathcal{M}$ . The above expression requires a minimization over  $V'_{m,i} \in \mathcal{M}$  with  $i = 1, \dots, F$ . In [14], the manifold constraints are enforced in an ALM strategy by using a matrix projection which efficiently computes the solution over several given manifold (e.g. Stiefel, unit sphere). Differently, in our case, data is embedded into a manifold without knowing a priori (i.e.  $V_r$ ). Therefore, it is customized to the specific low rank calculation problems at hand.

Now we formalize the manifold by introducing a neighbor-preserving embedding [39], [13], in which the intention is to find an estimation on a sub-manifold (a subset of a manifold). Such a formalization is similar to [13] that first calculates the weights in the process of dimension reduction by LLE [20]. In particular, the embedding generated by [13], [12] is exactly based on a sub-manifold given by a small set of samples. Therefore, instead of finding a whole manifold (LLE) of the data, we alternatively use or construct a sub-manifold. To do so, we try to find a projection in a sub-manifold based on its “true” neighbors measured by the Geodesic distance information [12], thus avoiding the perturbation of samples that are far from the input data.

## 3.3 Neighbor-preserving embedding

Let  $\mathcal{M}$  be the sample set representing a manifold and  $x$  be the embedding of  $\mathcal{M}$  via a mapping function  $\Phi(\cdot)$ .

**Definition 1.** The map function  $\Phi : x \rightarrow \mathcal{M}$  in the neighbour-preserving embedding method is conducted based on the Geodesic distance, which is defined as follows:

- 1) First we define  $\mathcal{M}_1 = \sum_{j=1}^K (1 - W_j) \mathcal{M}_j$  where  $W_j$  is the Geodesic distance of the sample  $x$  and the  $j^{th}$  sample in a sub-manifold set  $\mathcal{M}$ .
- 2) We define  $\mathcal{E} = \mathcal{M} - \mathcal{M}_1$ , and have:  
 $\Phi_{\alpha,\epsilon}(x, \mathcal{M}) = x_{\alpha,\epsilon} = \mathcal{M}_1 + \epsilon \cdot S_\alpha[\mathcal{E}]$ ;  $S_\alpha[x] = \text{sign}(x) \cdot \max\{|x| - \alpha, 0\}$   
where  $\alpha$  and  $\epsilon'$  are used to represent the shrinkage factor and the scaler for reconstruction error respectively.
- 3) As shown in [12], for a given point projected onto the sub-manifold, the bigger weights are reasonably assigned to its nearest points in the recovery process.

From Def. 1, the input sample can be projected onto a well-designed sub-manifold via an embedding function by fully exploiting the neighbour structure information [13]. As shown in LLE [20], [37], a local point on a sub-manifold can be represented by a small and compact set of nearest neighbors (i.e.  $K$ ) to approximate ISOMAP. Later in [39], it has been shown that the Geodesic distance used in ISOMAP is another effective way to locate the neighbors for a linear embedding. We first follow the idea in [25] to estimate the sub-manifold dimension by PCA. Next, we propose the map function  $\Phi$  to generate an embedding sample that lies on a given sub-manifold. Our idea is similar to [13] but the difference lies in its simplicity and feasibility to solve the problems at hand. Afterwards, the manifold structure is, for the first time, exploited to find a reasonable constraint for several problems jointly. This is the only part in the algorithm that the constraint manifold  $V'_m$  plays a role and replacing  $V'_m$  amounts to compute the proposed manifold embedding. Finally, MADMM adds an extra computational cost to our problem because variables are added given the cloning mechanism.

The soft-thresholding or shrinkage operator for scalars [24] is defined as:

$$S_\alpha[x] = \text{sign}(x) \cdot \max(|x| - \alpha, 0),$$

where  $\alpha \geq 0$ . When applied to vectors and matrices, the shrinkage operator acts element-wise. Based on the Definition 1, MADMM can be alternatively used to solve our problem and in the following we give details about the formalization of the optimization procedure.

---

**Algorithm 1:** Main Algorithm for misalignment robust in Low Rank images

---

- 1: INPUT:

  - 1)  $V_d \circ \tau = [\text{vec}(I_1), \text{vec}(I_2), \dots, \text{vec}(I_B)]$ ,  $I_i | i = 1, \dots, B$  represent B input images of the same object
  - 2)  $(V_d^0, E^0, \Delta\tau^0)$

- 2: repeat

  - 3: computing Jacobian matrices w.r.t transformations:  
 $J_i \leftarrow \frac{\partial}{\partial \zeta} \left( \frac{\text{vec}(I_i \circ \zeta)}{\|\text{vec}(I_i \circ \zeta)\|} \right) \Big|_{\zeta=\tau};$
  - 4: warp and normalize the images:  
 $V_d \circ \tau = \left[ \frac{\text{vec}(I_1 \circ \zeta)}{\|\text{vec}(I_1 \circ \zeta)\|}, \frac{\text{vec}(I_2 \circ \zeta)}{\|\text{vec}(I_2 \circ \zeta)\|}, \dots, \frac{\text{vec}(I_B \circ \zeta)}{\|\text{vec}(I_B \circ \zeta)\|} \right]$
  - 5: solving the manifold constraint on the transformation process on  $V_d \circ \tau + \sum J_i \Delta\tau_i \epsilon_i$   
the details of  $V'_m$  and  $V_r$  are shown in the Alg. 2 and Equ. 8. (inner loop)
  - 6: updating the transformation:  $\tau = \tau + \Delta\tau$
  - 7: until some stopping criterion
  - 8: OUTPUT: the solution  $(V_r^*, E^*, \Delta\tau^*)$  in our optimization framework.

