Linear Total Variation Approximate Regularized Nuclear Norm Optimization for Matrix Completion

Xu Han, Jiasong Wu, Lu Wang, Yang Chen, Lotfi Senhadji, Huazhong Shu

▶ **To cite this version:**

Xu Han, Jiasong Wu, Lu Wang, Yang Chen, Lotfi Senhadji, et al.. Linear Total Variation Approximate Regularized Nuclear Norm Optimization for Matrix Completion. Abstract and Applied Analysis, 2014, 2014, pp.765782. 10.1155/2014/765782 . inserm-01101215

HAL Id: inserm-01101215

https://inserm.hal.science/inserm-01101215

Submitted on 26 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 765782, 8 pages
http://dx.doi.org/10.1155/2014/765782

Research Article

Linear Total Variation Approximate Regularized Nuclear Norm Optimization for Matrix Completion

Xu Han,[1,2] Jiasong Wu,[1,2,3,4] Lu Wang,[2,3,4] Yang Chen,[1,2,3] Lotfi Senhadji,[2,3,4] and Huazhong Shu[1,2]

[1] *Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China*
[2] *Centre de Recherche en Information Médicale Sino-français (CRIBs), France*
[3] *INSERM, U1099, Rennes 35000, France*
[4] *Université de Rennes 1, LTSI, Rennes 35042, France*

Correspondence should be addressed to Xu Han; xuhan@seu.edu.cn

Received 15 February 2014; Accepted 7 May 2014; Published 28 May 2014

Academic Editor: Zhiwu Liao

Copyright © 2014 Xu Han et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Matrix completion that estimates missing values in visual data is an important topic in computer vision. Most of the recent studies focused on the low rank matrix approximation via the nuclear norm. However, the visual data, such as images, is rich in texture which may not be well approximated by low rank constraint. In this paper, we propose a novel matrix completion method, which combines the nuclear norm with the local geometric regularizer to solve the problem of matrix completion for redundant texture images. And in this paper we mainly consider one of the most commonly graph regularized parameters: the total variation norm which is a widely used measure for enforcing intensity continuity and recovering a piecewise smooth image. The experimental results show that the encouraging results can be obtained by the proposed method on real texture images compared to the state-of-the-art methods.

1. Introduction

The problem of matrix completion, which can be seen as the extension of recently developed compressed sensing (CS) theory [1–3], plays an important role in the field of signal and image processing [4–11]. This problem occurs in many real applications in computer vision and pattern recognition, such as image inpainting [12, 13], video denoising [14], and recommender systems [15, 16]. Reconstruction algorithms for matrix completion have received much attention. Cai et al. [17] proposed an algorithm, namely, the singular value thresholding (SVT) algorithm for matrix completion and related nuclear norm minimization problems. In [18], a simple and fast singular value projection (SVP) algorithm for rank minimization with affine constraints is exploited. Keshavan et al. [19] dealt with the matrix completion based on singular value decomposition followed by local manifold optimization. In order to achieve a better approximation of the rank of matrix, Hu et al. [11] presented an approach based on the truncated nuclear norm regularization (TNNR), which is defined by the difference between the nuclear norm and the sum of the largest few singular values. Since most of the existing matrix completion models aim to solve the low rank optimization via nuclear norm, we recall here this model. For an incomplete matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ of rank r, the model can be described as follows:

$$\min_{\mathbf{X}} \operatorname{rank}(\mathbf{X}) \quad \text{s.t. } \mathbf{X}_\Omega = \mathbf{M}_\Omega, \tag{1}$$

where $\mathbf{X} \in \mathbb{R}^{m\times n}$ and $\mathbf{M}_\Omega = \mathbf{M}_{ij}$, $(i, j) \in \Omega$, and Ω is the set of locations corresponding to the observed entries.

Unfortunately, the rank minimization problem in (1) is an NP-hard one, so the following convex relaxation is widely used:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{s.t. } \mathbf{X}_\Omega = \mathbf{M}_\Omega, \tag{2}$$

where $\|\cdot\|_*$ is the nuclear norm given by

$$\|\mathbf{X}\|_* = \sum_{k=1}^{\min(m,n)} \sigma_k, \tag{3}$$

where σ_k denotes the kth largest singular value of $\mathbf{X}$.

In this paper, our objective is to exploit the intrinsic geometry of the data distribution and incorporate it as an additional regularization term to deal with the images which are rich in texture. The total variation (TV) norm has demonstrated its usefulness as a graph regularizer in the field of image processing, so we propose here a method that combines the nuclear norm with the linear TV approximate norm to solve the problem of matrix completion. We call it the linear total variation approximate regularized nuclear norm (LTVNN) minimization problem. This combination optimization problem will be solved by simple and efficient optimization scheme based on the alternating direction method of multipliers (ADMM) model [20, 21].

The paper is organized as follows. In the next section, we introduce the proposed LTVNN model and we describe the optimization schemes. In Section 3, we establish the convergence results for the iterations given in Section 2. Experimental results on a set of images are provided in Section 4. Finally, we draw some conclusions in Section 5.

