

A Robust Framework For Eigenspace Image Reconstruction

Yuan XiaoTong¹ Zhu HongWen² Yang ShuTang²

1 School of Information Security Engineering, Shanghai Jiao Tong University

2 Department of Electronic Engineering, Shanghai Jiao Tong University

1954 Huashan Road, Shanghai, 200030 China

E-mail: { yxt, guolihua, styang }@sjtu.edu.cn

Abstract

Principal components analysis (PCA) has been proved to be a useful tool for many computer vision and signal processing problems. One drawback of traditional PCA is that they are based on least squares estimation techniques and hence fail to account for “outliers” which are commonly occurred in realistic training sets. To make PCA more robust to real-world problems such as image reconstruction addressed in this paper, we develop a two-step algorithm that can eliminate the outliers on both frame level and pixel level through the LASSO and RPCA separately. With LASSO optimization method, we may obtain sparse projected coefficients of the original image into the basis image space. According to the sparsity of these coefficients, sample outliers can be recognized automatically. Then, with the use of robust M-estimation, the influence of intra-sample outliers may be overwhelmed to great extent. Additionally, due to orthogonality of the principal components, the soft-threshold estimation can be applied to the LASSO to alleviate the computational costs, hence make our robust PCA method more applicable to large scale problems. An experiment on object image reconstruction is used to illustrate the advantage of our proposed technique over standard PCA.

1. Introduction

Principal Component Analysis (PCA) [1], like several other linear learning models and techniques, replace a set of p measured variables by a small set of derived variables. It has been widely used for the representation of high dimensional data such as appearance, shape, motion, et al. [2][3] in computer vision literature. Learned PCA representations have proven useful for solving problems such as image recognition, tracking, and background modeling [4][5][6]. Typically, PCA is posed as a problem of alternated least squares (ALS) estimation. Unfortunately, it is commonly known that least-squares techniques perform badly in many situations due to the statis-

tical outliers [7], which can arbitrarily bias the solution. To improve the robustness of the estimation processes, many works had been done in the statistics and neural networks [7][8]. In the past few years, robust estimation problem has also received enough attention in computer vision community. Recently, the theory of Robust Subspace Learning (RSL) has been developed in literature [5]. The RSL method solves the linear regression problem within a continuous optimization framework based on robust M-estimation. The RSL can be directly applied to PCA, referred to as RPCA. For face image data training, RPCA has obtained the robustness on pixel level, comparing to the frame level robustness achieved by the statistical physics approach [8].

The dimension reduction achieved by PCA is especially useful if the components can be readily interpreted. A frequently used informal approach is to ignore all loadings smaller than some threshold absolute value, effectively treating them as zero. This can be misleading [9]. A more formal way of making some of the loadings zero is to restrict the allowable loadings to a small set of values, for example -1, 0, 1. Vines [10] introduced simple component analysis as an alternative of this method. Jolliffe et al. [11] developed a new technique called SCoT-LASSO (Simplified Component Technique-LASSO) from the LASSO theory (Least Absolute Shrinkage and Selection Operator) [12], which is widely used for multiple regression. The SCoTLASSO method makes the shrinkage of the loadings more reliable and simple and it has achieved great success for the interpretation of the principal components.

In this article, we introduce a new framework, which share an idea central to both De la Torre's [5] and Jolliffe's work [11], to robustly extract the principal components. For illustrative purpose and without loss of generality, we will focus on appearance-based, image reconstruction problem, and show how our method can be used for the robust eigenspace representation. The main idea is that we firstly present an alternative approach to the sparse eigenspace projection using LASSO in order to achieve the robustness on frame level. Then with the

frame-level outliers kicked out, we will use the RPCA method, which can achieve the robustness on pixel level, to determine the basis images. The main contribution here is our modification of LASSO and the introduction of soft-threshold estimation into the numerical implementation.

The following section reviews some related work on RPCA and LASSO. Section 3 develops a sparse subspace projection framework. Some theories and the numerical realization of our algorithm are described in detail in section 4. In Section 5, some experimental results are provided to illustrate how this technique can be used to solve the robust object representation problem.