---

### 3.4 The MADMM algorithm

The main algorithm of MADMM is shown in Alg. 1. In the outer loop, we solve the  $\tau$ , while other variables such as  $V_r$  and  $V_m$  are solved in the inner loop. A separable structure based on  $L_{\mu,2}(\cdot)$  can be exploited by ADMM, which is:

$$(V_r^{[k+1]}, E^{[k+1]}, \Delta\tau^{[k+1]}, Y^{[k+1]}) = \arg \min_{V_r, E, \Delta\tau} L_{\mu,2}(V_m^{[k]}, E^{[k]}, \Delta\tau^{[k]}, Y^{[k]}). \quad (6)$$

Details concerning the solution of Equ. 6 are shown in Alg. 2. Different from original objective function in [24],  $V_m^{[k]}$  needs to be estimated first to do the SVD decomposition. Considering the constraint  $V_m^{[k]} = V_m'^{[k]}$ ,  $V_m'^{[k]}$  can be used to replace original  $V_m^{[k]}$  as shown in Alg. 2. In this case,  $V_m'^{[k+1]}$  is actually used to approximate  $V_m^{[k+1]}$  that stops on a manifold. Now we have a new objective as:

$$\begin{aligned} L_{\mu,2}(V_r, V_m, E, \Delta\tau, Y) &= \|V_r\|_* + \lambda * \|E\|_1 - \\ &< Y, ((V_r + E) - V_m') > + \frac{\mu}{2} \| (V_r + E) - V_m' \|^2 \\ &+ \sum_i^B \frac{\sigma_i}{2} \|V_{m,i} - V'_{m,i}\|^2 \end{aligned} \quad (7)$$

From Equ.7, calculating  $Y^{[k+1]}$ ,  $E^{[k+1]}$  and  $\Delta^{[k+1]}$  are not related to  $\sum_i^B \frac{\sigma_i}{2} \|V_{m,i} - V'_{m,i}\|^2$ , and so a similar method as that of [24] is used to solve our problem as that in Alg. 2. Now only  $V_m'^{[k+1]}$  is unsolved. We have  $V_m'^{[k+1]} = V_r^{[k+1]} + E^{[k+1]}$ , based on the derivative of Equ. 7,  $V_m'^{[k+1]}$  is solved as:

$$V_m'^{[k+1]} = \mathcal{T}_m \cdot V_m^{[k+1]} \quad (8)$$

where  $\mathcal{T}_m = (Y + (\sigma^* + \mu) \cdot I)^{-1} \cdot (\sigma^* + 2\mu)$ ,  $\sigma_i$  is the diagonal element of  $\sigma^*$ , and  $I$  is the identity matrix. From Equ. 8,  $\mathcal{T}_m$  is a unknown projection matrix on  $V_m^{[k+1]}$ . Based on manifold embedding, the projection is solved in an efficient way as that in Def. 1. More specifically,  $V_{m,i}^{[k+1]} = \Phi_{\alpha,\epsilon}(V_{m,i}^{[k+1]}, V_m^{[k+1]})$ .

The embedding performs well for a small set of nearest samples. It computes a set of weights for each point, eventually describing the point as a linear combination of its neighbors. The proposed manifold embedding tends to handle non-uniform sample densities inappropriately. The reason is that there is no fixed unit to prevent the weights from drifting due to the fact that various regions differ in sample densities. It is involved in the recovery process, leading to the robustness against severe illumination and corruption. Based on it, the current frame can be mostly recovered by using the neighbors information.

---

**Algorithm 2:** Variable solution based on the ADMM algorithm

---

- 1: INPUT:  $V_m'^{[k]}$  calculated in Def. 1.
  - 2: computing  
 $(U, \Sigma, V) = \mathcal{SVD}(V_m'^{[k]} + Y^{[k]}/\mu^{[k]} - E^{[k]})$
  - 3: computing  $V_r^{[k+1]} = US \frac{1}{\mu^{[k]}} |\Sigma| V^T$
  - 4: computing  
 $E^{[k+1]} = S \frac{1}{\mu^{[k]}} [V_d \circ \tau^{[k]} + \sum J_i \Delta\tau_i^{[k]} \epsilon_i \epsilon_i^T + Y^{[k]}/\mu^{[k]} - V_r^{[k+1]}]$
  - 5: computing  
 $\Delta\tau^{[k+1]} = \sum_i J_i (V_r^{[k+1]} + E^{[k+1]} - V_d \circ \tau^{[k]} - 1/(\mu^{[k]}) Y^{[k]}) \epsilon_i \epsilon_i^T$
  - 6:  $Y^{[k+1]} = Y^{[k]} + \mu^{[k]} L_u(V_r^{[k+1]}, E^{[k+1]}, \Delta\tau^{[k+1]}, Y^{[k]})$ ,
  - 7:  $\mu^{[k+1]} = \max(0.9\mu^{[k]}, \tilde{\mu})$ .
  - 8: computing  $V_m'^{[k+1]}$  based on Def. 1.
  - 9: OUTPUT: the solution  $(V_r^*, E^*, \Delta\tau^*, Y^*)$  to the recovery process in our optimization framework.
-