2. Proposed Method

2.1. Some Preliminaries. The total variation along the vertical and horizontal directions can be described as

$$D^v_{j,k}(\mathbf{X}) = \begin{cases} \mathbf{X}_{j,k} - \mathbf{X}_{j+1,k}, & 1 \le j < m \\ 0, & j = m, \end{cases} \tag{4}$$

$$D^h_{j,k}(\mathbf{X}) = \begin{cases} \mathbf{X}_{j,k} - \mathbf{X}_{j,k+1}, & 1 \le k < n \\ 0, & k = n. \end{cases} \tag{5}$$

So the total variation of $\mathbf{X}$ is the summation for the magnitude of the gradient of each pixel [22]:

$$\|\mathbf{X}\|_{\mathrm{TV}} = \sum_{j,k} \sqrt{\left(D^v_{j,k}\mathbf{X}\right)^2 + \left(D^h_{j,k}\mathbf{X}\right)^2}. \tag{6}$$

And the equvalent total variation formula as follows:

$$\|\mathbf{X}\|_{\mathrm{TV}} = \sum_{j,k} \left(\left|D^v_{j,k}\mathbf{X}\right| + \left|D^h_{j,k}\mathbf{X}\right|\right). \tag{7}$$

Here, we use the linear total variation approximate of (7) to approximate the second kind of total variation; that is,

$$\|\mathbf{X}\|_{\mathrm{LTVA}} = \sum_{j,k} \left(\left(D^v_{j,k}\mathbf{X}\right)^2 + \left(D^h_{j,k}\mathbf{X}\right)^2\right). \tag{8}$$

2.2. Proposed Model. As mentioned above, the key point of the proposed approach is the combination of the nuclear norm and the linear total variation approximate norm; therefore, the optimization problem is described as

$$\min_{\mathbf{X}} (1-\gamma)\|\mathbf{X}\|_* + \gamma\|\mathbf{X}\|_{\mathrm{LTVA}} \quad \text{s.t. } \mathbf{X}_\Omega = \mathbf{M}_\Omega, \tag{9}$$

where $0 \le \gamma \le 1$ is a penalty parameter, $\|\mathbf{X}\|_*$ is the nuclear norm defined in (3), and $\|\mathbf{X}\|_{\mathrm{LTVA}}$ is linear total variation norm approximate defined in (8), which can be reformulated as

$$\begin{aligned} \|\mathbf{X}\|_{\mathrm{LTVA}} &= \mathrm{Tr}\left[(\mathbf{X} - \mathbf{X}\phi_1)(\mathbf{X} - \mathbf{X}\phi_1)^T\right] \\ &\quad + \mathrm{Tr}\left[(\mathbf{X} - \phi_2\mathbf{X})(\mathbf{X} - \phi_2\mathbf{X})^T\right] \\ &= \|(\mathbf{X} - \mathbf{X}\phi_1)\|_F^2 + \|(\mathbf{X} - \phi_2\mathbf{X})\|_F^2, \end{aligned} \tag{10}$$

where "Tr" means the trace of the matrix, $\|\cdot\|_F$ denotes the Frobenius norm of the matrix, and ϕ_1 and ϕ_2 are, respectively, the column and row transform matrix given by

$$\begin{aligned} \phi_1 &= \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \underbrace{0 \quad 0 \quad \cdots \quad 1}_{(n-1)\times(n-1)} & & & & 1 \end{bmatrix}_{n\times n}, \\ \phi_2 &= \begin{bmatrix} 0 & \overbrace{1 \quad 0 \quad \cdots \quad 0}^{(m-1)\times(m-1)} \\ 0 & 0 \quad 1 \quad 0 \quad 0 \\ 0 & \vdots \quad \vdots \quad \ddots \quad \vdots \\ \vdots & 0 \quad 0 \quad \cdots \quad 1 \\ 0 & 0 \quad 0 \quad \cdots \quad 1 \end{bmatrix}_{m\times m}. \end{aligned} \tag{11}$$

So, the problem in (9) can be rewritten as

$$\begin{aligned} &\min_{\mathbf{X}} (1-\gamma)\|\mathbf{X}\|_* + \gamma\|(\mathbf{X} - \mathbf{X}\phi_1)\|_F^2 \\ &\quad + \gamma\|(\mathbf{X} - \phi_2\mathbf{X})\|_F^2 \\ &\text{s.t. } \mathbf{X}_\Omega = \mathbf{M}_\Omega. \end{aligned} \tag{12}$$

2.3. The Optimization Scheme. The alternating direction method of multipliers-ADMM [20, 21] is an efficient and scalable optimization model which exploits the structure of the optimization problem. In this section, we use ADMM to deal with the problem in (12), which can be reformulated as

$$\begin{aligned} &\min_{\mathbf{X},\mathbf{W}} (1-\gamma)\|\mathbf{X}\|_* + \gamma\|(\mathbf{W} - \mathbf{W}\phi_1)\|_F^2 \\ &\quad + \gamma\|(\mathbf{W} - \phi_2\mathbf{W})\|_F^2 \\ &\text{s.t. } \mathbf{X} = \mathbf{W}, \quad \mathbf{W}_\Omega = \mathbf{M}_\Omega, \end{aligned} \tag{13}$$

where $\|(\mathbf{W} - \mathbf{W}\phi_1)\|_F^2$ and $\|(\mathbf{W} - \phi_2\mathbf{W})\|_F^2$ are the indicator functions. The augmented Lagrange function of (13) is

$$\begin{aligned} \mathscr{L}(\mathbf{X}, \mathbf{Y}, \mathbf{W}, \lambda) &= (1-\gamma)\|\mathbf{X}\|_* + \gamma\|(\mathbf{W} - \mathbf{W}\phi_1)\|_F^2 \\ &\quad + \gamma\|(\mathbf{W} - \phi_2\mathbf{W})\|_F^2 + \frac{\lambda}{2}\|\mathbf{W} - \mathbf{X}\|_F^2 \\ &\quad + \mathrm{Tr}\left(\mathbf{Y}^T(\mathbf{W} - \mathbf{X})\right), \end{aligned} \tag{14}$$

where $\lambda > 0$ is the penalty parameter and $\mathbf{Y}$ is the multiplier. The solution can be obtained by incorporating the solutions

of each regularization problem separately which are defined as follows.

