

基于矩阵分解与列采样的快速对齐去噪算法

刘林林*, 唐科威

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2026年8月7日; 录用日期: 2026年9月1日; 发布日期: 2026年9月10日

摘 要

针对原始RASL算法在图像对齐去噪任务中需求解核范数正则化问题、每次迭代均需计算大矩阵奇异值分解(SVD)而导致计算复杂度过高、难以适用于大规模数据的不足, 本文提出一种基于矩阵分解和列采样的快速对齐去噪算法。首先, 将低秩矩阵分解为两个子矩阵的乘积, 并使用这两个子矩阵的Frobenius范数约束替代原有的核范数约束, 从而避免直接求解SVD, 显著提升了计算效率和去噪性能。其次, 在求解过程中引入列采样策略, 进一步降低内存开销与子问题的计算复杂度。在四个典型数据集上的数值实验表明, 本文算法在保持优异对齐去噪效果的同时, 计算效率相比原RASL算法获得显著提升。

关键词

RASL, 矩阵分解, 列采样

A Fast Alignment-Based Denoising Algorithm Based on Matrix Decomposition and Column Sampling

Linlin Liu*, Kewei Tang

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: August 7, 2026; accepted: September 1, 2026; published: September 10, 2026

Abstract

To address the shortcomings of the original RASL algorithm in image alignment and denoising tasks—namely, the need to resolve kernel norm regularization issues, the requirement to compute singular value decomposition (SVD) of large matrices at every iteration (resulting in excessive computational complexity), and its difficulty in handling large-scale data—this paper proposes a fast

*通讯作者。

alignment and denoising algorithm based on matrix decomposition and column sampling. First, we decompose a low-rank matrix into the product of two submatrices and replace the original kernel norm constraint with Frobenius norm constraints on these submatrices. This avoids the need to directly compute SVD, significantly improving computational efficiency and denoising performance. Second, we introduce a column sampling strategy during the solution process, further reducing memory overhead and the computational complexity of subproblems. Numerical experiments on four typical datasets demonstrate that our algorithm achieves significantly improved computational efficiency compared to the original RASL algorithm while maintaining excellent alignment and denoising performance.

Keywords

RASL, Matrix Decomposition, Column Sampling

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像对齐是计算机视觉中的基本问题, 广泛应用于人脸识别、医学图像配准、视频处理等领域。传统批量图像对齐方法的进展主要体现在图像相似性度量的不断细化上。Learned-Miller 的聚合算法[1] [2] 通过最小化对齐后图像在每个像素位置上像素值熵的总和来进行对齐; 若将对齐后的图像堆叠成一个大矩阵的列, 该方法要求矩阵的每一行几乎恒定。与之不同, 文献[3] [4]中的最小二乘聚合方法通过最小化图像对之间距离的平方和来进行对齐, 它要求矩阵的每一列几乎恒定。在理想情况下, 上述两种方法得到的对齐图像矩阵秩为 1。然而, 当图像存在较大的光照变化时, 该矩阵的秩可能高于 1, 导致两种方法的效果不理想。Vedaldi 等人[5]采用了一种最小化对数行列式度量的方法, 可视为对秩函数的平滑近似, 但该方法无法同时处理大的光照变化以及真实图像中常见的严重像素损坏与遮挡。

RASL (Robust Alignment by Sparse and Low-rank Decomposition) [6]将对齐问题建模为: 寻找一组变换, 使得变换后的图像矩阵具有低秩结构, 同时分离稀疏误差。该方法能够在强光照变化和部分遮挡下完成批量对齐与去噪任务。然而, RASL 的核心计算瓶颈在于, 每一步求解核范数最小化问题时都需要进行大规模的奇异值分解(SVD)。当图像数量 n 或像素维度 m 较大时(如当 n 、 $m > 1000$ 时), RASL 算法所产生的时间和内存开销难以接受。

低秩矩阵分解是降低计算复杂度的常用手段。将低秩矩阵 $L \in R^{m \times n}$ 分解为两个小矩阵的乘积 $L = AB$, 其中 $A \in R^{m \times r}$, $B \in R^{r \times n}$, 且 $r \ll \min\{m, n\}$, 可以避免直接对 L 进行奇异值分解(SVD)。并且[7]中表明, $\min \|L\|_*$ 在相同的约束条件下等价于 $\min \|A\|_F^2 + \|B\|_F^2$ 。