2. Related works

2.1 Energy Function and RPCA

Let $\{\vec{x}_i \in R^m, i=1,2,\dots,n\}$ be a given set of sample vectors from a m -dimension variable \vec{x} , $X = (\vec{x}_i) \in R^{m \times n}$, n is the number of samples. In the case of standard PCA, we will consider that the data is zero mean. What PCA does, when based on correlation matrix, is to find linear functions $\vec{p}_1' \vec{x}, \vec{p}_2' \vec{x}, \dots, \vec{p}_l' \vec{x}$ which successfully maximum sample variance subject to $P'P = I_l$, $P = [\vec{p}_1, \vec{p}_2, \dots, \vec{p}_l]$. The derived variable $\vec{p}_k' \vec{x}$ is the k th principal component (PC), for convenience purpose we just refer \vec{p}_k' as the k th PC in the context. In the statistics community, a widely used method for calculating the principal components is to formulate PCA as the least-squares estimation of the basis P that minimize the following energy function:

$$E_1(P) = \sum_{i=1}^n e_i = \sum_{i=1}^n \|\vec{x}_i - PP' \vec{x}_i\|_2^2 = \sum_{i=1}^n \sum_{j=1}^m (\vec{x}_{ji} - \sum_{s=1}^l \vec{p}_{js} \vec{p}_{is}' \vec{x}_{ji})^2 \quad (1)$$

, where $e_i = \|\vec{x}_i - PP' \vec{x}_i\|_2^2$ is the reconstruction error of \vec{x}_i . It is obviously known that the energy function (1) is not linear to P , thus may not be efficiently minimized. Alternatively, we can make the linear coefficients explicit and minimize

$$E_2(P, C) = \sum_{i=1}^n \|\vec{x}_i - PC_i\|_2^2 \quad (2)$$

In this case, EM algorithm can be used to estimate both P and C [13].

The standard eigenspace techniques mentioned above for estimating principal components are not robust to outliers that are common in training data [4]. This happens because the energy functions are derived from the least-squares framework. Since 1990s, more and more attention has been paid to the robustness of PCA in com-

puter vision. Based on the previous work on robust PCA, De la Torre and Black [5] exploit the relationship between outlier process and the robust statistics, and formulate the RPCA as the following M-Estimation:

$$E_3(P, C, \mu, \sigma) = \sum_{i=1}^n \sum_{j=1}^m \rho(\vec{x}_{ji} - \vec{\mu}_j - \sum_{s=1}^l \vec{p}_{js} \vec{c}_{si}) \quad (3)$$

, where μ is the robust mean vector, ρ is some specially chosen robust loss function [14]. This robust framework has achieved pixel level robustness in their background modeling experiments [5].

2.2 The LASSO Approach in PCA

Traditional PCA methods find an eigenspace that minimizes the projective reconstruction residual sum of X . However, if m is large or the PCs are highly correlated, then the variances of the components estimates may be unacceptably high. Standard methods for addressing this difficulty include ridge regression and particularly in cases where a more parsimonious model is desired, subset selection. As an alternative to standard ridge regression and subset selection techniques, Tibshirani [12] proposed the LASSO, which minimizes the residual sum of squares under a constraint on the L^1 -norm of the loading vector. Jolliffe's [11] apply the LASSO theory to PCA and developed the so called SCoTLASSO, in which components were estimated by successively maximize $\vec{p}_s' R \vec{p}_s$, subject to $\sum_{j=1}^m |\vec{p}_{js}| \leq t_s$, where $R = XX'$ is the covariance matrix

of data matrix X . As stated earlier, the LASSO fits are usually very sparse, hence very helpful to the interpretation of PCs.

3. A Modified LASSO PCA

Given a set of images, the eigenspace image reconstruction approaches construct a small set of basis images that characterize the majority of the variation in the training set and can be used to approximate any of the sample images. The sample images are always contaminated by the so called outliers. Generally, there are two kinds of outliers: sample outliers and intra-sample outliers. The sample outliers contaminate the entire sample images, and can be regarded as outliers on frame level. The intra-sample outliers, however, only affect some pixels in an image sample, thus can be regarded as outliers on pixel level. A common case of image reconstruction problem always involves both kinds of outliers. As mentioned in section 2, the intra-sample outliers can be solved by the RPCA method. Now, we will develop a LASSO-Based PCA to deal with sample outliers.