Row TV is as follows:

$$\begin{aligned}&\mathscr{L}_R(\mathbf{XR},\mathbf{YR},\mathbf{WR},\lambda)\\&=(1-\gamma)\|\mathbf{XR}\|_*+\gamma\left\|(\mathbf{WR}-\phi_2\mathbf{WR})\right\|_F^2\\&\quad+\frac{\lambda}{2}\|\mathbf{WR}-\mathbf{XR}\|_F^2+\mathrm{Tr}\left(\mathbf{YR}^T(\mathbf{WR}-\mathbf{XR})\right),\end{aligned}\tag{15}$$

where $\mathbf{XR}$ denotes the optimization result along the vertical direction of the total variation defined in (4).

Column TV is as follows:

$$\begin{aligned}&\mathscr{L}_C(\mathbf{XC},\mathbf{YC},\mathbf{WC},\lambda)\\&=(1-\gamma)\|\mathbf{XC}\|_*+\gamma\left\|(\mathbf{WC}-\mathbf{WC}\phi_1)\right\|_F^2\\&\quad+\frac{\lambda}{2}\|\mathbf{WC}-\mathbf{XC}\|_F^2+\mathrm{Tr}\left(\mathbf{YC}^T(\mathbf{WC}-\mathbf{XC})\right),\end{aligned}\tag{16}$$

where $\mathbf{XC}$ denotes the optimization result along the horizontal direction of the total variation defined in (5).

We deal with column linear TV optimization problem in (16) by the following steps in each iteration.

Step 1 (initial setting). $\mathbf{XC}_1=\mathbf{M}_\Omega$, $\mathbf{WC}_1=\mathbf{XC}_1$, $\mathbf{YC}_1=\mathbf{XC}_1$, with the tolerance ε.

Step 2 (computing $\mathbf{XC}_{k+1}$). Fix $\mathbf{WC}_k$ and $\mathbf{YC}_k$, and minimize (16) for obtaining $\mathbf{XC}_{k+1}$ as

$$\begin{aligned}\mathbf{XC}_{k+1}=\arg\min_{\mathbf{X}}&(1-\gamma)\|\mathbf{XC}\|_*+\gamma\left\|(\mathbf{WC}_k-\mathbf{WC}_k\phi_1)\right\|_F^2\\&+\frac{\lambda}{2}\left\|\mathbf{WC}-\mathbf{XC}_k\right\|_F^2+\mathrm{Tr}\left(\mathbf{YC}_k^T(\mathbf{WC}-\mathbf{XC}_k)\right).\end{aligned}\tag{17}$$

Ignoring the constant terms, (17) can be rewritten as

$$\begin{aligned}\mathbf{XC}_{k+1}=\arg\min_{\mathbf{X}}&(1-\gamma)\|\mathbf{XC}\|_*\\&+\frac{\lambda}{2}\left\|\mathbf{XC}-\left(\mathbf{WC}_k+\frac{1}{\lambda}\mathbf{YC}_k\right)\right\|_F^2.\end{aligned}\tag{18}$$

To solve (18), Cai et al. [17] introduce the soft-thresholding operator $\mathscr{D}_\tau$ which is defined as follows:

$$\begin{aligned}\mathscr{D}_\tau(\mathbf{X})&:=\mathbf{U}\mathscr{D}_\tau(\Sigma)\mathbf{V}^T,\\\mathscr{D}_\tau(\Sigma)&=\mathrm{diag}\left\{\max(\sigma_i-\tau)_+\right\},\end{aligned}\tag{19}$$

where $t_+=\max(0,t)$.

Using the operator $\mathscr{D}_\tau$ in (19), the solution of (18) can be obtained as

$$\mathbf{XC}_{k+1}=\mathscr{D}_{(1-\gamma)/\lambda}\left(\mathbf{WC}_k+\frac{1}{\lambda}\mathbf{YC}_k\right).\tag{20}$$

Step 3 (computing $\mathbf{WC}_{k+1}$). Fix $\mathbf{XC}_{k+1}$ and $\mathbf{YC}_k$ and calculate $\mathbf{WC}_{k+1}$ as follows:

$$\mathbf{WC}_{k+1}=\arg\min_{\mathbf{W}}\mathscr{L}\left(\mathbf{XC}_{k+1},\mathbf{YC}_k,\mathbf{WC},\lambda\right)\tag{21}$$

which is a quadratic function of $\mathbf{WC}$ and can be easily solved by setting the derivation of $\mathscr{L}(\mathbf{XC}_{k+1},\mathbf{YC}_k,\mathbf{WC},\lambda)$ to zeros, and then we get

$$\begin{aligned}\mathbf{WC}_{k+1}=&\left(\lambda\mathbf{XC}_{k+1}-\mathbf{YC}_k\right)\\&\times\left[2\gamma\left(\mathbf{I}-\phi_1-\phi_1^T+\phi_1^T\phi_1\right)+\lambda\mathbf{I}_{n\times n}\right]^{-1}.\end{aligned}\tag{22}$$

Then we fix the values at the observed entries:

$$\mathbf{WC}_{k+1}=\left(\mathbf{WC}_{k+1}\right)_{\Omega_m}+\mathbf{M}_\Omega,\tag{23}$$

where Ω_m denotes the set of the missing entries.

