An Adaptive Regularization Method for Image Denoising

Hongyi Wang
Changzhou Vocational Institute of Engineering, Changzhou, Jiangsu, China

ABSTRACT: This paper proposes a new regularization method to solve the problem of low rank optimization. Approximation compared with the nuclear norm, which is a rank regularization of successive approximation. The advantage of this model is that it can directly solve the rank of the regularized problem, and it only needs to calculate a singular value decomposition to improve the efficiency. Based on Morozov discrepancy principle, this paper analyses the choosing standard of parameters and proposes an adaptive algorithm. Finally, a series of experiments were done to illustrate this algorithm. Experimental results show that the proposed algorithm is compared with other algorithms, and the algorithm not only improves the denoising effect, but also shortens the time of operation.

Keywords: regularization problem; low rank; singular value decomposition; Morozov discrepancy principle

1 INTRODUCTION

In recent years, the low rank subspace estimation and segmentation have attracted the attention and researches of many researchers, such as machine learning, computer vision, statistical analysis, signal and image processing, etc.\cite{1,2,3} In those areas, many observational data, such as video surveillance, image, text and web data have a very high dimensionality leading to the "dimension disaster", which also makes reasoning, learning and recognition tasks incomplete. Although these data are in high dimensional space, its intrinsic dimensionality is generally low, and the data sample points are distributed in a low dimensional structure\cite{4}.

In high dimensional data analysis and subspace method such as principal component analysis (PCA) and linear discriminant analysis (LDA) and non-negative matrix factorization (NMF) etc. is very common. This is mainly because the subspace method is relatively simple, easy to implement and effective in the practical problems. In addition, the method can be used to transform the Hilbert space, and extended to deal with nonlinear problems. Therefore, a large number of subspace methods are proposed, which can separate the data samples into their respective subspaces and model each cluster to obtain their own low dimensional subspace. This method is called subspace segmentation, which has been successfully applied to many practical problems, such as motion segmentation, face clustering, image segmentation, image representation and compression, and hybrid system identification.

For subspace estimation and segmentation methods challenges\cite{5} are observed images or information usually polluted by noise or singular points, and even missing data. In order to solve these problems, many methods based on compressed sensing theory and rank minimization algorithm are proposed. In essence, those algorithms are required to minimize a non-convex optimization problem, namely, the norm and rank function on the minimization problem. Unfortunately, due to the nature of the $l_0$ norm and rank function, the optimization problem of \cite{6} them is NP difficult. In order to effectively solve the problem, the common method is to combine the $l_0$ norm of the mixed optimization problem with the rank function.

The number of partially relaxed their hull form, namely the norm and matrix nuclear norm (also known as Mark van $l_0$ norm trace) can get a convex optimization problem \cite{7}. It is well known that the two convex hull form: norm and matrix nuclear norm have strong ability were induced by the sparse and low rank structure \cite{8}.

*Corresponding author: why_math@126.com
2 ALGORITHM PROPOSED

In this paper, we mainly deal with the problem of periodic texture image denoising, that is, we live in a common symmetrical image, such as wall, organizational structure, board and so on. (See Figure 1)

![Figure 1. Image with periodic texture.](image)

For this kind of image denoising problem, the rank of the image matrix is a very good method for the regularization term:

$$\min_{u} \psi(u) = \|u - u_0\|_2^2 + \text{rank}(u)$$  \hspace{1cm} (1)

$u$ is the ideal image, and $u_0$ is the observed image. The pixel points of the image can be regarded as a matrix, and the rank of this matrix is difficult to calculate, so (1) in the optimization problem is difficult to directly determine the results. If the nuclear norm is relaxation of rank, under certain conditions, this relaxation can be a good approximation of the original image of convex rank. Then (1) the model can be transformed into the following model:

$$\min_{u} \psi(u) = \|u - u_0\|_2^2 + \lambda \|\sigma\|_1$$

Many optimization algorithms of this problem have been studied, and the iterative singular value decomposition (ISVT) \cite{9} is one of them, but the iterative singular value decomposition algorithm in each iteration step needs the singular value decomposition. This increases the computational complexity and running time of the program. In order to reduce the computational complexity, this paper proposes a new algorithm to avoid the problems mentioned above, improve the efficiency, and achieve the effect of denoising:

First of all, for the matrix $A$, the singular value decomposition, that is $A = SVD = \sum_{i=1}^{n} \sigma_i v_i^T$, $D$, $S$ is an orthogonal matrix, that is $SS^T = DD^T = I_n$, the matrix $S$ and $D$ respectively contains the matrix of the left and right singular vector. The $V$ is a diagonal matrix consisting of the singular value of the matrix $A$ and contains all the singular values $\sigma$, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$.

They can be expressed as follows:

$$S = (s_1, s_2, \ldots, s_n), \quad D = (d_1, d_2, \ldots, d_n), \quad V = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

Then we construct a characteristic function $p(\sigma)$ for each singular value $\sigma$, and the characteristic function satisfies $p(\sigma) = \begin{cases} 0 & \text{if } \sigma = 0 \\ 1 & \text{else} \end{cases}$, then $\sum_i p(\sigma_i)$ is the rank of the image $u$. And generally we choose the form of Gauss function $p(\sigma)$, that is $p(\sigma) = 1 - \exp(-\frac{\sigma^2}{s^2})$, in which the parameters of the $s$ Gauss function of the standard deviation parameters, controlling the speed of approaching 0. As shown in Figure 2, when $s \to 0$, the function $p(\cdot)$ is a good approximation of the characteristic function.

![Figure 2. Curve of characteristic function ($\sigma$).](image)

Finally, the rank of the image can be solved by using the characteristic function of each singular value, that is $\text{rank}(A) = \sum_i p(\sigma_i)$, so the denoising model (1) can be expressed as follows:

$$\min_{\sigma} \psi(\sigma) = \|S \cdot V - D - u_0\|_2^2 + \lambda \sum_i f(\sigma_i)$$ \hspace{1cm} (2)

Among them $V = \text{diag}(\sigma)$, the regularization parameter. The final denoising image is: $u = S \cdot V(\sigma) \cdot D$.

Compared with the kernel norm, the model (2) has a new relaxation term for rank regularization. Compared with ISVT, the advantage is that it only needs to compute the singular value decomposition in the iteration process.

According to the properties of the orthogonal matrix, the model (2) can be written as follows:

$$\min_{\sigma} \left\| \text{SVD} - u_0 \right\|_2^2 + \lambda \sum_i f(\sigma_i)$$

$$\Leftrightarrow \min_{\sigma} \left\| V - S' u_0 D^T \right\|_2^2 + \lambda \sum_i f(\sigma_i)$$

For discrete form can be expressed as:

$$\min_{i,j} (\sigma_i - (S' u_0 D^T)_{i,j})^2 + \lambda \sum_i f(\sigma_i)$$ \hspace{1cm} (3)
Here for the minimum value $\sigma$, that is $\sigma(3)$ on the derivative of zero, the following expressions are obtained:

$$
\sigma_i^j + \frac{\lambda}{s^2}\sigma_i \exp\left(-\frac{\sigma_i^j}{s^2}\right) = \sum_{i,j} (S^Tu_dD^T)_{i,j}
$$

On (4) can be derived from $\sigma$, the calculation, and can effectively solve this problem through the Newton iterative method.

3 ALGORITHM ANALYSIS

Algorithm 1

1. Initial value, $[S_0, V_0, D_0] = \text{SVD}(u_0), \lambda, s$
2. Iteration

$$
(1) \sigma_{new} = \arg \min_{\sigma} \|S \ast V \ast D - u_0\|^2 + \lambda \sum_i f(\sigma_i), \text{ by (4)}
$$

solving;

(2) To update $u_{new} = S \ast V(\sigma_{new}) \ast D$;

(3) If $|u_{new} - u_0| < \epsilon$, stop;

(4) Else $u_0 = u_{new}$, return (1).

