Image Set Representation with L_1-Norm Optimal Mean Robust Principal Component Analysis

Youxia Cao, Bo Jiang[(✉)], Jin Tang, and Bin Luo

School of Computer Science and Technology, Anhui University,
No. 111 Jiulong Road, Hefei, China
ahu_youxia@foxmail.com, {jiangbo,tj,luobin}@ahu.edu.cn

Abstract. Many problems in computer vision area can be formulated as image set representation and classification. One main challenge is that image set data usually contains various kinds of noises and outliers which usually make the recognition/learning tasks of image set more challengeable. In this paper, we propose a new L_1 norm optimal Mean Principal Component Analysis (L1-MPCA) to learn an optimal low-rank representation for image set. Comparing with original observed image set, L1-MPCA based low-rank representation is generally noiseless and thus can encourage more robust learning process. An effective update algorithm has been proposed to solve the proposed L1-MPCA model. Experimental results on several datasets demonstrate the effectiveness and robustness of the proposed L1-MPCA method.

1 Introduction

Object recognition/learning based on visual content information is a fundamental problem in computer vision area. Recently, image set based object recognition approaches have been widely studied and attracted more and more interest. For image set based object recognition, it aims to achieve object recognition/learning problem by using multiple images (or video) that belong to one object.

One problem for image set based object recognition process is how to effectively represent an image set [5,9,11,22,26]. In recent years, many methods have been proposed for this problem. One kind of popular methods is to use statistical models. These methods usually aim to represent an image set by using some distributions, such as Gaussian, GMM, etc. [24]. Based on these representation, the similarity measurement between two image sets can be computed by metric measurement between distributions [1,18,20,22]. Another kind of methods is to use linear subspace models which aims to represent an image set by using a linear or affine subspace. Based on these representation, one can compute the distance between two image sets by measuring the distance between two subspaces [9,10,15,25]. Recent studies also aim to represent an image set using a nonlinear manifold or several sub-manifolds. Then, they generally use the metric learning method of manifolds to achieve image set recognition/learning tasks [3,4,7,11,21,23,26]. Some other methods have also been proposed [6].

Previous works generally focus on developing a method for image set feature extraction and classification problems. In this paper, we focus on image set itself and propose a robust low-rank representation for image set. Our method is motivated by the following observation. Motivated by recent work on low-rank representation [12,19], we propose a new low-rank representation, called L_1 norm based robust optimal Mean Principal Component Analysis (L1-MPCA), for image set recovery and representation. The aim of L1-MPCA is to integrate the mean calculation into L1-norm PCA low-rank approximation objective and thus the optimal mean can be obtained to enhance the low-rank approximation and representation. An effective optimization algorithm has been derived to solve the proposed L1-MPCA. Comparing with original observed image sets, L1-MPCA of image sets are generally noiseless and more regular, which significantly encourages the robust learning and recognition process. Experimental results on four datasets demonstrate the effectiveness and robustness of the proposed L1-MPCA methods.

2 Brief Review of Optimal Mean PCA

In this section, we give a brief introduction of Optimal Mean Principal Component Analysis (OMPCA) model [19]. Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_n) \in \mathbb{R}^{p \times n}$ be the input data matrix containing the collection of n data column vectors in p dimension space. In image set representation, each column $\mathbf{x}_i$ denotes one linearized array of pixels gray levels. The aim of Optimal Mean Principal Component Analysis (OMPCA) [19] is to find the optimal low-dimensional matrices $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \cdots \mathbf{u}_k) \in \mathbb{R}^{p \times k}$, $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_k) \in \mathbb{R}^{n \times k}$ and mean vector $\mathbf{b} \in \mathbb{R}^p$ by minimizing,

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{b}} \quad \sum_{i=1}^{n} \|\mathbf{x}_i - \mathbf{U}\mathbf{v}_i - \mathbf{b}\|_2^2 = \|\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}}\|_F^2 \qquad s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \quad (1)$$

where $\mathbf{1} = (1, 1, \cdots, 1) \in \mathbb{R}^n$. Let $\mathbf{Z} = \mathbf{U}\mathbf{V}^{\mathrm{T}} + \mathbf{b}\mathbf{1}^{\mathrm{T}}$, then $\mathbf{Z}$ provides a kind of low-rank representation for original input data $\mathbf{X}$. It is known that the squared loss function used in the above MPCA is very sensitive to outliers. In order to overcome this problem, Nie et al. [19] also propose a kind of Robust MPCA by using $L_{2,1}$ norm and solve the optimization problem as

$$\min_{\mathbf{U}, \mathbf{V}, \mathbf{b}} \quad \sum_{i=1}^{n} \|\mathbf{x}_i - \mathbf{U}\mathbf{v}_i - \mathbf{b}\|_2 = \|\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}}\|_{2,1} \qquad s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \quad (2)$$