Step 4 (computing $\mathbf{YC}_{k+1}$). Fix $\mathbf{XC}_{k+1}$ and $\mathbf{WC}_{k+1}$ and calculate $\mathbf{YC}_{k+1}$ as follows:

$$\mathbf{YC}_{k+1}=\mathbf{YC}_k+\lambda\left(\mathbf{WC}_{k+1}-\mathbf{XC}_{k+1}\right).\tag{24}$$

Until the stop condition: $\|\mathbf{XC}_{k+1}-\mathbf{XC}_k\|_F\le\varepsilon$.

Row TV problem defined by (15) can be solved in a similar way to that of column TV problem. The only difference is the $\mathbf{WR}_{k+1}$ in the second step, which is given by

$$\begin{aligned}\mathbf{WR}_{k+1}=&\left[2\gamma\left(\mathbf{I}-\phi_2-\phi_2^T+\phi_2^T\phi_2\right)+\lambda\mathbf{I}_{m\times m}\right]^{-1}\\&\times\left(\lambda\mathbf{XR}_{k+1}-\mathbf{YR}_k\right).\end{aligned}\tag{25}$$

And the stop condition is $\|\mathbf{XR}_{k+1}-\mathbf{XR}_k\|_F\le\varepsilon$.

Finally, we obtained $\mathbf{X}_{k+1}$ as the average of $\mathbf{XC}_{k+1}$ and $\mathbf{XR}_{k+1}$; that is,

$$\mathbf{X}_{k+1}=\frac{\mathbf{XC}_{k+1}+\mathbf{XR}_{k+1}}{2}.\tag{26}$$

3. Convergence Analysis

In this section, we give the proof of the convergence of column total variation (16) and the convergence of row total variation is similar to the column total variation. Here, the objection function (16) about column variation is as follows:

$$\begin{aligned}&\min_{\mathbf{X}}f_\tau(\mathbf{X})\\&\text{s.t. }\mathbf{X}_\Omega=\mathbf{M}_\Omega\\&f_\tau(\mathbf{X})=\tau\|\mathbf{X}\|_*+\frac{1}{2}\mathrm{Tr}\left[(\mathbf{X}-\mathbf{X}\phi_1)(\mathbf{X}-\mathbf{X}\phi_1)^T\right],\\&\tau=\frac{1-\gamma}{2\gamma}.\end{aligned}\tag{27}$$

Lemma 1. *Let* $\mathbf{Z}\in\partial f_\tau(\mathbf{X})$ *and* $\mathbf{Z}'\in\partial f_\tau(\mathbf{X}')$. *Then*

$$\left\langle\mathbf{Z}-\mathbf{Z}',\mathbf{X}-\mathbf{X}'\right\rangle\ge\left\|\mathbf{X}-\mathbf{X}'\right\|_F^2.\tag{28}$$

The details of the proof can be found in [17].

(a) Original image (256 × 256) (b) Random masked image (c) $\gamma = 0$, PSNR: 8.934

(d) $\gamma = 0.5$, PSNR: 29.183 (e) $\gamma = 1$, PSNR: 8.593 (f) Word masked image

(g) $\gamma = 0$, PSNR: 14.706 (h) $\gamma = 0.5$, PSNR: 33.421 (i) $\gamma = 1$, PSNR: 14.759

FIGURE 1: The recovered results with 60% random mask and word mask for $\gamma = 0$, 0.5 and 1 by LTVNN.

Theorem 2. *Assuming that the sequence of step size obeys* $0 < \inf \lambda_k < \sup \lambda_k < (2\boldsymbol{\alpha}/\boldsymbol{\beta})$, $\boldsymbol{\alpha} = \langle (\mathbf{X}^k - \mathbf{X}^*)(\mathbf{I} - \phi_1 - \phi_1^T + \phi_1\phi_1^T), \mathbf{X}^k - \mathbf{X}^* \rangle$ *and* $\boldsymbol{\beta} = \|\mathbf{X}^k - \mathbf{X}^*\|_F^2$. *Here,* $\mathbf{X}^*$ *denotes the optimization result and* $\mathbf{X}^k$ *denotes the kth iteration object variable; then by the iteration procedure defined in Section 2.3, we can obtain the unique optimization result, that is,* $\mathbf{X}^*$. *And the details of the proof can be found in the Appendix.*

4. Experiments

In this section, we test the proposed method on a set of images. The algorithm was implemented with MATLAB programming language on a PC machine, which sets up Microsoft Windows 7 operating system and has an Intel Core I5 CPU with speed of 2.79 GHz and 2 GB RAM.