In algorithm 1, two Parameters $\lambda$ and $S$ need to assign the initial value, and how to choose the two initial values is an important and difficult work. Then we discuss $\lambda$ and $S$ the choice of parameters and standards.

For the regularization parameter $\lambda$ choice is very difficult, if weight parameter $\lambda$ is too small, algorithm focuses on the approximation term and isn’t stable; if the parameter is too large, algorithm focuses on the regularization term, the result is inaccurate to the original solution. Based on the principle of Morozov deviation, the parameter $\lambda$ of the standard are proposed.

Because the final optimal solution is dependent on the choice of parameters, so the problem (2) can be written as follows:

$$
\min_{\sigma} J(\sigma, \lambda) = \|S \cdot \text{diag}(\sigma^j) \cdot D - u_0\|^2 + \lambda \sum_i p(\sigma_i^j)
$$

Solving the Euler equation, as well as

$$
\sigma_i^j + \lambda p(\sigma_i^j)' = b
$$

solving, i.e., $$(1 + \lambda p(\sigma_i^j)') \frac{d\sigma_i^j}{d\lambda} + p(\sigma_i^j)' = 0 \tag{5}$$

among

$$
b = \sum_{i,j} (S^Tu_d)_{i,j}d_{i,j}, \quad p(\sigma_i^j)' = \frac{2\sigma_i^j}{s^2} \exp\left(-\frac{\sigma_i^j}{s^2}\right),$$

$$
p(\sigma_i^j)' = \frac{2}{s^2} \exp\left(-\frac{\sigma_i^j}{s^2}\right) - \frac{4\sigma_i^j}{s^4} \exp\left(-\frac{\sigma_i^j}{s^2}\right)
$$

For the right $\sigma$, the parameters $\lambda$ can be selected by the Morozov deviation principle.

According to the Morozov deviation principle, the function $\phi(\lambda)$ is defined.

$$
\phi(\lambda) = \|S \cdot \text{diag}(\sigma^j) \cdot D - u_0\|^2 - \delta^2
$$

Solving $\phi(\lambda) = 0$, using Newton iterative method

$$
\lambda_{k+1} = \lambda_k - \frac{\phi(\lambda_k)}{\phi'(\lambda_k)}
$$

And satisfy, $\frac{d\phi(\lambda)}{d\lambda} = 0$ according to (5)

$$
\phi_k = \sigma_i \frac{d\sigma_i}{d\lambda} = -\sigma_i(1 + \lambda p(\sigma_i^j))^{-1} p(\sigma_i^j)'
$$

We can be seen $p(\sigma)$ as a convex function, especially when the parameters $s \to 0$, you can get the global optimal solution. But for smaller and smaller parameters $S, s_0 > s_1 > \ldots > s_n$, this is a smooth optimization problem, and the global solution can be solved.

Algorithm 2 on parameters and $s$ adaptive algorithm

1. Initial value, $[S_0, V_0, D_0] = \text{SVD}(u_0), \lambda, s_0$
2. Iteration

$$
(1) \sigma_{new} = \arg \min_{\sigma} \|S \ast V(\sigma_{new}) \ast D - u_0\|^2 + \lambda \sum_i f(\sigma_i), \text{ by (4)}
$$

solving;

(2) According to the Morozov deviation principle, the parameters are updated by (6);

$$
(3) u_{new} = S \ast V(\sigma_{new}) \ast D;
$$

(4) If $|u_{new} - u_0| < \epsilon$, stop;

(5) Else $u_0 = u_{new}, s = 1/2 \ast s_0$, return (1).

Algorithm 3 singular value iteration algorithm [9]

Select a drop sequence

iteration $[s, v, d] = \text{SVD}(u)$;

$$
u = s \ast \max(v - 1/\lambda \ast \text{eye(size(v))), 0} \ast d'.
$$

4 EXPERIMENTAL ANALYSIS AND RESULTS

In order to verify the effectiveness of the algorithm, it is mainly aimed at some of the periodic texture of the image (such as wall, tissue, board, etc.) to denoise.