3.1 Choice of The L^1 Constrains

The traditional LASSO-Based PCA constrain the L^1 norm of \bar{p}_s . Here, we describe PCA as the following least-square estimation under the constrains of the L^1 norm of \bar{c}_i :

$$\min : E_4(P, C, \mu, \sigma) = \sum_{i=1}^n \sum_{j=1}^m \frac{1}{2} (\bar{x}_{ji} - \bar{\mu}_j - \sum_{s=1}^l \bar{p}_{js} \bar{c}_{si})^2 \quad (4)$$

$$s.t. \quad \|\bar{c}_i\|_1 \leq t_i, \quad i = 1, \dots, n$$

The constraint value t_i determines the approximation accuracy and can thus be viewed as a model-selection parameter. We know that \bar{c} is a l dimension vector, and it is easy to tell that $l \leq n$. Generally n (the number of the sample images) is far less than m (the pixels numbers of one sample image), so we get $l \ll m$, which means the variables of the constrains in (4) are reduced dramatically comparing to that of traditional LASSO-based PCA [11] and is possible to be applicable to large scale problems like image reconstruction.

3.2 Some Interpretations

As is known that LASSO will make the estimated parameters sparse. We will tell that the sparsity of the output \bar{c}_i in constrained programming problem (4) can be helpful to the recognition of sample outlier images. In the following we give some detailed description. As is known that \bar{c}_i is the projection coordinates of the i th sample image into the linear space spanned by the basis images. From section 2.1 we know that the PCs successively maximize the sample variance and the first few PCs may obtain the majority of the variance. Meanwhile, the variance of sample images is introduced, to some extent, by the sample outliers. Thus, to an outlier image, all the components of its \bar{c}_i are close to zero except the first few ones and the shrinkage of \bar{c}_i via LASSO can be very helpful to interpret the relationship between original sample images and the basis images. Through this shrinkage, according the extent of sparsity, we can easily extract the sample outlier images from the training image sequence. Now, we give an example to illustrate the above conclusion. Figure 1 shows a set of synthetic images, from appearance, it is easy to tell that the last image is an outlier sample. Calculation shows that the first 2 PCs obtain 72.3% of the total sample variance (Table 1). By solving problem (4) without the constrains, we give the estimation of each \bar{c}_i in table 2, from which it can

be observed that the last 4 components of $\hat{\bar{c}}_6$ is very close to zero while $\hat{\bar{c}}_1 \sim \hat{\bar{c}}_5$ do not have such property. Through the LASSO, those components are close to zero can be forced to be exactly zero, and outlier can be easily exposed according to the sparsity of $\hat{\bar{c}}_i$, as described precisely in section 4 and 5 a little later.



Fig.1 A set of synthetic images with last one as a sample outlier

Table1 Proportion to the total variance obtained from the first $i(=1, \dots, 6)$ PCs.

First 1	First 2	First 3	First 4	First 5	First 6
48.9%	72.3%	85.6%	93.4%	100%	100%

Table2. Components of each $\hat{\bar{c}}_i$, obtained by solving (4) without constrains

$\hat{\bar{c}}_i$	-0.0904	-0.1435	-0.1573	-0.1638	-0.1596	0.7146
	-0.2612	-0.1034	-0.0074	0.1559	0.1928	0.0234
	0.1205	-0.1074	-0.1146	0.0341	0.0745	-0.0071
	-0.0175	0.0837	-0.0796	-0.0244	0.0378	-0.0001
	0.0026	-0.0169	0.0338	-0.0783	0.0592	-0.0004
	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000

On the other hand, in the SCoTLASSO framework, the mechanism of sparsity is somewhat different. Here, the components of \bar{p}_s are forced to be vanished. Through this kind of shrinkage, it may be possible for us to interpret some relationship among the pixels, but so far we do not know whether this relationship is helpful to the extraction of the outliers.

3.3 Outline of Our Reconstruction Algorithm

Generally, in actual image reconstruction problems, the sample outliers and intra-sample outliers always co-exist. In Theory, the RPCA framework is robust to both kinds of outliers. However, some threshold values in this framework are hard to be determined in the co-exist situation. Here, for the purpose of reconstructing sample images robustly and practically, we now develop a two-step reconstruction algorithm:

Step1: Find out the sample outliers via M-LASSO-PCA;

Step2: With the sample outliers kicked out, RPCA is applied to deal with intra-sample outliers and robustly reconstruct the sample images.

The outline of our algorithm is very clear. The key problems left are the numerical implementation issues.