We deal with three channels (r, g, b) of color images separately and combine the results together to get the final outcome. We use peak signal-to-noise ratio (PSNR) values to evaluate the performance:

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{255^2}{\text{MSE}}\right), \tag{29}$$

where MSE denotes mean squared error,

$$\text{MSE} = \frac{1}{3mn}\left[\text{error}^2(r) + \text{error}^2(g) + \text{error}^2(b)\right]. \tag{30}$$

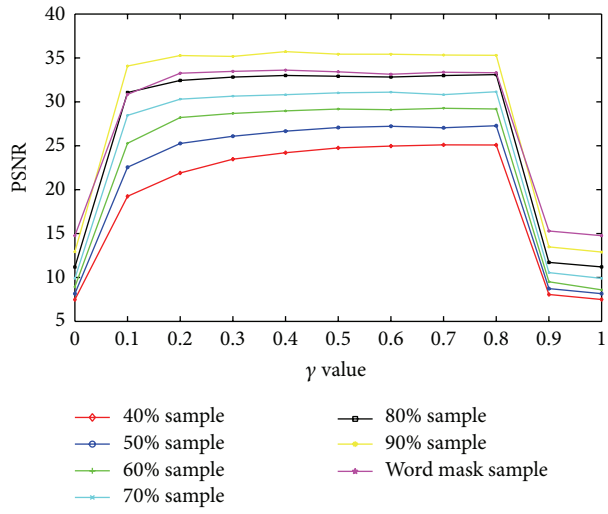


Figure 2: The recovered PSNR for Pepper under different random sample ratio and word mask sample with γ from 0 0 to 1 by LTVNN.

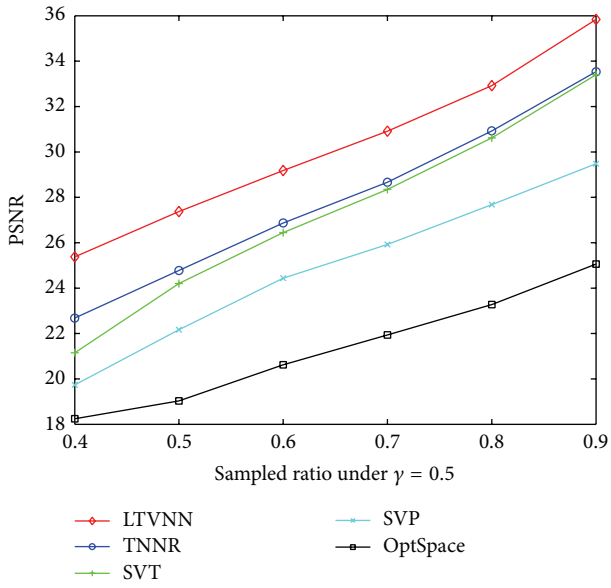


Figure 3: Recovered PSNR for Pepper under $\gamma = 0.5$ with different random sample ratio by LTVNN, TNNR, SVT, SVP, and OptSpace.

In the experiments, we consider two situations: random mask sample and word mask sample. Figure 1 describes the recovered results with 60% random mask and word mask for $\gamma = 0, 0.5$ and 1 by LTVNN. Figure 2 shows the recovered PSNR for Pepper under different random sample ratios and word mask sample for γ from 0 to 1 with step of 0.1 by LTVNN. It can be observed from these two figures that the best result is obtained for the value of γ near to 0.5, which corresponds to the case where the two norms (nuclear and LTV) are equivalently used in (9). For the two extreme cases: $\gamma = 0$ (only the nuclear norm is taken into consideration) and $\gamma = 1$ (only the linear total variation approximate norm is considered), the algorithm loses its efficiency.

We also compare our method (LTVNN) with other matrix completion methods including TNNR [10, 11], SVT [12], SVP [13], and OptSpace [14]. Figure 3 plots the recovered PSNR for Pepper for $\gamma = 0.5$ with different random sample ratios (from 40% to 90%) by LTVNN and other four methods (TNNR, SVT, SVP, and OptSpace). It can be seen from Figure 3 that the proposed LTVNN method achieves much higher PSNR than the other methods. Figure 4 shows the comparison of PSNR of recovered methods for Lena under word mask with $\gamma = 0.5$ by LTVNN and the other methods. Table 1 lists the PSNR results under word mask sample with $\gamma = 0.5$ for different images by LTVNN and the other methods. From Figure 4 and Table 1, we can see that the proposed method outperforms the other matrix completion methods under word mask for different images.

5. Conclusion

In this paper, we have proposed a new model that combines the nuclear norm and total variation norm for the matrix completion problem, which was then solved by ADMM model. Experimental results demonstrate the effectiveness of the proposed algorithm compared to other methods.

Appendix

Before we give the proof of Theorem 2, we supplement one proof about

$$\left\langle \left(\mathbf{X}-\mathbf{X}'\right)\left(\mathbf{I}-\phi_1-\phi_1^T+\phi_1\phi_1^T\right), \mathbf{X}-\mathbf{X}'\right\rangle \geq 0. \tag{A.1}$$

Without loss of generality, we take an example matrix $\xi = (\mathbf{X}-\mathbf{X}') \in \mathbb{R}^{4\times 4}$ and the corresponding transform matrix $(\mathbf{I}-\phi_1-\phi_1^T+\phi_1\phi_1^T) = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Then,