TABLE 2

Statistics of errors in eye centers, calculated as the distances from the estimated eye centers to their center.

methods	mean errors	Error Std.	Max Error
Initial	1.69	0.428	2.23
RASL	0.16	0.36	1.0
MADMM	0.14	0.35	1.0

## 4 EXPERIMENTS

We test our low rank decomposition method on various datasets including Extended Yale face database B [4], AR [19], USPS [16], and Windows. The images in those databases suffer from rotation, occlusion and lighting variations. Here, we add the manifold constraints to find the low rank results from the input images. The Geodesic distance for sub-manifold is calculated based on  $V_r$ , which can be easily implemented in our following experiments. The number of the neighbors ( $K$ ) in the Geodesic distance calculation is empirically set to 7. We empirically set  $\alpha = 0.05$ ,  $\epsilon' = 0.85$  in all our experiments.

To validate the performance of MADMM, we use RASL [24] as the baseline, which is the state-of-the-art algorithm in LRD. It should be noted that MADMM is implemented based on the source code of RASL. To fairly compare with RASL, the parameter settings of our method are the same as those of RASL.

### 4.1 MADMM for faces

**MADMM for Extended Yale Face Database B:** In the Extended Yale Face Database B, each subject contains 64 illumination conditions and is resized to 42x45. The images in the database were captured using a purpose-built illumination rig. This rig is fitted with 64 computer controlled strobes. The 64 images of a subject in a particular pose were acquired at camera frame rate (30 frames/second) in about 2 seconds, so there is only small change in head pose and facial expression for those 64 (+1 ambient) images. To increase the difficulty for the LRD problem, we randomly rotate and shift the face images, and some results (INPUT) are illustrated in Fig.2.

With respect to the alignment, both methods achieve more or less the same performance. For the average faces after alignment, both of methods achieve clear low rank results, which indicates that the misalignment problem is well solved by RASL and MADMM. As evident in Table 2, they achieve small misalignment errors for the faces suffered from rotation, lighting variations and shifting.

Differently from [24] focusing on the misalignment performance, we pay much attention on the recovery effectiveness. As can be seen, MADMM achieves much better performances than those of RASL, especially when viewing the results on the last three rows where we significantly eliminate the illuminations from the original images. Considering the severe illumination and misalignment contained in the dataset, it is convincing that the results sufficiently validate the effectiveness of MADMM. More results on this database are also shown in the appendix part.

**MADMM for AR Face Database:** We next test MADMM on the AR database, which contains 126 persons with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. Each person

participated in two sessions, separated by two weeks (14 days) time. For each person, we pick up 26 pictures with 64x64 pixels from the Session 1 to calculate their low rank results. Different from Yale database B, the faces are severely occluded by glasses, and scarf. It can be observed from Fig.5 that MADMM still achieves much better low rank images than RASL, especially for the subjects of big scarf and expression variations. And the eyes and mouths are almost recovered from the input images as shown in the last row. This is remarkable, since RASL does not have the property as that of MADMM. MADMM is believed to be used for extensive applications, i.e., recognition [40], [41], [42].

### 4.2 MADMM for digits

In the USPS handwritten digit database, we use a popular subset containing 10-class digit images, and perform MADMM and RASL on its training set. As shown in Fig.3, the samples in the first column are the input data, and the third column and the last column report the low rank results calculated by MADMM and RASL, respectively. More specifically, MADMM recovers all the images of digit 1, but RASL fails on three images. Similarly, MADMM works well on digit 4, whereas RASL fails to recover one low rank image. The experiment on the digit shows again the advantages of our method when the variations such as rotation, shift, and affine transformation occur in the data. More results on the database are shown in the appendix part.

### 4.3 MADMM for Window Images

Moreover, we test our method on Window images. To better demonstrate the robustness of the proposed MADMM method, we manipulate some variations on input images by cropping patches or changing illuminations as shown in Fig.4. We again observe from Fig. 4 that MADMM achieves much better results on the first and third images. Together with consistent results obtained from faces and digits datasets, we could draw a conclusion that MADAMM is indeed effective for low rank calculation if the data include rotation, occlusion, and illumination variations. More results are included in the appendix part.

**Speed of MADMM:** Regarding the computational cost, MADMM is efficient, but it is not as fast as RASL. For example, the running time for window images are respectively 280ms and 87ms for MADMM and RASL, on a PC with Intel i5 CPU and 4G RAM. The testing package on this dataset will be available at our website ([mpl.buaa.edu.cn](http://mpl.buaa.edu.cn)) or the attached files.

## 5 DISCUSSIONS AND FUTURE WORK

This paper focuses on a new insight into the ADMM based LRD method from the perspective of manifold constraint. Based on a simple neighbor-preserving embedding method, an ADMM scheme is proposed to solve our problem. Finally, we solve manifold constraint based on a projective matrix, which is efficiently calculated in an embedding process, resulting in an efficient solution to our problem. The proposed approach is successfully applied for face, digit, window objects, showing a consistently increase on performance compared to the state of the art.