Comparing with Frobenious norm loss function, $L_{2,1}$-norm loss function performs robustly w.r.t outliers because it uses a non-squared loss function.

3 L_1-Norm Based Robust MPCA

The above $L_{2,1}$-norm OMPCA is robustness to outliers. However, it is sensitive to the corruptions or large errors existing in each image $\mathbf{x}_i$ because of L2 norm

loss function for each image data. Our aim in this section is to propose a new kind of robust OMPCA by using L_1 norm loss function instead of $L_{2,1}$ norm loss function.

Model formulation. Formally, let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) \in \mathbb{R}^{p \times n}$ be the image set data, our L_1 norm based MPCA (L1-MPCA) is formulated as,

$$\min_{\mathbf{U},\mathbf{V},\mathbf{b}} \quad \sum_{i=1}^{n} \|\mathbf{x}_i - \mathbf{U}\mathbf{v}_i - \mathbf{b}\|_1 = \|\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}}\|_1 \qquad s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \quad (3)$$

where L_1 norm loss function is defined as $\|\mathbf{A}\|_1 = \sum_i \sum_j |\mathbf{A}_{ij}|$. It is known that the L_1 norm loss function will make the the proposed L1-MPCA robust to both corruptions noise/large errors and outliers. Note that the above L1-MPCA can be regarded as a natural extension of the tractional L1-PCA model [2,13] by further removing optimal mean automatically from the input data set $\mathbf{X}$.

Optimization. We present an effective updating algorithm to solve L1-MPCA model. Firstly, Eq. (3) can be rewritten equivalently as

$$\min_{\mathbf{U},\mathbf{V},\mathbf{E},\mathbf{b}} \quad \|\mathbf{E}\|_1 \quad s.t. \quad \mathbf{E} = \mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}}, \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \qquad (4)$$

We use the Augmented Lagrange Multiplier (ALM) method to solve this problem. ALM solves a sequences of subproblems

$$\min_{\mathbf{U},\mathbf{V},\mathbf{E},\mathbf{b}} \quad \|\mathbf{E}\|_1 + \mathrm{Tr} \quad \Omega^{\mathrm{T}}(\mathbf{E} - \mathbf{X} + \mathbf{U}\mathbf{V}^{T} + \mathbf{b}\mathbf{1}^{\mathrm{T}}) + \frac{2}{\mu}\|\mathbf{E} - \mathbf{X} + \mathbf{U}\mathbf{V}^{T} + \mathbf{b}\mathbf{1}^{\mathrm{T}}\|_F^2$$
$$s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \qquad (5)$$

where Ω is Lagrange multipliers and μ is the penalty parameter. There are two major parts of this algorithm, i.e., solving the sub-problem and updating parameters (Ω, μ).

First, we rewrite the objective function of Eq. (5) as

$$\min_{\mathbf{U},\mathbf{V},\mathbf{E},\mathbf{b}} \quad \|\mathbf{E}\|_1 + \frac{2}{\mu}\|\mathbf{E} - (\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}} + \frac{\Omega}{\mu})\|_F^2 \qquad s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \quad (6)$$

Then, we iteratively solve the following sub-problems until convergence.

(1) Solve $\mathbf{U}, \mathbf{V}, \mathbf{b}$ while fixing $\mathbf{E}$. The problem becomes

$$\min_{\mathbf{U},\mathbf{V},\mathbf{b}} \|(\mathbf{X} - \mathbf{E} - \frac{\Omega}{\mu}) - \mathbf{b}\mathbf{1}^{\mathrm{T}} - \mathbf{U}\mathbf{V}^{\mathrm{T}}\|_F^2 \qquad s.t. \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} \qquad (7)$$

This is standard MPCA [19] and can be solved effectively using a closed-form solution.