$$\begin{aligned}
&\mathrm{Tr}\left[\left(\mathbf{I}-\phi_1-\phi_1^T+\phi_1^T\phi_1\right)\xi^T\xi\right] \\
&\quad = 2\left(\xi_{1,1}^2+\xi_{2,1}^2+\xi_{3,1}^2+\xi_{4,1}^2\right) \\
&\qquad -\left(\xi_{1,1}\xi_{1,2}+\xi_{2,1}\xi_{2,2}+\xi_{3,1}\xi_{3,2}+\xi_{4,1}\xi_{4,2}\right) \\
&\qquad +2\left(\xi_{1,2}^2+\xi_{2,2}^2+\xi_{3,2}^2+\xi_{4,2}^2\right) \\
&\qquad -\left(\xi_{1,1}\xi_{1,2}+\xi_{2,1}\xi_{2,2}+\xi_{3,1}\xi_{3,2}+\xi_{4,1}\xi_{4,2}\right) \\
&\qquad -\left(\xi_{1,2}\xi_{1,3}+\xi_{2,2}\xi_{2,3}+\xi_{3,2}\xi_{3,3}+\xi_{4,2}\xi_{4,3}\right) \\
&\qquad +2\left(\xi_{1,3}^2+\xi_{2,3}^2+\xi_{3,3}^2+\xi_{4,3}^2\right) \\
&\qquad -\left(\xi_{1,2}\xi_{1,3}+\xi_{2,2}\xi_{2,3}+\xi_{3,2}\xi_{3,3}+\xi_{4,2}\xi_{4,3}\right) \\
&\quad = \left(\xi_{1,1}^2+\xi_{2,1}^2+\xi_{3,1}^2+\xi_{4,1}^2\right)+\left(\xi_{1,1}-\xi_{1,2}\right)^2 \\
&\qquad +\left(\xi_{2,1}-\xi_{2,2}\right)^2+\left(\xi_{3,1}-\xi_{3,2}\right)^2+\left(\xi_{4,1}-\xi_{4,2}\right)^2 \\
&\qquad +\left(\xi_{1,2}-\xi_{1,3}\right)^2+\left(\xi_{2,2}-\xi_{2,3}\right)^2+\left(\xi_{3,2}-\xi_{3,3}\right)^2 \\
&\qquad +\left(\xi_{4,2}-\xi_{4,3}\right)^2+\left(\xi_{1,3}^2+\xi_{2,3}^2+\xi_{3,3}^2+\xi_{4,3}^2\right) \\
&\quad \geq 0,
\end{aligned} \tag{A.2}$$

(a) Original image (256 × 256)
(b) Word masked image
(c) LTVNN (PSNR: 32.920)
(d) TNNR (PSNR: 30.189)
(e) SVT (PSNR: 30.760)
(f) SVP (PSNR: 22.159)
(g) OptSpace (PSNR: 26.492)

Figure 4: Comparison of PSNR of recovered methods for Lena under word mask with $\gamma = 0.5$ by LTVNN, TNNR, SVT, SVP, and OptSpace.

Table 1: PSNR results under word mask sample with $\gamma = 0.5$ for different images by LTVNN, TNNR, SVT, SVP, and OptSpace.

	LTVNN	TNNR [10, 11]	SVT [12]	SVP [13]	OptSpace [14]
Mandrill (256 × 256)	29.495	27.845	27.736	18.881	26.264
Pepper (256 × 256)	33.421	31.019	30.188	23.648	27.141
Barbara (240 × 192)	31.170	29.261	29.352	22.901	23.427
Barbara (512 × 512)	32.680	30.682	28.855	22.113	26.872
Girl (256 × 256)	36.797	34.298	33.848	27.128	30.977
Couple (256 × 256)	36.916	35.176	35.241	29.649	32.815
Airplane (512 × 512)	31.883	30.083	25.506	19.573	26.222
House (256 × 256)	34.340	33.288	32.646	22.374	27.505
Sailboat (512 × 512)	30.858	29.103	27.079	26.778	20.891

so the term $\langle(\mathbf{X}-\mathbf{X}')(\mathbf{I}-\phi_1-\phi_1^T+\phi_1\phi_1^T),\mathbf{X}-\mathbf{X}'\rangle \geq 0$. The proof of Theorem 2 is as follows.

Proof. Let $(\mathbf{X}^*,\mathbf{Y}^*)$ be primal-dual optimization for the problem (27). The optimality conditions give

$$\begin{aligned}\mathbf{0} &= \mathbf{Z}^k - \mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}\right),\\ \mathbf{0} &= \mathbf{Z}^* - \mathscr{P}_\Omega\left(\mathbf{Y}^*\right),\end{aligned} \tag{A.3}$$

where $\mathbf{Z}^k \in \partial f_\tau(\mathbf{X}^k)$ and $\mathbf{Z}^* \in \partial f_\tau(\mathbf{X}^*)$. Then from (A.3), we deduce that

$$\left(\mathbf{Z}^k-\mathbf{Z}^*\right)-\mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}-\mathbf{Y}^*\right)=0 \tag{A.4}$$

and combine (A.4) with Lemma 1 that

$$\begin{aligned}&\left\langle \mathbf{X}^k-\mathbf{X}^*, \mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}-\mathbf{Y}^*\right)\right\rangle\\ &\quad=\left\langle \mathbf{Z}^k-\mathbf{Z}^*, \mathbf{X}^k-\mathbf{X}^*\right\rangle\\ &\quad\geq\left\langle \left(\mathbf{X}^k-\mathbf{X}^*\right)\left(\mathbf{I}-\phi_1-\phi_1^T+\phi_1^T\phi_1\right), \mathbf{X}^k-\mathbf{X}^*\right\rangle.\end{aligned} \tag{A.5}$$