## ACKNOWLEDGEMENTS

B. Zhang and L. Yang acknowledge the support of the Natural Science Foundation of China, under Contracts 61272052 and

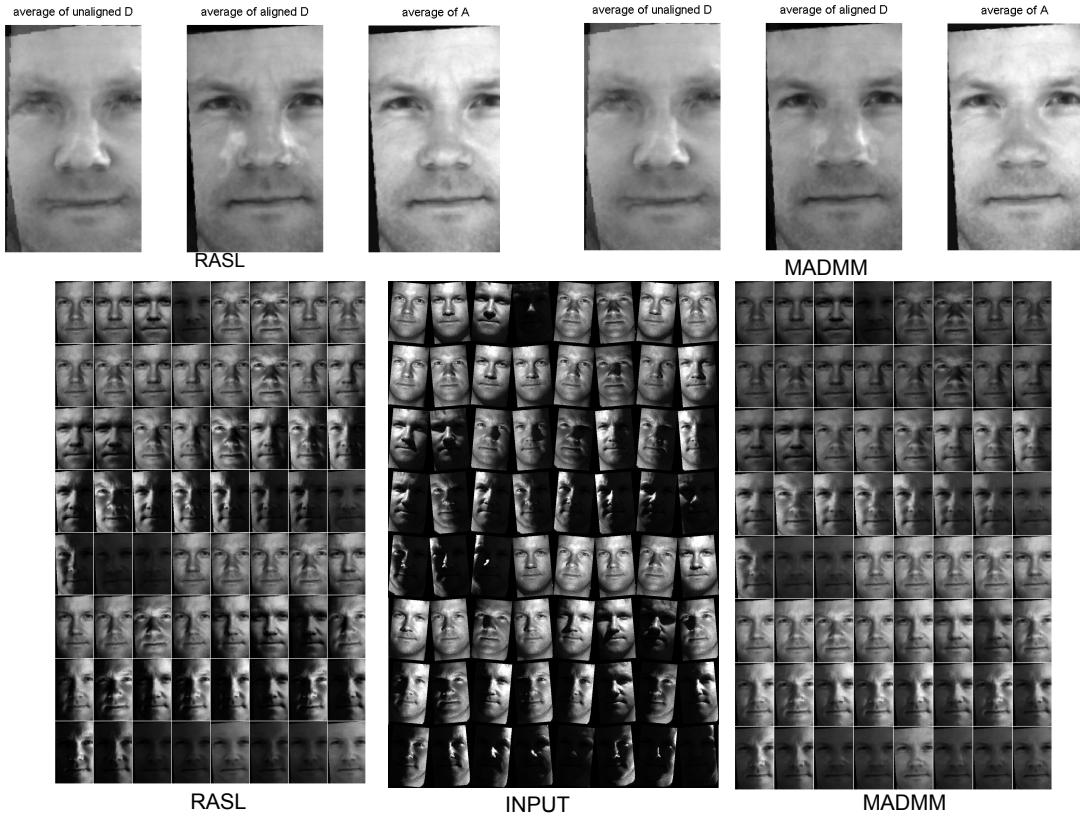


Fig. 2. Comparative low rank calculation by RASL and MADMM on the Yale-B database. In the first row, the average of input, alignment, and low rank results are shown for RASL and MADMM. In the second row, the first column is obtained by RASL, the second column is the input samples, and the last column is obtained by MADMM