因此可以用 Frobenius 范数替代核范数, 获得更易处理的优化问题。另一方面, 随机列采样技术通过选取少量具有代表性的图像列来近似整个数据矩阵的低维子空间[8], 可进一步降低每次迭代的计算开销。结合上述两种策略, 本文算法能够实现大规模图像对齐去噪问题的快速求解。

具体而言, 本文在 RASL 框架下提出以下两项改进:

(1) 显式低秩分解: 将待恢复的低秩对齐图像矩阵 L 直接写为两个矩阵的乘积 $L = AB$, 并在目标函数中加入正则项 $\|A\|_F^2 + \|B\|_F^2$, 从而避免了对大规模 SVD 的求解。

(2) 列采样加速: 在每次内循环中, 随机选取 c 列 $c \ll n$ 构造子问题, 利用这些少量列近似估计低秩

子空间并更新变换参数, 从而显著降低内存占用与单次迭代的计算复杂度。

2. 改进算法描述

2.1. 问题描述

设 $I_1, I_2, \dots, I_n$ 是输入的 n 张图像, 那么存在域变换 $\tau_1, \tau_2, \dots, \tau_n$, 使得变换后的图像 $I_1 \circ \tau_1, I_2 \circ \tau_2, \dots, I_n \circ \tau_n$ 在像素级别上是良好对齐的, 也就是说矩阵 $D \circ \tau = [\text{vec}(I_1^0) | \text{vec}(I_2^0) | \dots | \text{vec}(I_n^0)] \in R^{m \times n}$ 是低秩的, 其中 $I_j^0 = I_j \circ \tau_j$, $j = 1, 2, \dots, n$, $D = [\text{vec}(I_1) | \text{vec}(I_2) | \dots | \text{vec}(I_n)] \in R^{m \times n}$, 而 τ 表示将图像对齐的 n 个变换 $\tau_1, \tau_2, \dots, \tau_n$ 的集合。RASL 的目标函数是:

$$\min_{L, E, \tau} \|L\|_* + \lambda \|E\|_1, \text{ s.t. } D \circ \tau = L + E. \quad (1)$$

而本文将低秩部分 L 分解为 $L = AB$, 其中 $A \in R^{m \times r}$, $B \in R^{r \times n}$, 秩 r 的设定影响低秩模型的表达能力。若 r 设定过小, 则低秩基底无法充分捕捉对齐后图像的结构, 容易导致欠拟合; 若 r 设定过大, 则不仅会引入噪声, 还会显著增加计算开销。所以结合本文实验所用数据集, 综合权衡精度与运算效率, 本文统一经验性地设定 $r = 10$ 。

将原 RASL 模型中的核范数约束 $\|L\|_*$ 成正则项 $\|A\|_F^2 + \|B\|_F^2$, 有效避免了每一次子问题更新时需要求解 SVD, 得到如下优化模型:

$$\min_{A, B, E, \tau} \|A\|_F^2 + \|B\|_F^2 + \lambda \|E\|_1, \text{ s.t. } D \circ \tau = AB + E. \quad (2)$$

针对采用这里, 由于约束条件 $D \circ \tau = AB + E$ 是关于 τ 的一个非线性约束条件, 使得整个优化问题是非凸的, 难以直接求解。本文采取 RASL 算法中的思路, 将约束条件进行线性化, 使得整个优化模型可以用增广拉格朗日乘子法进行求解。

当 τ 的变化很小时, 可以通过在当前估计值 τ 处进行一阶泰勒展开来近似这个非线性约束条件, 这里我们假设 G 是某个 p 参数群, $\tau = [\tau_1 | \dots | \tau_n] \in R^{p \times n}$ 视为所有变换的参数化表示。对于 $\Delta \tau = [\Delta \tau_1 | \dots | \Delta \tau_n] \in R^{p \times n}$, 当初始错位不大时, 在 τ_i 处进行一阶泰勒展开有:

$$D \circ (\tau + \Delta \tau) \approx D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \epsilon_i^T, \quad (3)$$