We observe (23) that $\mathscr{P}_\Omega\mathbf{X}^* = \mathscr{P}_\Omega\mathbf{W}$,

$$\begin{aligned}&\left\|\mathscr{P}_\Omega\left(\mathbf{Y}^k-\mathbf{Y}^*\right)\right\|_F\\ &\quad=\left\|\mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}-\mathbf{Y}^*\right)+\lambda_k\mathscr{P}_\Omega\left(\mathbf{W}-\mathbf{X}^k\right)\right\|_F\\ &\quad=\left\|\mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}-\mathbf{Y}^*\right)+\lambda_k\mathscr{P}_\Omega\left(\mathbf{X}^*-\mathbf{X}^k\right)\right\|_F.\end{aligned} \tag{A.6}$$

Here, we set $r_k = \|\mathscr{P}_\Omega(\mathbf{Y}^k-\mathbf{Y}^*)\|_F$; then

$$\begin{aligned}r_k^2 &= r_{k-1}^2 - 2\lambda_k\left\langle \mathscr{P}_\Omega\left(\mathbf{Y}^{k-1}-\mathbf{Y}^*\right), \mathbf{X}^k-\mathbf{X}^*\right\rangle\\ &\quad+\lambda_k^2\left\|\mathscr{P}_\Omega\left(\mathbf{X}^*-\mathbf{X}^k\right)\right\|_F^2\\ &\leq r_{k-1}^2 - 2\lambda_k\left\langle \left(\mathbf{X}^k-\mathbf{X}^*\right)\left(\mathbf{I}-\phi_1-\phi_1^T+\phi_1^T\phi_1\right), \mathbf{X}^k-\mathbf{X}^*\right\rangle\\ &\quad+\lambda_k^2\left\|\mathbf{X}^k-\mathbf{X}^*\right\|_F^2\\ &= r_{k-1}^2 - \left(2\lambda_k\boldsymbol{\alpha}-\lambda_k^2\boldsymbol{\beta}\right),\end{aligned} \tag{A.7}$$

where $\boldsymbol{\alpha} = \langle(\mathbf{X}^k-\mathbf{X}^*)(\mathbf{I}-\phi_1-\phi_1^T+\phi_1^T\phi_1),\mathbf{X}^k-\mathbf{X}^*\rangle \geq 0$, $\boldsymbol{\beta} = \|\mathbf{X}^k-\mathbf{X}^*\|_F^2 \geq 0$.

Based on (A.7), when $(2\lambda_k\boldsymbol{\alpha}-\lambda_k^2\boldsymbol{\beta}) > 0$, that is, $\lambda_k \in (0, 2\boldsymbol{\alpha}/\boldsymbol{\beta})$, the term $\|\mathscr{P}_\Omega(\mathbf{Y}^k-\mathbf{Y}^*)\|_F$ is nonincreasing and converges to limit. The parameter λ_k is very easy for satisfying this condition when λ_k is smaller constant. And we can obtain other properties as follows.

Let $\lambda_k = \boldsymbol{\alpha}/\boldsymbol{\beta}$, and then $2\lambda_k\boldsymbol{\alpha}-\lambda_k^2\boldsymbol{\beta} = \boldsymbol{\alpha}^2/\boldsymbol{\beta}$. Due to the fact that $\boldsymbol{\alpha}^2/\boldsymbol{\beta}$ converges to zero, so $\boldsymbol{\alpha}^2$ is infinite small about $\boldsymbol{\beta}$ and converges to zero. Now we reconsider (A.2); evidently the first column in ξ converges to zero; that is, $\xi_{1,1} \to 0$, $\xi_{2,1} \to 0, \xi_{3,1} \to 0, \xi_{4,1} \to 0$. The second column converges to the first column and then converges to zero; that is, $\xi_{1,2} \to \xi_{1,1} \to 0, \xi_{2,2} \to \xi_{2,1} \to 0, \xi_{3,2} \to \xi_{3,1} \to 0, \xi_{4,2} \to \xi_{4,1} \to 0$. The third column converges to the second column and then converges to zero; that is, $\xi_{1,3} \to \xi_{1,2} \to 0, \xi_{2,3} \to \xi_{2,2} \to 0, \xi_{4,3} \to \xi_{4,2} \to 0, \xi_{1,2} \to \xi_{1,1} \to 0$, so through the iteration $\mathbf{X}^k$ converges to $\mathbf{X}^*$ except the last column due to the definition in (4) and (5); the last column and the last row are set to zero.

Fortunately, this problem does not have side effect for global result. Theorem 2 is established. □

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Basic Research Program of China under Grant 2011CB707904, by the National Natural Science Foundation of China under Grants 61201344, 61271312, and 61073138, by the Ministry of Education of China under Grants 20110092110023 and 20120092120036, the Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, and by Natural Science Foundation of Jiangsu Province under Grant BK2012329. This work is supported by INSERM postdoctoral fellowship.

References

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2005.

[4] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2008.

[5] E. J. Candès and T. Tao, "The power of convex relaxation: near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2009.

[6] A. Eriksson and A. van den Hengel, "Efficient computation of robust low-rank matrix approximations in the presence of missing data using the L1 norm," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 771–778, June 2010.

[7] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using Low rank matrix completion," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1791–1798, June 2010.

[8] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," in *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV '09)*, pp. 2114–2121, Kyoto, Japan, 2009.

[9] T. Okatani, T. Yoshida, and K. Deguchi, "Efficient algorithm for low-rank matrix factorization with missing components and

performance comparison of latest algorithms," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 842–849, Barcelona, Spain, November 2011.

[10] D. Zhang, Y. Hu, J. Ye, X. Li, and X. He, "Matrix completion by truncated nuclear norm regularization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 2192–2199, 2012.

[11] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, pp. 2117–2130, 2013.