4. Computational Issues

So far, many numerical algorithms have been presented for the LASSO and LASSO-based PCA [11] [12][15]. For the problem issued in this paper, we find out that the soft-threshold estimator [15] and the projected gradient approach [16] works well because programming (3) and (4) is always under the condition that $P'P = \mathbf{I}_l$. Now we will describe our method in details.

4.1 Soft-Threshold Estimation(STE)

For the purpose of introducing STE into M-LASSO-PCA, we firstly present two theorems (The proofs are elided due to the limit of the pages).

Theorem 1: Suppose that $\bar{\mu}$ and P are known in (3), $\bar{x}_i = P\bar{c}_i + \bar{\mu} + e_i$, e_{ji} , $j=1, \dots, m$ are zero-mean i.i.d., ρ is a convex function, then $\hat{c}_i - \bar{c}_i \rightarrow N(0, \alpha_i^2 \mathbf{I}_l)$ in distribution, where \hat{c} is the M-estimator of \bar{c}_i , α_i is some constant determined by ρ function and e_i .

Theorem 2: Suppose that $\bar{\mu}$ and P are known in (4), then the estimation of the i th project coefficients vector \bar{c}_i is equivalent to the following constrained least square estimation:

$$\min : E_4(P, C, \mu, \sigma) = \sum_{j=1}^m \frac{1}{2} (\bar{x}_{ji} - \bar{\mu}_j - \sum_{s=1}^l \bar{p}_{js} \bar{c}_{si})^2 \quad (5)$$

$$s.t. \|\bar{c}_i\|_1 \leq t_i$$

Suppose that $\bar{\mu}$, P and \bar{c}_i has been estimated as $\hat{\mu}$, \hat{P} and \hat{c}_i^0 in (4) (without the consideration of constraints) by some method, since quadratic function x^2 is convex, with the above theorem (m is very large and we approximately present that $\hat{c}_i^0 - \bar{c}_i \sim N(0, \alpha_i^2 \mathbf{I}_l)$), it is easy to get an asymptotic optimal solution \hat{c}_i for the constrained programming problem (4) using the so called “soft-threshold” estimator [15]. According to STE, the LASSO estimate of \hat{c}_i has the form:

$$\hat{c}_{si} = \text{sign}(\hat{c}_{si}^0) \max(0, |\hat{c}_{si}^0| - \gamma_i), \quad s=1, \dots, l$$

, where the constant γ_i is chosen so that $\|\hat{c}_i\|_1 = t_i$.

4.2 Some Implementation Details

4.2.1 Determination of $\hat{\mu}$, \hat{P} and \hat{c}_i^0 . Now, we start our numerical implementation from the calculation

of $\hat{\mu}$, \hat{P} and \hat{c}_i^0 , which is the optimal solution to a standard PCA problem (4) (without the consideration of those constraints). From some well known conclusions [1], we get that $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \bar{x}_i$ and \hat{P} is formed by the first l dominant eigenvectors of covariance matrix $R = (X - \hat{\mu})(X - \hat{\mu})'$ and $\hat{c}_i^0 = \hat{P}' x_i$. In our paper, we determine \hat{P} using the singular value decomposition(SVD) method which is much more efficient than directly calculating the eigenvectors of R .

4.2.2 Implementation of STE. The key issue of the soft-threshold estimation is to find out γ_i . Letting $(|\hat{c}_{si}^0| - \gamma_i)^+$ be the positive part of $(|\hat{c}_{si}^0| - \gamma_i)$, and $\|\hat{c}_i^0\|_1 = t_i^0$. We first note that

$$t_i^0 - t_i = \sum_{s=1}^l |\hat{c}_{si}^0| - \sum_{s=1}^l |\hat{c}_{si}| = \sum_{s=1}^l \left\{ |\hat{c}_{si}^0| - (|\hat{c}_{si}^0| - \gamma_i)^+ \right\}$$

$$= \sum_{s=1}^l |\hat{c}_{si}^0| I(|\hat{c}_{si}^0| \leq \gamma_i) + \gamma_i \sum_{s=1}^l I(|\hat{c}_{si}^0| > \gamma_i)$$