(2) Solve $\mathbf{E}$ while fixing $\mathbf{U}, \mathbf{V}, \mathbf{b}$. The problem becomes

$$\min_{\mathbf{E}} \quad \|\mathbf{E}\|_1 + \frac{2}{\mu}\|\mathbf{E} - (\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}} + \frac{\Omega}{\mu})\|_F^2 \qquad (8)$$

It is well known that, this problem has closed-form solution,

$$\mathbf{E}_{ij} = \text{sign}(\mathbf{K}_{ij})(|\mathbf{K}_{ij}| - \frac{1}{\mu})_{+}, \quad \mathbf{K} = \mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{b}\mathbf{1}^{\mathrm{T}} + \frac{\Omega}{\mu} \qquad (9)$$

(3) At the end of each ALM iteration, Ω, μ are updated as

$$\begin{aligned} \Omega &= \Omega + \mu(\mathbf{X} - \mathbf{U}\mathbf{V}^{\mathrm{T}} - \mathbf{E}) \\ \mu &= \rho\mu \end{aligned} \qquad (10)$$

where $\rho > 1$.

4 Application: Image Set Representation and Classification

In this section, we apply the proposed L1-MPCA in image set representation and classification tasks. Our image set representation and classification method contains two main steps.

First, given an image set $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_n)$, we first use the proposed L1-MPCA to compute the optimal $\mathbf{U}^*, \mathbf{V}^*$ and mean vector $\mathbf{b}^*$. We then obtain the optimal low-rank representation $\mathbf{Z}$ as

$$\mathbf{Z} = \mathbf{X} - \mathbf{U}^*\mathbf{V}^{*\mathrm{T}} - \mathbf{b}^*\mathbf{1}^{\mathrm{T}}$$

Comparing with the original image set data $\mathbf{X}$, the noises of images and outliers in image set $\mathbf{X}$ can be well suppressed in its low-rank representation $\mathbf{Z}$.

Second, based on low-rank representation $\mathbf{Z}$, we can use some image set feature extraction and learning methods such as Covariance Discriminative Learning (CDL) [22], Covariate-relation graph (CRG) [6] and Manifold-Manifold Distance (MMD) [23] to conduct image set classification tasks.

5 Experiments

To evaluate the effectiveness of the proposed L1-MPCA method, we apply it in image set representation and classification tasks. For image set learning methods, we use some recent methods: Covariance Discriminative Learning (CDL) [22], Covariate-relation graph (CRG) [6], Manifold-Manifold Distance (MMD) [23], Set to Set Distance Metric Learning (SSDML) [27] and Canonical Correlations (DCC) [15]. According to [23], MMD method does the subspaces learning with 95% data energy based on PCA. For the discriminative learning method of CDL, we choose PLS to do the learning task. For SSDML method, we set $\nu = 1, \lambda_1 = 0.001$ and $\lambda_2 = 0.1$.

5.1 Datasets and Settings

In our experiments, we test our L1-MPCA on four datasets including YouTube Celebrities (YTC) [14], ETH-80 [17], Honda/UCSD [16] and CMU MoBo [8]. In each image set data, we resize all the images into 20×20 intensity images. The datasets are described as following.

– ETH-80 [17] dataset has image sets of 8 categories and each category contains 10 objects with 41 views per object, spaced equally over the viewing hemisphere, for a total of 3280 images. For each subject of this dataset, we randomly choose 5 sets for training and the rest 5 object for testing.
– YouTube-Celebrities (YTC) [14] dataset contains 1910 video clips of 47 celebrities (actors and politicians), most of the videos are low resolution and highly compression, which leads to noisy, low-quality image frames. Each clip contains hundreds of frames. For this dataset, we randomly chose 3 sets for training and 6 sets for testing.
– Honda/UCSD [16] dataset consists of 59 video sequences belonging to 20 different persons. Each sequence contains about 400 frames covering large variations. Each individual in our database has at least two videos. For this dataset, we randomly select one sequence for training set and the rest for testing.
– CMU MoBo [8] dataset has 96 sequences of 24 persons and the sequences are captured from different walking situations inclined walk, and slow walk holding a ball (to inhibit arm swing). Each video further divided into four illumination sets, the first set for training and the rest sets for testing.