其中 $J_i \triangleq \frac{\partial}{\partial \zeta} \text{vec}(I_i \circ \zeta) \Big|_{\zeta=\tau_i} \in R^{m \times p}$ 是第 i 幅图像关于变换参数 τ_i 的雅可比矩阵, $\{\epsilon_i\}$ 表示 R^n 的标准基。此时优化模型变为:

$$\min_{A, B, E, \Delta \tau} \|A\|_F^2 + \|B\|_F^2 + \lambda \|E\|_1, \text{ s.t. } D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \epsilon_i^T = AB + E. \quad (4)$$

2.2. 使用 Frobenius 范数的理论说明

在这一小节, 我们将重点讨论为什么可以用 $L = AB$ 的两个分解矩阵 A 和 B 的 Frobenius 范数代替 L 的核范数进行低秩正则化约束, 这种使用 Frobenius 范数进行低秩正则化的理论依据来源于[9]中的引理 5.1, 即在相同的约束条件下, 最小化 $\|L\|_*$ 等价于最小化 L 的两个分解矩阵 A 和 B 的 Frobenius 范数。并且对于本文所涉及的数据集(如人脸、建筑数据集等), 其变换后的对齐图像矩阵通常具有显著的本质低秩特性, 所以在对齐去噪背景下, 将核范数进行替换以加速原始 RASL 算法是合理的。值得指出的是, 使用 Frobenius 范数替代核范数将原问题中的凸低秩约束转化为了 A 和 B 双线性乘积项的约束, 存在陷入局部

最优的潜在风险, 但是本文采用 ADMM (交替方向乘子法) 进行迭代求解, 对于 A 和 B 这两个变量对应的子问题均为最小二乘问题, 且拥有显式的闭式解。在四个数据集上的对齐去噪结果表明, 在这种条件下, 算法在实际应用中拥有良好的稳定收敛行为。

2.3. 列采样加速策略

原 RASL 模型每次迭代需对全部 n 列进行操作, 当 n 较大时计算量巨大。本文采用直接对数据矩阵 D 进行列采样的策略, 将 D 划分为两个部分: 采样列 D_s 和未采样列 D_u 。

在每轮外循环开始时, 均匀随机选取列索引子集 $(\Omega \subset \{1, \dots, n\})$, 大小 c 。令 $D_s \in R^{m \times c}$ 为采样列, $D_u \in R^{m \times (n-c)}$ 为剩余未采样列。对应的变换参数、雅可比矩阵等也按相同索引分割。

将 A 看作是基底, B 看作系数矩阵, 将 B 也分割为两部分 $B_s \in R^{r \times c}$ 和 $B_u \in R^{r \times (n-c)}$ 。在内循环中, 先使用采样列 D_s 更新得到对应的基底 A 、采样列的系数矩阵 B_s 、对应变换增量 $\Delta \tau_s$ 和稀疏误差矩阵 E_s 。在得到 A 和 B_s 后, 利用 A 和约束条件, 通过最小二乘法得到 B_u 。然后再用得到的已知矩阵 A 、 B 更新稀疏误差矩阵 E 和变换参数 $\Delta \tau$ 。由于 $c = \gamma * n$ (其中 γ 为采样比例), 使得计算量显著减少。

3. 算法求解和收敛性分析

3.1. 算法求解

第一步: 使用采样部分 D_s 求解下面的优化问题:

$$\min_{A, B_s, E_s, \Delta \tau_s} \|A\|_F^2 + \|B_s\|_F^2 + \lambda \|E_s\|_1, \quad \text{s.t.} \quad D_s \circ \tau + \sum_{i=1}^c J_i \Delta \tau_s \varepsilon_i \varepsilon_i^T = AB_s + E_s. \quad (5)$$

我们采用 ADMM 算法求解模型, 则上面模型的增广拉格朗日函数为:

$$\begin{aligned} \mathcal{L}_\mu(A, B_s, E_s, \Delta \tau_s, Y_s) = & \|A\|_F^2 + \|B_s\|_F^2 + \lambda \|E_s\|_1 + \left\langle Y_s, D_s \circ \tau_s + \sum_{i=1}^c J_i \Delta \tau_s \varepsilon_i \varepsilon_i^T - AB_s - E_s \right\rangle \\ & + \frac{\mu_1}{2} \left\| D_s \circ \tau_s + \sum_{i=1}^c J_i \Delta \tau_s \varepsilon_i \varepsilon_i^T - AB_s - E_s \right\|_F^2 \end{aligned}$$