$$= \sum_{s=1}^{K_i} b_{si} + \gamma_i (l - K_i)$$

, where $b_{l_1} \leq \dots \leq b_{l_l}$ are the ordered values of $|\hat{c}_{li}^0|, \dots, |\hat{c}_{li}^0|$ and $K_i = \max\{s : b_{si} \leq \gamma_i\}$. Let $a_{0i} = 0$ and

$$a_{ki} = \sum_{s=1}^k b_{si} + b_{ki} (l - k), \quad k=1, \dots, l, \quad \text{so that}$$

$$0 = a_{0i} \leq a_{1i} \leq \dots \leq a_{li} = t_i^0 \quad \text{and it is easy to show that:}$$

$$0 \leq K_i = \max\{s : a_{si} \leq t_i^0 - t_i\}$$

$$, \text{ from this we may calculate } \gamma_i = \frac{(t_i^0 - t_i) - \sum_{s=1}^{K_i} b_{si}}{l - K_i}.$$

4.2.3 Projected Gradient Approach for RPCA. There are two kinds of commonly used numerical implementation for RPCA: iteratively reweighted least squares (IRLS) and gradient descent (GD). In this paper, we choose the GD method as our candidate algorithm. Here, the biggest challenge during the gradient decent process is to make sure that P remain an orthogonal matrix. This difficulty is solved by applying the projected gradient approach to the calculating of the partial derivation of $E_4(P, C, \mu, \sigma)$ respecting to P (see [16] for more detailed description). The learning “rules” for updating P , C and μ successively are given as follows:

$$P^{k+1} = P^k - \eta_l (P^k)' \left(\frac{\partial E_4}{\partial P} \right) - \left(\frac{\partial E_4}{\partial P} \right)' P^k$$

$$+ (\mathbf{I}_l - P^k (P^k)') \frac{\partial E_4}{\partial P}$$

$$C^{k+1} = C^k - \eta_2 \left(\frac{\partial E_4}{\partial C} \right), \quad \mu^{k+1} = \mu^k - \eta_3 \left(\frac{\partial E_4}{\partial \mu} \right)$$

, where

$$\frac{\partial E_4}{\partial P} = -WC', \quad \frac{\partial E_4}{\partial C} = -P'W, \quad \frac{\partial E_4}{\partial \mu} = -W\mathbf{1}_n, \quad W \in R^{m \times n} \text{ and}$$

$$W_{ji} = \rho'(\bar{x}_{ji} - \bar{\mu}_j - \sum_{s=1}^l \bar{p}_{js} \bar{c}_{si}), j=1, \dots, m; \quad i=1, \dots, n.$$

η_1, η_2, η_3 are step size, they can be some constants or estimated by the local quadratic approximation, as described in [5]. Here, we choose the Geman-McClure error

function $\rho(x, \sigma) = \frac{x^2}{x^2 + \sigma^2}$ as the robust function, where

“scale” parameter σ controls the convexity of the function and can be determined through Median Absolute Deviation (MAD) [4]. The initial value for P^0, C^0 and μ_0 can be chosen via SVD method satisfying that $(P^0)'P^0 = \mathbf{I}_l$.

5. Experimental Results

In this section, we will illustrate the application of M-LASSO-PCA and RPCA to extract outliers in a set of sample images and reconstruct the images robustly. First we introduce the concept of SOC (sample outlier coefficient) that will be used in our experiment.

Definition Let α_i be the count of the zero components of $\hat{\tilde{c}}_i$, then we define the proportion $\theta_i = \frac{\alpha_i}{\dim(\hat{\tilde{c}}_i)}$ as

the sample outlier coefficient (SOC).

From the process of soft-threshold estimation we know that $\hat{\tilde{c}}_i$ is determined by the model-selection parameter t_i , thus θ_i is a function of t_i . Generally, t_i is chosen to be less than and very close to t_i^0 , as an alternative we

may fix the proportion $\frac{t_i}{t_i^0} = \mu < 1, i=1, \dots, n$. In this

situation, we also write the determined SOC θ_i as $\theta_i(\mu)$. We may compare the SOC with some threshold value T to determine whether an image is a sample outlier or not. In our experiment, we choose $T = 0.5$.

Figure 2(a) shows a collection of 9 images (128×128, gray level) from the Columbia object image library (<http://www1.cs.columbia.edu/CAVE/research/softli b/>), among which seven are belonging to the duck class and the rest two are belonging to the coke-tin class. In common sense the two coke-tin images are viewed to be sample outliers. Meanwhile, the first duck image is also contaminated by some obvious intra-sample outliers.