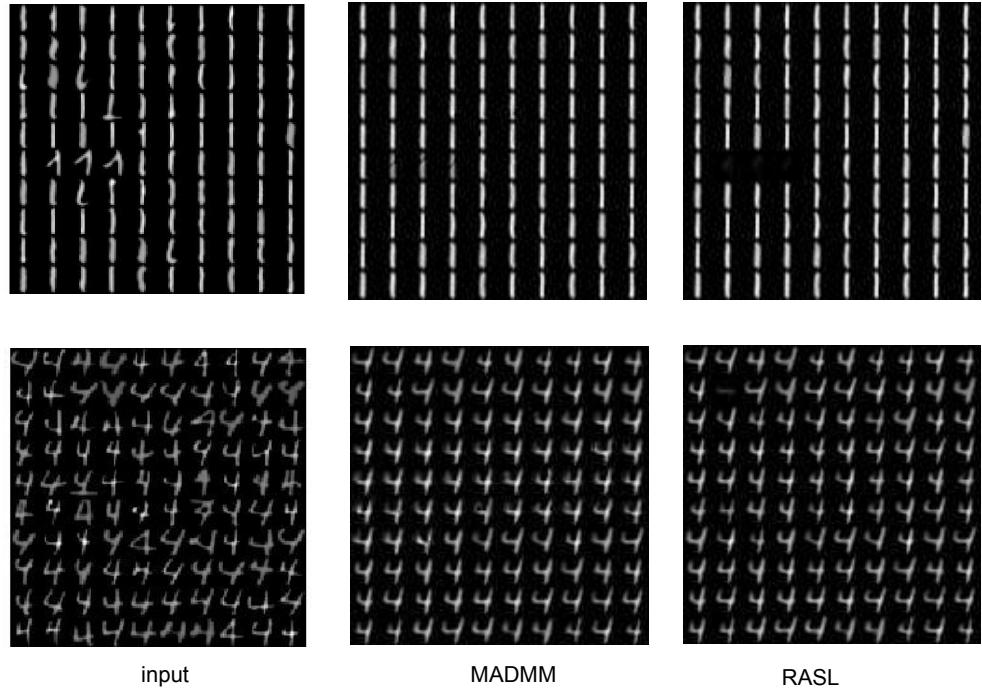


Fig. 3. Comparative low rank calculation by RASL and MADMM on the USPS database

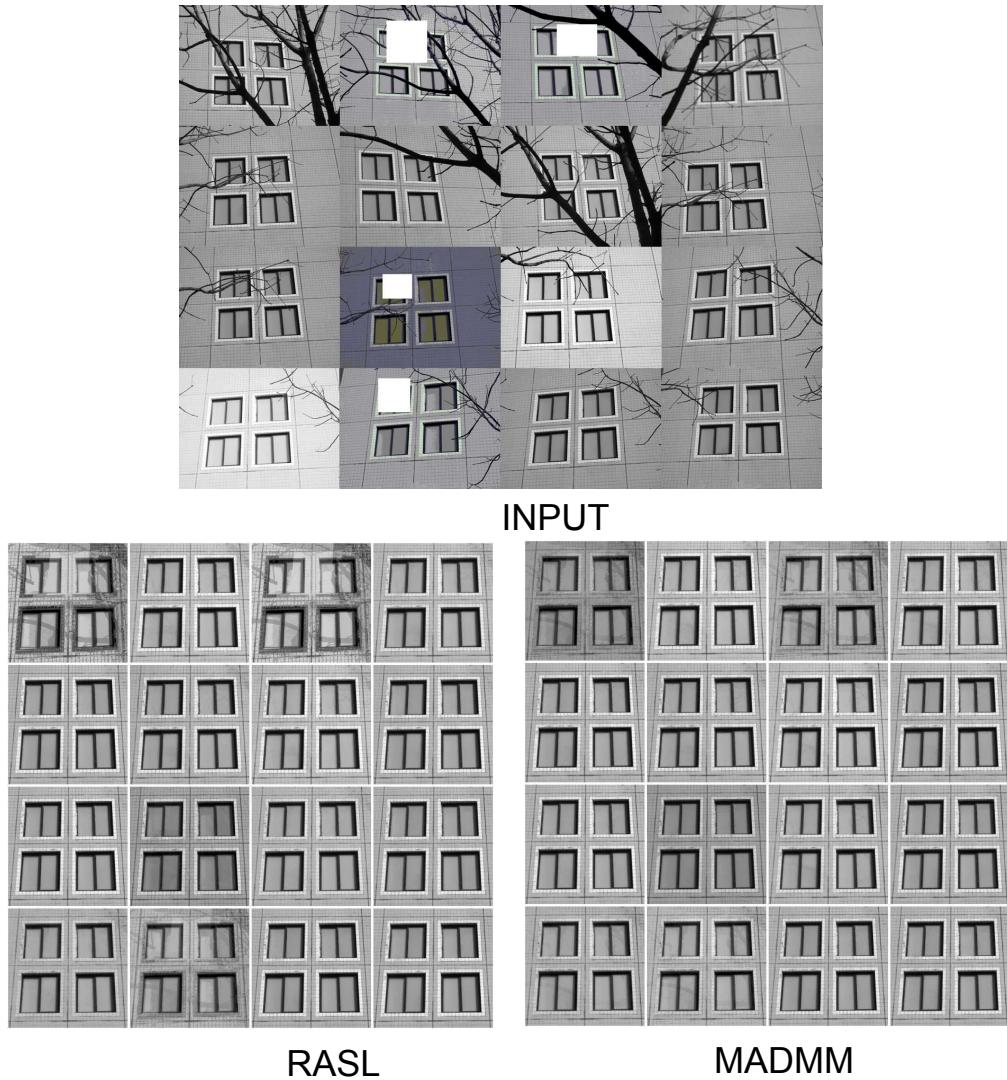


Fig. 4. Comparative low rank calculation of RASL and MADMM

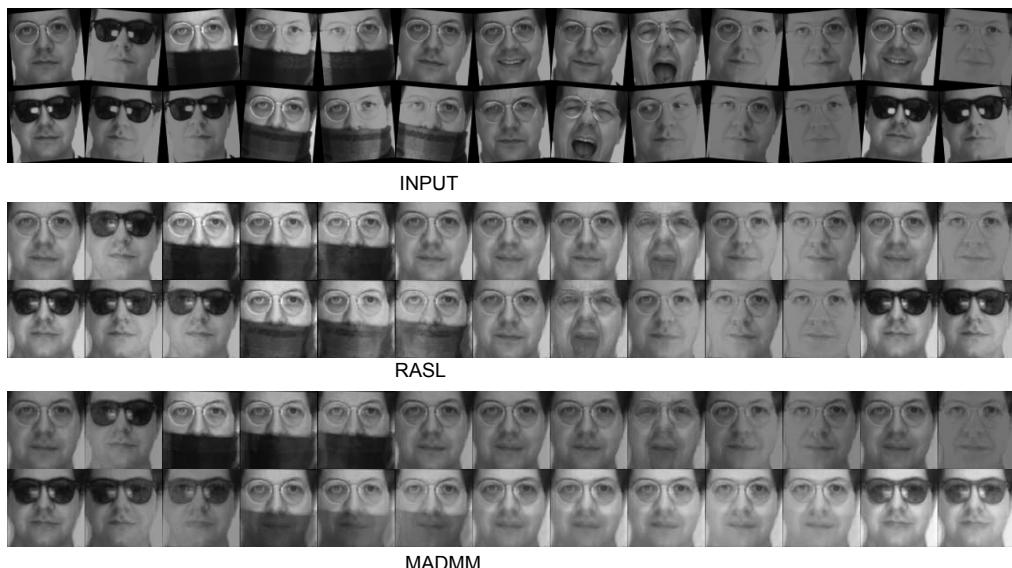


Fig. 5. Comparative low rank decomposition by RASL and MADMM on the AR database. The top row shows the input images. The second and third rows show the results by RASL and MADMM, respectively.

61473086, and of the Program for New Century Excellent Talents of the University of Ministry of Education of China. Thanks Qixiang Ye and Ruiping Wang from CAS for the suggestions to improve the quality of manuscript. Thank for Rongrong Ji from Xiamen university for discussion on the work.