其中 $Y_s \in R^{m \times c}$ 为拉格朗日乘子, $\mu_1 > 0$ 为惩罚参数。

(1) 固定其他变量, 更新 A :

$$\min_A \|A\|_F^2 + \frac{\mu_1}{2} \left\| AB_s - \left(D_s \circ \tau_s + \sum_{i=1}^c J_{i,s} \Delta \tau_s \varepsilon_i \varepsilon_i^T - E_s + \frac{Y_s}{\mu_1} \right) \right\|_F^2$$

这是一个关于 A 的最小二乘问题, 有闭式解, 关于 A 求导并令导数等于 0, 解得:

$$A_{k+1} = \mu_{1,k} W_k B_{s,k}^T \left(2I + \mu_{1,k} B_{s,k} B_{s,k}^T \right)^{-1}. \quad (6)$$

(2) 固定其他变量, 更新 B_s :

$$\min_{B_s} \|B_s\|_F^2 + \frac{\mu_1}{2} \left\| AB_s - \left(D_s \circ \tau_s + \sum_{i=1}^c J_{i,s} \Delta \tau_s \varepsilon_i \varepsilon_i^T - E_s + \frac{Y_s}{\mu_1} \right) \right\|_F^2$$

这是一个关于 B_s 的最小二乘问题, 有闭式解, 关于 B_s 求导并令导数等于 0, 解得:

$$B_{s,k+1} = \mu_{1,k} \left(2I + \mu_{1,k} A_{k+1}^T A_{k+1} \right)^{-1} A_{k+1}^T W_k. \quad (7)$$

(3) 固定其他变量, 更新 E_s :

$$\min_{E_s} \lambda \|E_s\|_1 + \frac{\mu_1}{2} \left\| E_s - \left(D_s \circ \tau_s + \sum_{i=1}^c J_{i,s} \Delta \tau_s \varepsilon_i \varepsilon_i^T - AB_s + \frac{Y_s}{\mu_1} \right) \right\|_F^2$$

可以用软阈值算子求解:

$$E_{s,k+1} = S_{\lambda/\mu_{1,k}}(T_{s,k}) \quad (8)$$

其中 $T_{s,k} = D_s \circ \tau_s + \sum_{i=1}^c J_{i,s} \Delta \tau_{s,k} \varepsilon_i \varepsilon_i^T - A_{k+1} B_{s,k+1} + \frac{Y_{s,k}}{\mu_{1,k}}$, $S_\alpha[x] = \text{sign}(x) \cdot \max(|x| - \alpha, 0)$ 。

(4) 固定其他变量, 更新 $\Delta \tau_s$

定义残差矩阵 $R_s = AB_s + E_s - D_s \circ \tau_s - \frac{Y_s}{\mu_1}$, 约束中 $\Delta \tau_{s,i}$ 仅出现在第 i 列, 且相互独立。因此对于每个 $i=1, \dots, c$, 关于 $\Delta \tau_{s,i}$ 子问题为:

$$\min \frac{\mu_1}{2} \|J_{i,s} \Delta \tau_{s,i} - R_{s,i}\|_2^2$$

其中 $R_{s,i}$ 是 R_s 的第 i 列。这个最小二乘问题的解为:

$$\Delta \tau_{s,i} = J_{i,s}^\dagger (R_{s,i}) \quad (9)$$

其中 $J_{i,s}^\dagger$ 是 $J_{i,s}$ 的 Moore-Penrose 广义逆。

(5) 更新拉格朗日乘子 Y_s :

$$Y_{s,k+1} = Y_{s,k} + \mu_{1,k} \left(D_s \circ \tau_s + \sum_{i=1}^c J_{i,s} \Delta \tau_{s,k+1} \varepsilon_i \varepsilon_i^T - A_{k+1} B_{s,k+1} - E_{s,k+1} \right) \quad (10)$$