5.2 Results Analysis

To evaluate the benefit of the proposed L1-MPCA low-rank representation method, we compare our method with original data and L1-PCA method [2,13]. Figure 1 summarizes the average classification results on four datasets, respectively. (1) Comparing with original image set $\mathbf{X}$, the proposed L1-MPCA method can significantly improve the image set classification results, which clearly demonstrates the desired benefit and effectiveness of the proposed L1-MPCA method on conducting image set representation problem and thus leads to better classification result. (2) The proposed L1-MPCA methods generally performs better than L1-PCA method [2]. This clearly demonstrates the benefit of the proposed L1-MPCA by further considering the optimal mean vector value in low-rank representation.

5.3 Robust to Noise

To evaluate the robustness of L1-MPCA method to the noise possibly appearing in the testing image set data, we randomly add some noise to the image set datasets. Here, we add two kinds of noises including salt & pepper and block
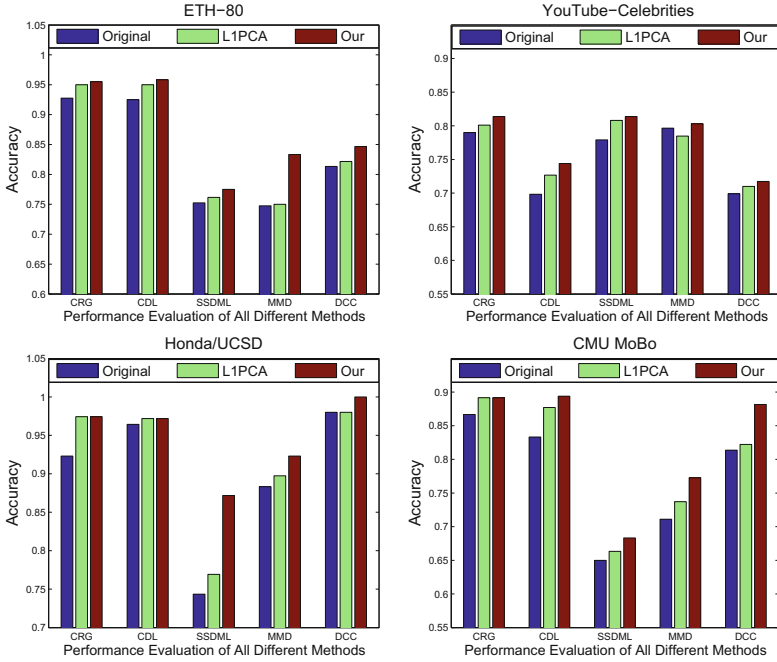
Fig. 1. Average accuracies of different methods on four datasets.

Table 1. Classification accuracies on ETH80 dataset with different noises.

Methods	Noisead	Salt & Pepper					Block noise				
		0.1	0.2	0.3	0.4	0.5	5	15	25	35	45
CDL	Original	0.800	0.700	0.625	0.583	0.425	0.833	0.750	0.708	0.637	0.405
	L1-PCA	0.875	0.739	0.667	0.617	0.458	0.901	0.843	0.763	0.665	0.567
	L1-MPCA	**0.917**	**0.845**	**0.738**	**0.708**	**0.667**	**0.922**	**0.861**	**0.783**	**0.696**	**0.607**
CRG	Original	0.875	0.800	0.775	0.653	0.550	0.875	0.833	0.747	0.675	0.475
	L1-PCA	0.901	0.833	0.797	0.732	0.625	0.911	0.866	0.774	0.697	0.583
	L1-MPCA	**0.916**	**0.875**	**0.811**	**0.767**	**0.673**	**0.937**	**0.875**	**0.801**	**0.721**	**0.595**
MMD	Original	0.712	0.625	0.550	0.525	0.475	0.666	0.583	0.500	0.466	0.431
	L1-PCA	0.750	0.667	0.625	0.542	0.491	0.776	0.673	0.575	0.533	0.461
	L1-MPCA	**0.811**	**0.708**	**0.650**	**0.593**	**0.500**	**0.807**	**0.767**	**0.702**	**0.647**	**0.573**
SSDML	Original	0.700	0.650	0.575	0.473	0.438	0.708	0.671	0.583	0.501	0.467
	L1-PCA	0.733	0.666	0.583	0.483	0.443	0.750	0.683	0.596	0.511	0.475
	L1-MPCA	**0.767**	**0.708**	**0.611**	**0.573**	**0.542**	**0.783**	**0.713**	**0.633**	**0.575**	**0.511**
DCC	Original	0.733	0.613	0.553	0.437	0.339	0.700	0.627	0.566	0.483	0.408
	L1-PCA	0.803	0.722	0.637	0.583	0.524	0.788	0.718	0.643	0.522	0.466
	L1-MPCA	**0.833**	**0.797**	**0.703**	**0.627**	**0.573**	**0.800**	**0.758**	**0.650**	**0.567**	**0.511**

Table 2. Classification accuracies on YTC dataset with different noises.