## REFERENCES

- [1] Tianzhu Zhang, Si Liu, Narendra Ahuja, Ming-Hsuan Yang, Bernard Ghanem, Robust Visual Tracking Via Consistent Low-Rank Sparse Learning. *International Journal of Computer Vision* 111(2): 171-190 (2015)
- [2] Ming Yin, Junbin Gao, Zhouchen Lin, Qinfeng Shi, and Yi Guo, Dual Graph Regularized Latent Low-rank Representation for Subspace Clustering, *IEEE Trans. Image Processing*, Vol. 24. No. 12, pp. 4918-4933, 2015.
- [3] Ming Yin, Junbin Gao, Zhouchen Lin, Laplacian Regularized Low-Rank Representation and Its Applications, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2016.
- [4] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):643-660, 2001.
- [5] Chih-Fan Chen, Chia-Po Wei, and Yu-Chiang Frank Wang. Low-rank matrix recovery with structural incoherence for robust face recognition. In *CVPR*, pages 2618–2625, 2012.
- [6] Ming Yin, Shuting Cai, and Junbin Gao. Robust face recognition via double low-rank matrix recovery for feature extraction. In *IEEE International Conference on Image Processing (ICIP)*, pages 3770–3774, 2013.
- [7] Zhenyue Zhang and Keke Zhao. Low-rank matrix approximation with manifold regularization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(7):1717–1729, 2013.
- [8] Liansheng Zhuang, Haoyuan Gao, Zhouchen Lin, Yi Ma, Xin Zhang, and Nenghai Yu. Non-negative low rank and sparse graph for semisupervised learning. In *CVPR*, pages 2328–2335. IEEE, 2012.
- [9] D. Babacan, M. Luessi, R. Molina, and A.K. Katsaggelos. Sparse bayesian methods for low-rank matrix estimation. *IEEE Transactions on Signal Processing*, 60(8):3964–3977, August 2012.
- [10] G. Cabanes and Y. Bennani. Learning topological constraints in self-organizing map, in proc. of ICONIP, 2010.
- [11] M. Chang, L. Ratnay, and D. Roth. Guiding semi-supervision with constraint-driven learning. *ACL*, pages 280–287, 6 2007.
- [12] J. Chen, X. Chen, J. Yang, S. Shan, R. Wang, and W. Gao. Optimization of a training set for more robust face detection,. *Pattern Recognition*, 42(11), Pages 2828?2840, 2009.
- [13] J. Chen, R. Wang, S. Yan, S. Shan, X. Chen, and W. Gao. Enhancing human face detection by resampling examples through manifolds., *IEEE Transactions on Systems, Man, and Cybernetics-Part A*, pages vol. 37, no. 6, pp. 1017–1028, 2007.
- [14] A. Del Bue, J. Xavier, L. Agapito, and M. Paladini. Bilinear modeling via augmented lagrange multipliers (balm). *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(8):1496 –1508, 2012.
- [15] I. Ramirez, P. Sprechmann, and G. Sapiro Classification and clustering via dictionary learning with structured incoherence and shared features. In Proc. IEEE Computer Vision and Pattern Recognition, 2010.
- [16] J.J. Hull A database for handwritten text recognition research. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1994
- [17] Z. Kalal, K. Mikolajczyk, and J. Matas. Tracking-learning-detection. *IEEE Trans. PAMI*, pages pp:1409–1422, 2012.
- [18] K. Koh, S.-J. Kim, and S. P. Boyd. An interior-point method for large-scale l1-regularized logistic regression. *Journal of Machine Learning Research*, 2007.
- [19] A. Martinez and R. Benavente. The ar face database. *CVC technical report*. CVC Technical Report 24, 1998.
- [20] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000.
- [21] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, , and Y. Ma. Towards a practical face recognition system: Robust alignment and illumination by sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012
- [22] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2009.
- [23] A. Del Bue, J. Xavier, L. Agapito, and M. Paladini. Bilinear Factorization via Augmented Lagrange Multipliers in 11th European Conference on Computer Vision, *ECCV 2010*.
- [24] YiGang Peng, Arvind Ganesh, John Wright, Wenli Xu, Yi Ma. *RASL: Robust Alignment by Sparse and Low-Rank Decomposition for Linearly Correlated Images* *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012.
- [25] T. Lin and H. Zha, *Riemannian Manifold Learning*, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 5, pp. 796-809, 2008
- [26] Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E. *Labeled faces in the wild: A database for studying face recognition in unconstrained environments*. *Technical Report 07-49*, University of Massachusetts, Amherst
- [27] E. Learned-Miller, *Data driven image models through continuous joint alignment*, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, pp. 236-250, 2006.
- [28] G. B. Huang, V. Jain, and E. Learned-Miller, Unsupervised joint alignment of complex images, in Proc. of IEEE International Conference on Computer Vision, 2007.
- [29] M. Cox, S. Lucey, S. Sridharan, and J. Cohn, Least squares congealing for unsupervised alignment of images, in Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2008.
- [30] A. Vedaldi, G. Guidi, and S. Soatto, Joint alignment up to (lossy) transformations, in Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2008.
- [31] Baochang Zhang, Alessandro Perina, Vittorio Murino, Alessio Del Bue: Sparse representation classification with manifold constraints transfer. *CVPR 2015: 4557-4565*
- [32] M. Fazel, H. Hindi, and S. Boyd, Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices, in Proc. of American Control Conference.
- [33] Hongyang Zhang, Zhouchen Lin, Chao Zhang, and Edward Chang, Exact Recoverability of Robust PCA via Outlier Pursuit with Tight Recovery Bounds, *AAAI2015*
- [34] Canyi Lu, Zhouchen Lin, and Shuicheng Yan, Smoothed Low Rank and Sparse Matrix Recovery by Iteratively Reweighted Least Squared Minimization, *IEEE Trans. Image Processing*, Vol. 24, No. 2, pp. 646-654, 2015
- [35] H.Quang Minh, Loris Bazzani, and Vittorio Murino. A unifying framework for vector-valued manifold regularization and multi-view learning. Proceedings of the 30th International Conference on Machine Learning (ICML), pages 100-108, 2013.
- [36] Zhiwu Huang, Ruiping Wang, Shiguang Shan, Xilin Chen, Hybrid Euclidean-and-Riemannian Metric Learning for Image Set Classification, in Proceedings of Asian Conference on Computer Vision, Singapore, Nov. pages 562-577, 2014
- [37] Ruiping wang, Shiguang shan, Xilin Chen,Q. Dai, W. Gao, manifold-manifold distance and its application to face recognition with image sets. *IEEE Trans. image processing*, vol.21, 2012.
- [38] J. B. Tenenbaum, V. de Silva, J. C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, *Science* 290(2000): 2319-2323.
- [39] C. Varini, A. Degenhard, and T. Nattkemper Isolle: Locally linear embedding with geodesic distance. In *Proc. Knowledge Discovery in Database: PKDD*, pages 331-342, 2005.
- [40] Xiaoshuang Shi, Zhenhua Guo, Zhihui Lai, Face recognition by sparse discriminant analysis via jointL2,1-norm minimization, *Pattern Recognition*, 2014, 47(7):2447-2453.
- [41] Zhihui Lai,Waikeung Wong,Yong Xu, Sparse alignment for robust tensor learning, *IEEE Transactions on Neural Networks and Learning Systems*,2014,25(10): 1779-1792.
- [42] Zhihui Lai, Yong Xu, Zhong Jin, David Zhang, Human gait recognition via sparse discriminant projection learning, *IEEE Transactions on Circuits and Systems for Video Technology*, 2014,24(10): 1651-1662.

## 6 APPENDIX\*

Here we show more results of MADMM in Fig. 6 and Fig. 7.