第二步: 用第一步得到的基底 A 通过投影求解系数矩阵 B

对于每一张图像, 通过最小化下面目标函数可以得到 B 的每一列 b_i :

$$\min_{b_i} \|b_i\|_2^2 + \frac{\mu_1}{2} \left\| \left(D_i \circ \tau_i + \Delta \tau_i \varepsilon_i - E_i + \frac{Y_{s,i}}{\mu_1} \right) - Ab_i \right\|_2^2$$

该问题有闭式解:

$$b_i = (\mu_1 A^T A + 2I)^{-1} (\mu_1 A^T M_i) \quad (11)$$

其中 $M_i = D_i \circ \tau_i + \Delta \tau_i \varepsilon_i - E_i + \frac{Y_{s,i}}{\mu_1}$ 。

第三步: 得出 A 和 B 后, 进一步更新稀疏矩阵 E 和变换增量 $\Delta \tau$ 。

已知 A 和 B 后, 优化模型为:

$$\min_{E, \Delta \tau} \lambda \|E\|_1, \quad D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \varepsilon_i \varepsilon_i^T = AB + E.$$

用 ADMM 求解该模型, 则增广拉格朗日函数为:

$$L_\mu(E, \Delta \tau, Y_1) = \lambda \|E\|_1 + \left\langle Y_1, D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \varepsilon_i \varepsilon_i^T - AB - E \right\rangle + \frac{\mu_2}{2} \left\| D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \varepsilon_i \varepsilon_i^T - AB - E \right\|_F^2$$

固定其他变量更新 E :

$$\min_E \lambda \|E\|_1 + \frac{\mu_2}{2} \left\| E - \left(D \circ \tau + \sum_{i=1}^n J_i \Delta \tau \varepsilon_i \varepsilon_i^T - AB + \frac{Y_1}{\mu_2} \right) \right\|_F^2$$

此问题可以用软阈值算子求解:

$$E_{k+1} = S_{\lambda/\mu_{2,k}}(T_k) \quad (12)$$

$$\text{其中 } T_k = D \circ \tau + \sum_{i=1}^n J_i \Delta \tau_k \varepsilon_i \varepsilon_i^T - AB + \frac{Y_{1,k}}{\mu_{2,k}}$$

固定其他变量更新 $\Delta \tau$:

参考第一步中更新 $\Delta \tau_s$ 的过程, 对于每个 i 单独求解下面的子问题:

$$\min_{\Delta \tau_i} \frac{\mu_2}{2} \|J_i \Delta \tau_i - R_i\|_F^2$$

其中 $R = AB + E - D \circ \tau - \frac{Y_1}{\mu_2}$, R_i 是 R 的第 i 列。

这是一个最小二乘问题, 它的解为:

$$\Delta \tau_i = J_i^\dagger R_i \quad (13)$$

这里 $J_i^\dagger$ 表示 J_i 的 Moore-Penrose 广义逆。

固定其他变量, 更新拉格朗日乘子 Y_1 :

$$Y_{1,k+1} = Y_{1,k} + \mu_{2,k} \left(D \circ \tau + \sum_{i=1}^n J_i \Delta \tau_{k+1} \varepsilon_i \varepsilon_i^T - AB - E_{k+1} \right) \quad (14)$$

算法的流程如表 1:

Table 1. System resulting data of standard experiment

表 1. 算法求解过程

本文算法求解过程
输入: 图像 $I_1, \dots, I_n \in R^{w \times h}$, 初始变换 $\tau_1, \dots, \tau_n$, 秩 r , 采样列数 c , 参数 $\lambda > 0$
输出: 低秩基底 A , 系数矩阵 B , 稀疏误差 E , 变换参数 τ
初始化: $A = 0$, $B = 0$, $E = 0$, $Y_1 = 0$, $\mu = \mu_0$
1. 列采样: 随机选取列索引子集 $\Omega \subset \{1, \dots, n\}$, 大小为 c , 采样部分为 $D_s \in R^{m \times c}$, 未采样部分为 $D_u \in R^{m \times (n-c)}$, 采样部分对应的变换参数为 τ_s , 雅可比矩阵为 $\{J_{i,s}\}_{i \in \Omega}$
2. 计算雅可比矩阵与归一化图像:
对于每个采样图像 $i \in \Omega$, $D_s = [d_{i_1,s}, \dots, d_{i_c,s}]$, $J_{i,s} = \frac{\partial}{\partial \zeta} \left(\frac{\text{vec}(I_i \circ \zeta)}{\ \text{vec}(I_i \circ \zeta)\ _2} \right) \bigg _{\zeta=\tau_i}$, $d_{i,s} = \frac{\text{vec}(I_i \circ \tau_i)}{\ \text{vec}(I_i \circ \tau_i)\ _2}$ 。
3. 内循环
① 求解 $\min_{A, B_s, E_s, \Delta \tau_s} \ A\ _F^2 + \ B_s\ _F^2 + \lambda \ E_s\ _1$ s.t. $D_s \circ \tau + \sum_{i=1}^c J_i \Delta \tau_s \varepsilon_i \varepsilon_i^T = AB_s + E_s$ 。分别使用(6), (7), (8), (9), (10)更新变量和乘子;
② 固定阶段一得到的 A , 用(11)求解系数矩阵 B 的每一列 b_i ;
③ 已知 A, B 后, 用(12), (13)分别更新稀疏矩阵 E 和变换增量 $\Delta \tau$ 。
4. 更新变换参数:
$\tau_i \leftarrow \tau_i + \Delta \tau_i$

3.2. 收敛性分析

在这一小节, 我们将讨论所提算法的收敛性分析, 由于本文重点的改进方向是对于低秩正则化项的改进, 采用了相同的泰勒线性化更新策略, 因此继承了原始 RASL 算法的收敛理论保证。并且本文算法的优化目标函数对于新增变量 A 和 B 都是凸的, 而且有显式的闭式解, 同时第 4 节的实验结果表明, 本文算法在四个经典数据集上均表现出稳定的收敛行为, 能够有效地完成图像的对齐与去噪任务。

4. 实验

为了全面评估本文所提方法的有效性, 我们在四个具有不同特点的数据集上和 RASL 算法进行对比实验, 本节将详细叙述实验的数据集介绍、实验参数设置以及实验结果的定性评价和定量评价。

4.1. 实验设置

4.1.1. 数据集介绍

为验证方法有效性, 我们在以下四个数据集上进行了实验:

(1) 数字 3 数据集(Digits-3) [10]: 该数据集包含了从 MNIST 手写数字数据库中选取 300 张“3”字图像, 每张图像大小为 28×28 像素。该数据集存在不同程度的旋转和书写风格差异, 用于评估算法对简单形变的对齐能力。

(2) Gore 视频数据集[10]: 该数据集包含了从 AI Gore 演讲视频中连续截取 140 帧图像, 每帧经裁剪为 80×60 像素。该数据集存在头部运动、表情变化和轻微遮挡, 用于评估算法对视频序列的稳定对齐能力。

(3) LFW 人脸数据集[11]: 该数据集包含了从 Labeled Faces in the Wild 数据集中选取“George W. Bush”和“Colin Powell”两个人物共 80 张图像, 每张图像裁剪为 80×60 像素。该数据集存在光照变化、姿态差异和部分遮挡, 用于评估算法在自然场景下的鲁棒性。

(4) Windows 建筑数据集[12]: 该数据集包含了从 16 张不同视角下拍摄的建筑立面图像, 每张图像大小为 200×200 像素, 部分图像被树枝遮挡。该数据集用于评估算法对平面单应变换的适应能力。

4.1.2. 实验参数设置

本文算法参数设置如下: 秩 $r=10$, 采样比例 $\gamma=0.3$, 正则参数 $\lambda=1/\sqrt{m}$ (m 为像素数), 惩罚参数初始值 $\mu_0=1.25/\|D\|_F^2$, 增长因子 $\rho=1.25$, 最大惩罚参数 $\mu_{\max}=10^6$, 收敛容差 $\varepsilon=10^{-5}$, 原始 RASL 使用相同的外循环参数, 其内循环采用 ALM 求解。