Fig.2 (b) shows the 9 (here we choose $l = n = 9$) basis images obtained from standard PCA. Table 3 lists the projected coefficients $\hat{\tilde{c}}_i$ ($i=1, \dots, n$) estimated by M-LASSO-PCA and their corresponding θ_i . Through the comparison of θ_i with T , we can automatically recognize that the 4th and the 9th image are sample outliers, just as we can see through our eyes (Fig.2(c)).

Fig.3 (a) shows the original images with the sample outliers eliminated. Fig. 3(b) ~ (d) illustrate the reconstruction results of frame 1 using standard PCA and RPCA(by minimizing (3)) separately. The difference between the RPCA results (c) and (d) lies in that the former uses the fig. 2 (a) as training set while the latter uses fig.3 (a) as training set. For our experiment 200 times of iterations is enough to guarantee the convergence of the learning rules in 4.2.3. It is apparent to see, from the three result images, that M-LASSO_PCA+RPCA is much more robust to outliers than the standard PCA or pure RPCA.

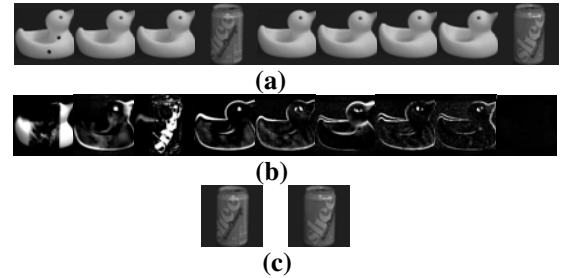


Fig. 2 (a) Original images (b) Basis images by n PCA (c) Outlier images.

Table3 Project coefficients estimated by M-LASSO-PCA

$\hat{\tilde{c}}_i$	0.129	0.141	0.147	-0.475	0.148	0.142	0.128	0.112	-0.471
	0.149	0.108	0.028	0	-0.008	-0.039	-0.110	-0.127	0.000
	-0.004	-0.006	0	-0.119	0	0	0.003	0.002	0.121
	0.0666	0.001	-0.054	0	-0.063	-0.047	0.030	0.066	0
	0.027	-0.032	-0.031	0	0.002	0.048	-0.001	-0.002	0
	0.017	-0.041	0.024	0	0.013	-0.027	0.016	-0.009	0
	0	-0.009	0.008	0	0.001	0	-0.033	0.026	0
	0	0	-0.014	0	0.024	-0.007	0	0	0
	0	0	0	0	0	0	0	0	0
$\ \hat{\tilde{c}}_i\ _1$	0.396	0.341	0.309	0.595	0.264	0.312	0.328	0.352	0.592
$\ \hat{\tilde{c}}_i^0\ _1$	0.440	0.379	0.343	0.661	0.293	0.346	0.365	0.392	0.658
θ_i	0.33	0.22	0.22	0.78	0.22	0.33	0.22	0.22	0.67

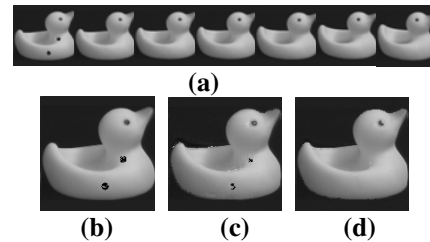


Fig. 3 (a) Duck image set with sample outliers elimi-

nated; (b) Reconstructed image (frame 1) using standard PCA on set (a) in fig. 3; (c) Reconstructed image (frame 1) from RPCA on set (a) in fig. 2; (d) Reconstruction image (frame1) from RPCA on image set (a) in fig. 3.

In this image reconstruction experiment, there are 2 sample outliers out of 9 sample images. A natural extension of our experiment is to increase the number of sample outliers and observe how the SOC changes. Here, we still use image set formed by 9 samples from toy duck class and coke-tin class and let the number of coke-tin samples increase from 1 to 8, or equally saying the coke-tin Class Proportion (CP) increase from 0.11 to 0.89. For the purpose of convenience, we make frame 9 always be the coke-tin image. Figure 4 is the function graph of $\theta_9(0.9)$ taking coke-tin CP as variance. From this figure we can see when the coke-tin CP is less than 0.5, which means coke-tin images are sample outliers, the corresponding $\theta_9(0.9)$ is larger than $T = 0.5$ and vice versa. This promising experimental result shows that our M-LASSO-PCA is a general method for the detection of sample outliers in a given image set.