Methods	Noisead	Salt & Pepper					Block noise				
		0.1	0.2	0.3	0.4	0.5	5	15	25	35	45
CDL	Original	0.651	0.573	0.432	0.358	0.238	0.650	0.510	0.413	0.366	0.227
	L1-PCA	0.703	0.654	0.587	0.422	0.317	0.713	0.654	**0.583**	**0.471**	0.328
	L1-MPCA	**0.709**	**0.666**	**0.595**	**0.433**	**0.344**	**0.733**	**0.682**	0.573	0.453	**0.393**
CRG	Original	0.742	0.666	0.521	0.453	0.373	0.707	0.611	0.532	0.444	0.309
	L1-PCA	0.771	0.683	0.511	**0.475**	**0.383**	0.783	0.627	0.575	0.467	0.350
	L1-MPCA	**0.783**	**0.683**	**0.537**	0.465	0.377	**0.802**	**0.741**	**0.666**	**0.583**	**0.405**
MMD	Original	0.701	0.621	0.575	0.466	0.393	0.696	0.605	0.528	0.499	0.336
	L1-PCA	0.766	0.666	**0.611**	0.505	0.423	0.750	0.653	**0.583**	0.511	0.471
	L1-MPCA	**0.783**	**0.683**	0.601	**0.542**	**0.447**	**0.778**	**0.683**	0.566	**0.533**	**0.505**
SSDML	Original	0.701	0.621	0.531	0.401	0.389	0.711	0.627	0.550	0.448	0.342
	L1-PCA	0.773	0.646	0.562	0.473	0.382	0.766	0.631	0.575	**0.499**	0.377
	L1-MPCA	**0.792**	**0.696**	**0.583**	**0.481**	**0.393**	**0.773**	**0.674**	**0.586**	0.473	**0.383**
DCC	Original	0.656	0.567	0.441	0.397	0.217	0.683	0.578	0.450	0.349	0.207
	L1-PCA	0.697	0.577	0.465	**0.453**	0.311	0.721	0.649	0.557	0.467	0.350
	L1-MPCA	**0.703**	**0.650**	**0.583**	0.437	**0.343**	**0.727**	**0.673**	**0.579**	**0.483**	**0.366**

Table 3. Classification accuracies on Honda dataset with different noises.

Methods	Noisead	Salt & Pepper					Block noise				
		0.1	0.2	0.3	0.4	0.5	5	15	25	35	45
CDL	Original	0.871	0.799	0.730	0.666	0.432	0.883	0.777	0.653	0.577	0.450
	L1-PCA	0.875	0.783	**0.741**	0.696	0.601	0.911	0.833	0.713	0.641	0.522
	L1-MPCA	**0.916**	**0.833**	0.731	**0.700**	**0.626**	**0.921**	**0.850**	**0.743**	**0.650**	**0.583**
CRG	Original	0.910	0.811	0.696	0.633	0.473	0.899	0.751	0.583	0.434	0.405
	L1-PCA	0.948	0.833	0.701	0.650	0.515	0.911	0.811	0.637	0.557	0.466
	L1-MPCA	**0.950**	**0.844**	**0.733**	**0.666**	**0.533**	**0.937**	**0.823**	**0.710**	**0.595**	**0.511**
MMD	Original	0.871	0.734	0.590	0.550	0.433	0.846	0.711	0.573	0.466	0.437
	L1-PCA	0.866	0.750	0.611	0.576	0.433	0.897	0.801	0.695	0.505	0.444
	L1-MPCA	**0.883**	**0.786**	**0.650**	**0.592**	**0.498**	**0.901**	**0.811**	**0.722**	**0.583**	**0.466**
SSDML	Original	0.743	0.666	0.511	0.433	0.399	0.718	0.635	0.544	0.473	0.366
	L1-PCA	0.769	0.683	0.573	0.450	0.409	0.766	0.677	0.595	0.500	0.415
	L1-MPCA	**0.794**	**0.693**	**0.610**	**0.493**	**0.433**	**0.808**	**0.731**	**0.606**	**0.593**	**0.511**
DCC	Original	0.950	0.883	0.766	0.613	0.526	0.943	0.837	0.727	0.579	0.433
	L1-PCA	0.967	**0.922**	0.776	0.696	0.633	0.950	0.871	0.733	0.652	0.566
	L1-MPCA	**0.983**	0.911	**0.797**	**0.701**	**0.650**	**0.974**	**0.883**	**0.766**	**0.666**	**0.583**