[12] N. Komodakis and G. Tziritas, "Image completion using global optimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 442–452, 2006.

[13] C. Rasmussen and T. Korah, "Spatiotemporal inpainting for recovering texture maps of partially occluded building facades," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, vol. 3, pp. 125–128, September 2005.

[14] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using Low rank matrix completion," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1791–1798, June 2010.

[15] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 426–434, Las Vegas, Nev, USA, August 2008.

[16] H. Steck, "Training and testing of recommender systems on data missing not at random," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 713–722, Washington, DC, USA, July 2010.

[17] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

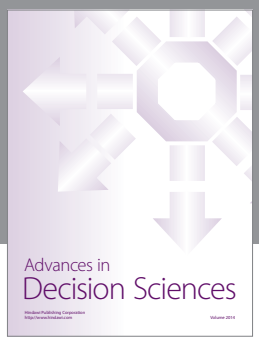
[18] P. Jain, R. Meka, and I. Dhillon, "Guaranteed rank minimization via Singular Value Projection," in *Proceedings of the 24th Annual Conference on Neural Information Processing Systems (NIPS '10)*, Vancouver, Canada, December 2010.

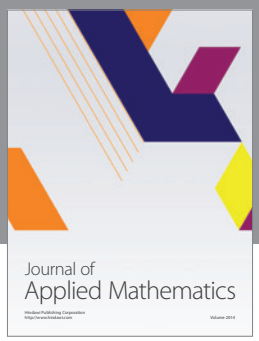
[19] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2980–2998, 2010.

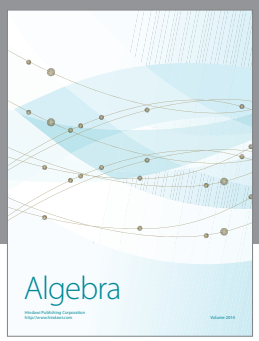
[20] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proceedings of the 25th Annual Conference on Neural Information Processing Systems (NIPS '11)*, December 2011.

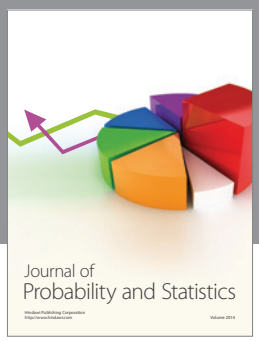
[21] J. Yang and X. Yuan, "Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization," *Mathematics of Computation*, vol. 82, no. 281, pp. 301–329, 2013.

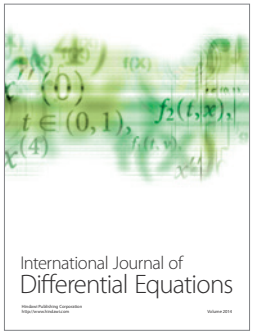
[22] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, 1992.

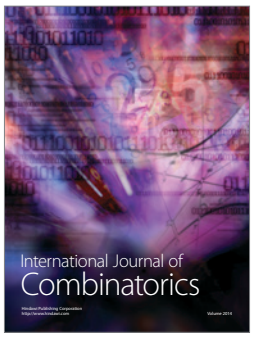


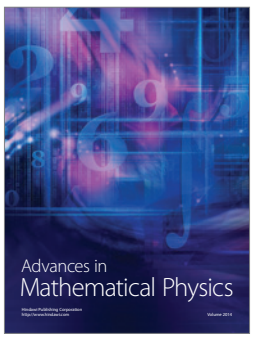


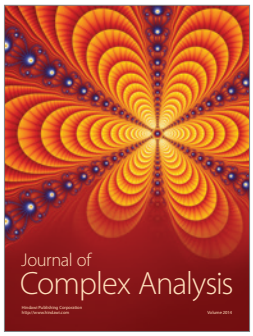


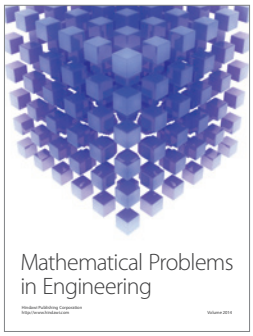


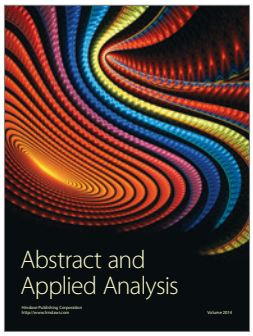


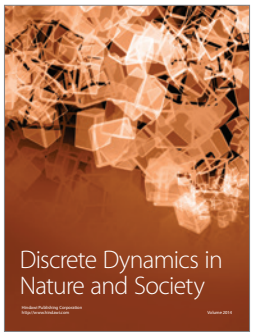


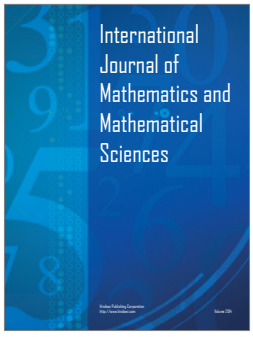


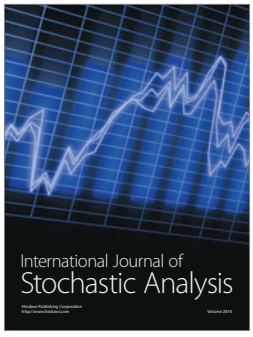












Hindawi
Submit your manuscripts at
http://www.hindawi.com