The first test image is modified for size $250 \times 250$, parameters $\lambda$ and $S$ initial values were taken for 0.5 and 1, the 2 algorithms can effectively solve the problem of image denoising. This paper will put forward the adaptive regularization rank smooth approximation algorithm with the nuclear norm proposed in singular value decomposition algorithm to compare and rank regularization algorithm.
Experimental results as shown in Figure 3 and 4, using the peak signal to noise ratio (PSNR) as the objective evaluation of the experimental results, the greater the peak signal to noise ratio, the better the effect of denoising, as shown in Table 1. Compared with singular value decomposition, rank regularization algorithm not only improves the denoising effect, but also the efficiency of the algorithm.

Table 1. Signal to noise ratio (SNR) of different algorithms

<table>
<thead>
<tr>
<th></th>
<th>ISVT</th>
<th>Rank regularization algorithm</th>
<th>Adaptive algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR (wall image)</td>
<td>21.5949</td>
<td>22.0663</td>
<td>22.6292</td>
</tr>
<tr>
<td>PSNR (checkerboard image)</td>
<td>18.1243</td>
<td>22.1243</td>
<td>22.4524</td>
</tr>
</tbody>
</table>

Adaptive algorithm in rank regularization algorithm based on the Morozov principle proper choice of parameters $\lambda$, weights the approximation term and a regularization term. For the image size of the image $250 \times 250$, the adaptive algorithm is less than the time required for the singular value algorithm.

The comparison results of the peak signal to noise ratio and the time required for the periodic image wall under different picture size and iteration number are given, as shown in Table 2 and Table 3.

Table 2. Three different algorithms in image size not at the same time, the signal to noise ratio and the time required.

<table>
<thead>
<tr>
<th>Picture size</th>
<th>ISVT</th>
<th>Rank regularization algorithm</th>
<th>Adaptive algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>200×200</td>
<td>20.0756</td>
<td>21.0150</td>
<td>21.5668</td>
</tr>
<tr>
<td>250×250</td>
<td>21.5949</td>
<td>22.0663</td>
<td>22.6292</td>
</tr>
<tr>
<td>500×500</td>
<td>20.9257</td>
<td>21.9678</td>
<td>23.1731</td>
</tr>
</tbody>
</table>

Table 3. The signal to noise ratio and the time required for the three algorithms at different times of iteration.

<table>
<thead>
<tr>
<th>Iterations</th>
<th>ISVT</th>
<th>Rank regularization algorithm</th>
<th>Adaptive algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>20.6395</td>
<td>21.0561</td>
<td>21.9956</td>
</tr>
<tr>
<td>100</td>
<td>20.8106</td>
<td>21.3672</td>
<td>21.5246</td>
</tr>
<tr>
<td>200</td>
<td>21.3985</td>
<td>21.4665</td>
<td>22.1151</td>
</tr>
<tr>
<td>500</td>
<td>15.8331</td>
<td>21.5725</td>
<td>22.1484</td>
</tr>
</tbody>
</table>

Through the Table 2 and 3, we found that as the image size increases and the number of iterations change much, singular value algorithm required time increase significantly, and rank regularization algorithm not only effectively to noise at the same time also shortens the required time.

5 SUMMARY

This paper is mainly to introduce a kind of optimal regularization of rank new smooth approximation problem, deal with periodic texture image to the nuclear norm to approximate rank regular problem. The advantage of this model is that it can directly solve the rank of the regularized problem, and it only needs to calculate a singular value decomposition to improve the efficiency. The results show the algorithm not only improves the denoising effect, and shorten the time of operation. In order to improve the effect of denoising, the Morozov principle is used to analyze the choice of the weighing coefficient between the approximation term and the regularization term.

REFERENCES