noise. For each kind of noise, we add various levels of the noises and test our method on these noise image set data. Tables 1, 2, 3 and 4 show the accuracies of all the traditional methods across different noise level. From the results, we can note that: (1) As the level of noise increasing. Our L1-MPCA method still maintains better performance comparing with the original image set data. This obviously indicates the noise removing ability of the proposed L1-MPCA method. (2) L1-MPCA performs better than L1-PCA method [2], indicating the more robustness of the L1-MPCA method on noise data reconstruction.

Table 4. Classification accuracies on CMU dataset with different noises.

Methods	Noisead	Salt & Pepper					Block noise				
		0.1	0.2	0.3	0.4	0.5	5	15	25	35	45
CDL	Original	0.797	0.691	0.618	0.543	0.344	0.808	0.750	0.666	0.487	0.301
	L1-PCA	0.833	0.766	0.650	0.575	0.483	0.827	0.775	0.683	0.506	0.466
	L1-MPCA	**0.850**	**0.799**	**0.723**	**0.644**	**0.490**	**0.866**	**0.818**	**0.750**	**0.566**	**0.482**
CRG	Original	0.825	0.766	0.671	0.543	0.321	0.803	0.733	0.683	0.450	0.320
	L1-PCA	0.873	0.800	**0.743**	0.628	0.495	0.850	0.766	0.679	0.637	0.449
	L1-MPCA	**0.875**	**0.810**	0.723	**0.683**	**0.506**	**0.866**	**0.797**	**0.683**	**0.650**	**0.455**
MMD	Original	0.711	0.622	0.533	0.505	0.341	0.666	0.550	0.471	0.416	0.283
	L1-PCA	0.750	0.683	0.621	0.575	0.483	0.776	0.707	0.650	0.571	0.361
	L1-MPCA	**0.811**	**0.755**	**0.643**	**0.591**	**0.513**	**0.807**	**0.732**	**0.666**	**0.583**	**0.421**
SSDML	Original	0.696	0.611	0.550	0.466	0.377	0.700	0.650	0.541	0.483	0.322
	L1-PCA	0.703	0.647	0.595	0.517	0.466	0.717	**0.666**	0.606	0.533	0.450
	L1-MPCA	**0.731**	**0.650**	**0.601**	**0.543**	**0.483**	**0.733**	0.650	**0.622**	**0.571**	**0.500**
DCC	Original	0.783	0.711	0.637	0.550	0.366	0.803	0.733	0.677	0.543	0.391
	L1-PCA	0.791	0.710	**0.672**	0.606	0.450	0.845	0.750	0.683	0.577	0.411
	L1-MPCA	**0.810**	**0.737**	0.666	**0.611**	**0.473**	**0.850**	**0.758**	**0.683**	**0.601**	**0.483**

6   Conclusion

In this paper, we propose a new method, called L_1 norm Mean PCA (L1-MPCA) model for image set representation and learning problems. L1-MPCA is robust to both noises and outliers, which encourages robust image set learning tasks. An effective update algorithm has been proposed to solve the proposed L1-MPCA model. Experimental results on several datasets show the benefit and robustness of the proposed L1-MPCA method. In our future, we will further consider the manifold structure of data in our L1-MPCA model.