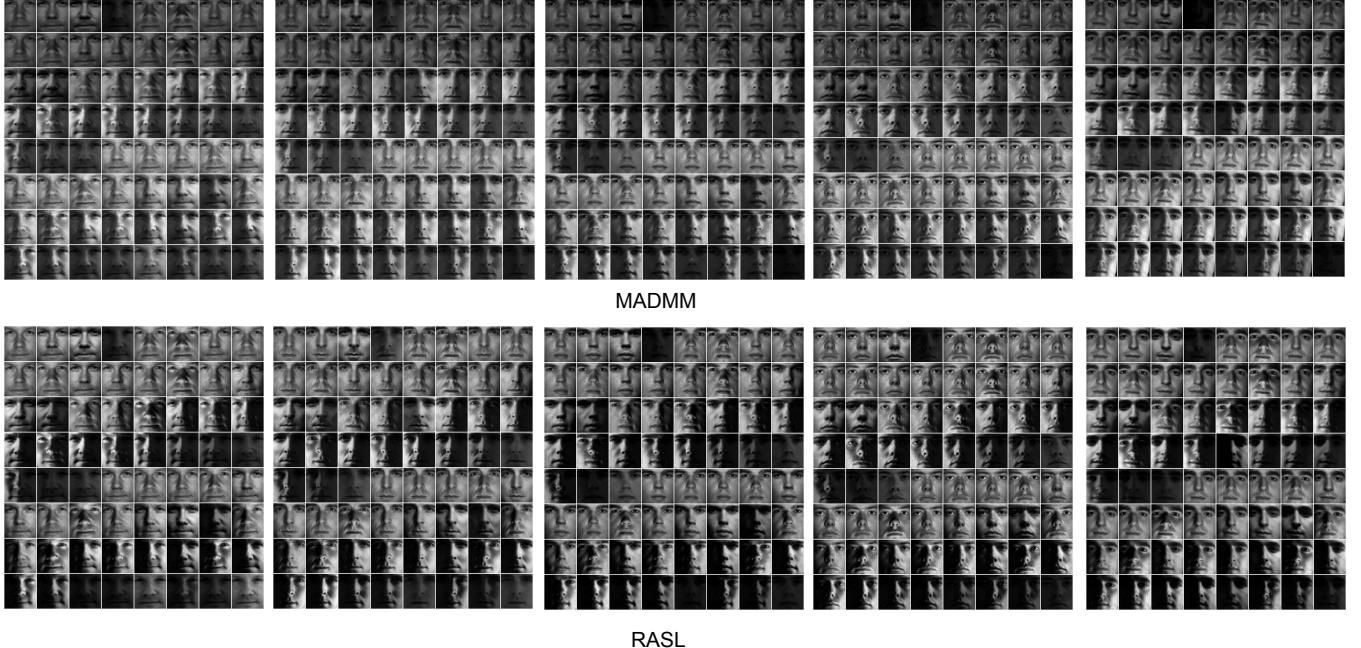


Fig. 6. Comparison of RASL and MADMM on the Yale-B database; The samples in the first row are the results of RASL, and those in the second row are the results of MADMM

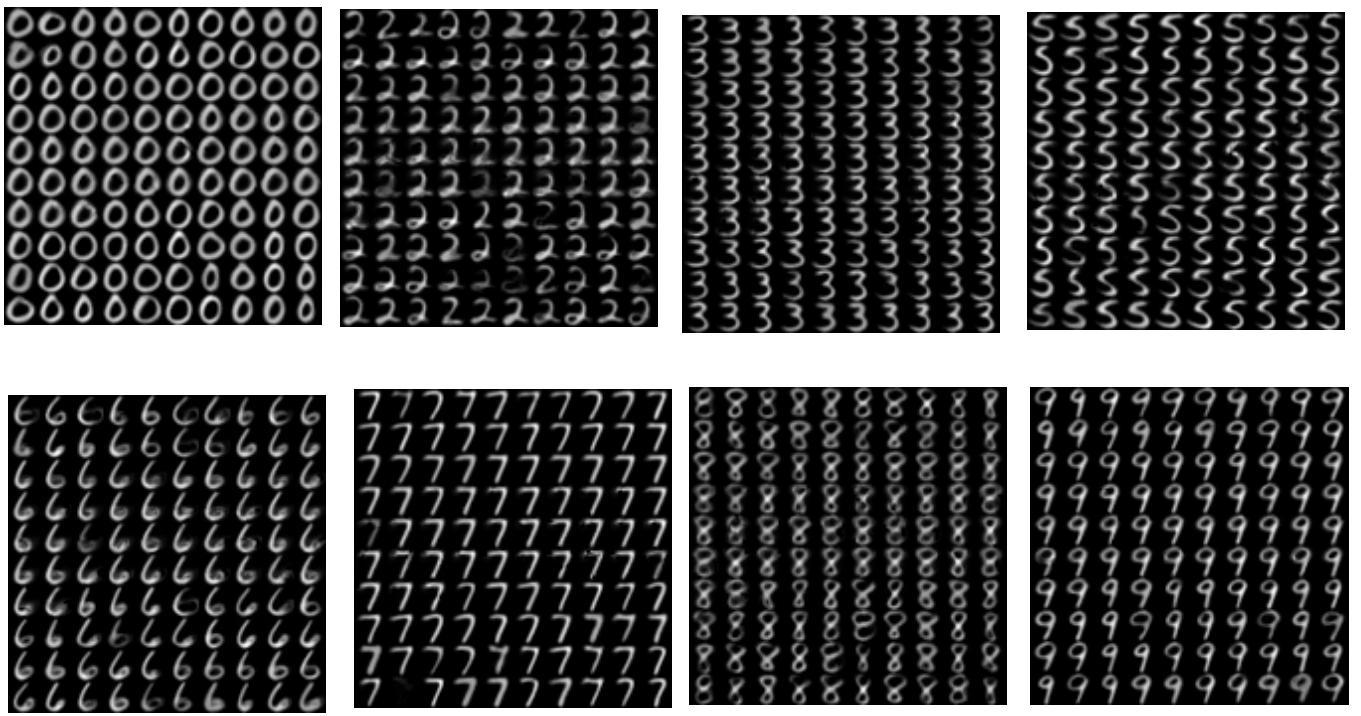


Fig. 7. The MADMM results on the USPS database