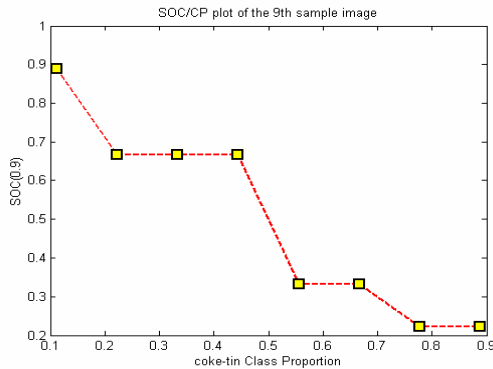


Fig. 4 Plot of the relationship between $\theta_9(0.9)$ and coke-tin Class Proportion.

6. Conclusion

In this paper, we address the problem of robust image reconstruction and present an algorithm that can efficiently overwhelm the influence of both the sample outliers and the intra-sample outliers to the reconstruction results. To our point of view, in the common case where both kinds of outliers co-exist, it is more reliable to deal with them separately, as we have done through LASSO and robust M-estimation, instead of using some single algorithm such as RPCA. Our two-step algorithm works well in real-word experiment and the reconstruction results are quite satisfying.

Some further research work will be carried out to apply our image reconstruction algorithm to the image matching problems such as human face recognition.

7 Acknowledgement

This research is partly supported by the National High Technology Research and Development Program of China, under grant 2002AA145090.

Reference

- [1] Jolliffe, I.T. 1986. Principal Component Analysis. Springer-Verlag:NewYork.
- [2] Black, M.J., Yaccob, Y., Jepson, A., and Fleet, D.J. 1997. Learning parameterized models of image motion. In Conference on Computer Vision and Pattern Recognition, pp. 561-567.
- [3] De la Torre, F. and Black, M.J. 2002. Robust parameterized component analysis: Theory and applications to 2d facial modeling. In European Conference on Computer Vision, pp. 653-669.
- [4] Black, M.J. and Jepson, A. 1998. EignTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation. International Journal of Computer Vision, 26(1), pp.63-84.
- [5] De la Torre, F. and Black, M.J. 2003. A Framework for Robust Subspace Learning. International Journal of Computer Vision, 54(1/2), 117-142.
- [6] Moghaddam B. and Pentland A. 1997. Probabilistic visual learning for object representation. IEEE tran. Pattern Analysis and Machine Intelligence, 19(7): 137-143.
- [7] Rousseeuw, P.J. and Leroy, A.M. 1987. Robust Regression and Outlier Detection. John Wiley and Sons.
- [8] Xu, L. and Yuille, A. 1995. Robust principal component analysis by self-organizing rules based on statistical physics approach. IEEE tran. Neural Networks. 6(1): 131-143.
- [9] Cadima, J. and Jolliffe, I.T. 1995. Loading and Correlations in the Interpretation of Principal Components. Journal of Applied Statistics, 22, 203-214.
- [10] Vines, S.K. 2000. Simple principal components. Applied Statistics, 49(4), pp. 441-451.
- [11] Jolliffe, I.T., Trendafilov, N.T. and Uddin M. 2003. A Modified Principal Component Technique Based on the LASSO. Journal of Computational and Graphical Statistics, Vol. 12, No.3, pp.531-547.
- [12] Tishirani, R. 1996, Regression Shrinkage and Selection via the Lasso, Journal of the Royal Statistical Society(Series B), 58, 267-288.
- [13] Roweis, S. 1997. EM algorithms for PCA and SPCA. In Neural Information Processing Systems, pp. 626-632.
- [14] Black, M.J. and Rangarajan, A. 1996. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. International Journal of Computer Vision, 25(19), pp.57-92.
- [15] Osborne M., Presnell B. and Turlach B. 2000. On the LASSO and its dual. Journal of Computational and Graphical Statistics, Vol. 9, pp. 319-337.
- [16] Chu M.T., Trendafilov N.T. 2001. The Orthogonally Constrained Regression Revisited. Journal of Computational and Graphical Statistics, Vol. 10, No.4, pp 746-771.