References

1. Arandjelovic, O., Shakhnarovich, G., Fisher, J., Cipolla, R., Darrell, T.: Face recognition with image sets using manifold density divergence. In: 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 581–588. IEEE (2005)
2. Cao, Y., Jiang, B., Chen, Z., Tang, J., Luo, B.: Low-rank image set representation and classification. In: Liu, C.-L., Hussain, A., Luo, B., Tan, K.C., Zeng, Y., Zhang, Z. (eds.) BICS 2016. LNCS (LNAI), vol. 10023, pp. 321–330. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49685-6_29

3. Cevikalp, H., Triggs, B.: Face recognition based on image sets. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2567–2573. IEEE (2010)
4. Chen, S., Sanderson, C., Harandi, M.T., Lovell, B.C.: Improved image set classification via joint sparse approximated nearest subspaces. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 452–459. IEEE (2013)
5. Chen, S., Wiliem, A., Sanderson, C., Lovell, B.C.: Matching image sets via adaptive multi convex hull. In: 2014 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1074–1081. IEEE (2014)
6. Chen, Z., Jiang, B., Tang, J., Luo, B.: Image set representation and classification with covariate-relation graph. In: IEEE Conference on Asian Conference and Pattern Recognition (ACPR), pp. 750–754. IEEE (2015)
7. Cui, Z., Shan, S., Zhang, H., Lao, S., Chen, X.: Image sets alignment for video-based face recognition. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2626–2633. IEEE (2012)
8. Gross, R., Shi, J.: The CMU motion of body (MoBo) database. Technical report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)
9. Hamm, J., Lee, D.D.: Grassmann discriminant analysis: a unifying view on subspace-based learning. In: Proceedings of the 25th International Conference on Machine Learning, pp. 376–383. ACM (2008)
10. Harandi, M.T., Sanderson, C., Shirazi, S., Lovell, B.C.: Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2705–2712. IEEE (2011)
11. Hu, Y., Mian, A.S., Owens, R.: Sparse approximated nearest points for image set classification. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 121–128. IEEE (2011)
12. Jiang, B., Ding, C., Luo, B., Tang, J.: Graph-laplacian PCA: closed-form solution and robustness. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3492–3498. IEEE (2013)
13. Ke, Q., Kanade, T.: Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming. In 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 739–746. IEEE (2005)
14. Kim, M., Kumar, S., Pavlovic, V., Rowley, H.: Face tracking and recognition with visual constraints in real-world videos. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8. IEEE (2008)
15. Kim, T.-K., Kittler, J., Cipolla, R.: Discriminative learning and recognition of image set classes using canonical correlations. IEEE Trans. Pattern Anal. Mach. Intell. **29**(6), 1005–1018 (2007)
16. Lee, K.-C., Ho, J., Yang, M.-H., Kriegman, D.: Video-based face recognition using probabilistic appearance manifolds. In: 2003 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. I–313. IEEE (2003)
17. Leibe, B., Schiele, B.: Analyzing appearance and contour based methods for object categorization. In: 2003 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. II–409. IEEE (2003)
18. Lu, J., Wang, G., Moulin, P.: Image set classification using holistic multiple order statistics features and localized multi-kernel metric learning. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 329–336. IEEE (2013)
19. Nie, F., Yuan, J., Huang, H.: Optimal mean robust principal component analysis. In: Proceedings of the 31st International Conference on International Conference on Machine Learning, vol. 32, pp. 1062–1070 (2014)

20. Shakhnarovich, G., Fisher, J.W., Darrell, T.: Face recognition from long-term observations. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002 Part III. LNCS, vol. 2352, pp. 851–865. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-47977-5_56

21. Wang, R., Chen, X.: Manifold discriminant analysis. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 429–436. IEEE (2009)

22. Wang, R., Guo, H., Davis, L.S., Dai, Q.: Covariance discriminative learning: a natural and efficient approach to image set classification. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2496–2503. IEEE (2012)

23. Wang, R., Shan, S., Chen, X., Gao, W.: Manifold-manifold distance with application to face recognition based on image set. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8. IEEE (2008)

24. Wang, W., Wang, R., Huang, Z., Shan, S., Chen, X.: Discriminant analysis on Riemannian manifold of Gaussian distributions for face recognition with image sets. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2048–2057. IEEE (2015)

25. Yamaguchi, O., Fukui, K., Maeda, K.-I.: Face recognition using temporal image sequence. In: 1998 Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 318–323. IEEE (1998)

26. Yang, M., Zhu, P., Van Gool, L., Zhang, L.: Face recognition based on regularized nearest points between image sets. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1–7. IEEE (2013)

27. Zhu, P., Zhang, L., Zuo, W., Zhang, D.: From point to set: extend the learning of distance metrics. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 2664–2671. IEEE (2013)