4.1.3. 实验评估指标

为了对比两种算法的运行时间, 我们在表 2 给出了两种算法在四个数据集上的运行时间对比, 同时为了量化不同图像之间的对齐一致性, 本文引入均方误差(MSE)作为定量指标。具体计算方法为: 首先对数据集中所有已完成对齐的图像进行像素均值计算, 得到一张平均图像; 然后逐一计算每一张对齐后图像与该平均图像之间的均方误差; 最后, 将该数据集下所有图像的 MSE 取平均值, 作为评估该数据集整体对齐精度的指标。我们在表 3 给出了两种算法在四个数据集上的均方误差对比, 均方误差值越小, 表明对齐后的图像在像素结构上的一致性越高, 对齐效果越好。

4.2. 实验结果

图 1~4 分别展示了在四个数据集上原始 RASL 与本文算法的对齐结果对比。每个图中包含了原图、RASL 对齐图和对齐去噪图、本文算法对齐图和对齐去噪图。同时表 2 对比了 RASL 和本文算法在四个

数据集上的运行时间, 表 3 给出了两种算法的均方误差对比。

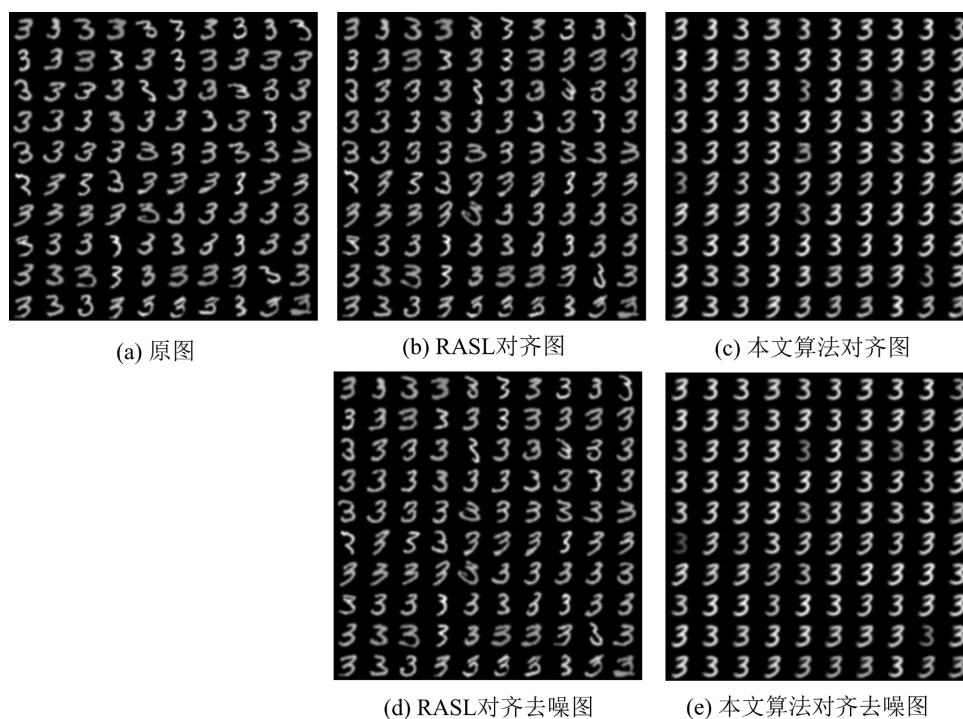


Figure 1. Comparison chart of the two methods on the “Digits-3” dataset

图 1. 两种方法在数字 3 数据集上的对比图

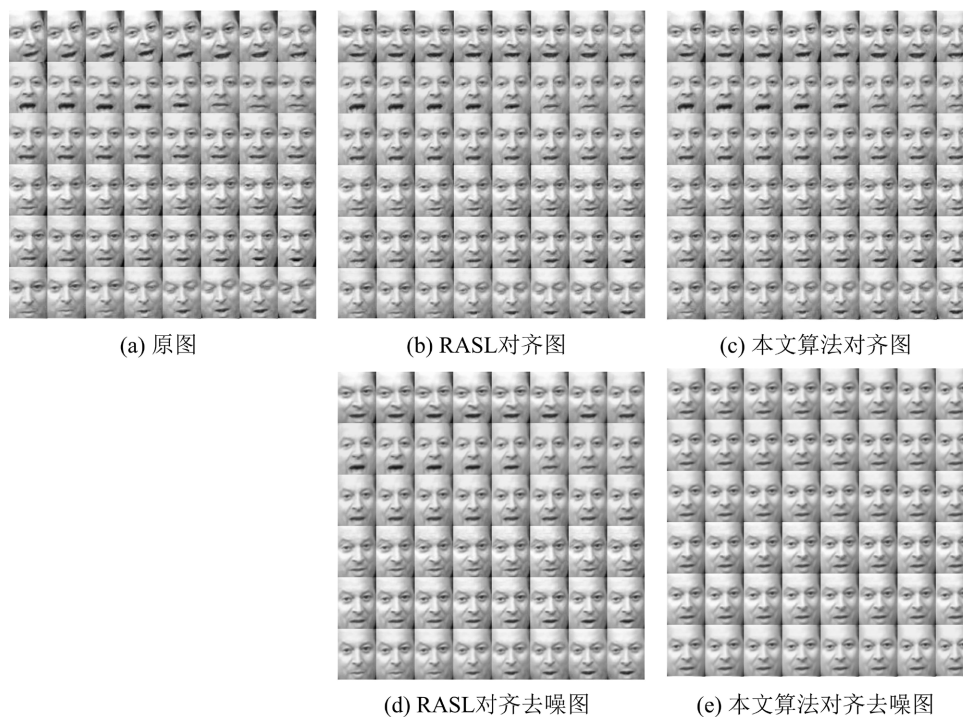


Figure 2. Comparison of the two methods on the Gore video dataset

图 2. 两种方法在 Gore 视频数据集上的对比图

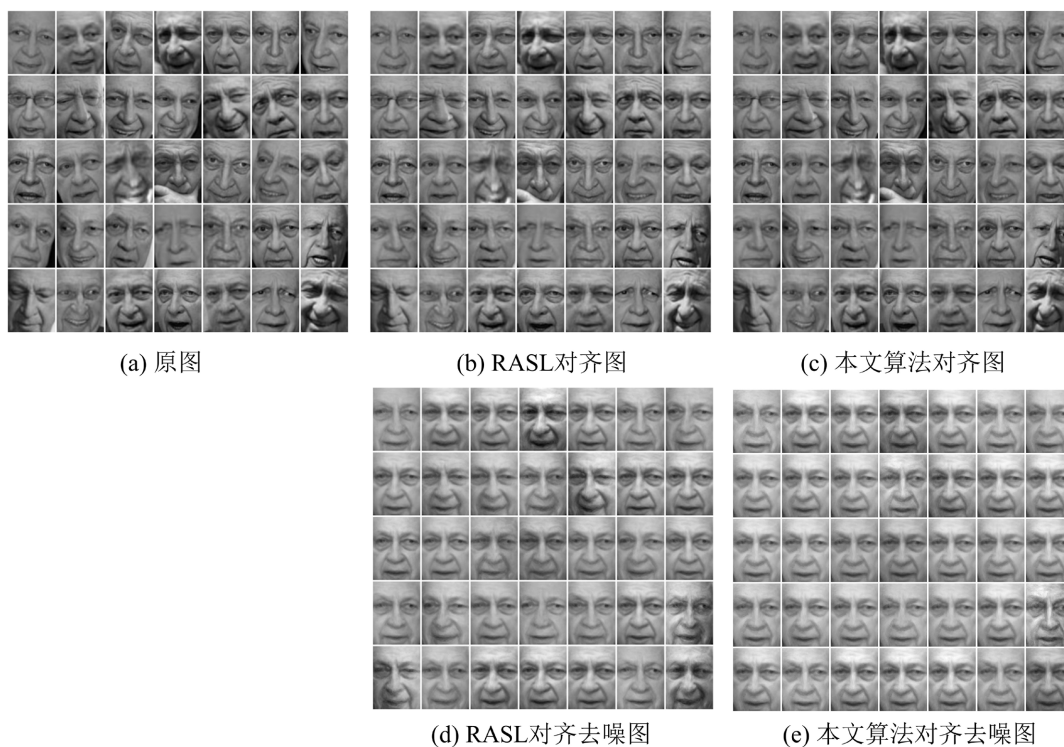


Figure 3. Comparison of the two methods on the LFW face dataset

图 3. 两种方法在 LFW 人脸数据集上的对比图

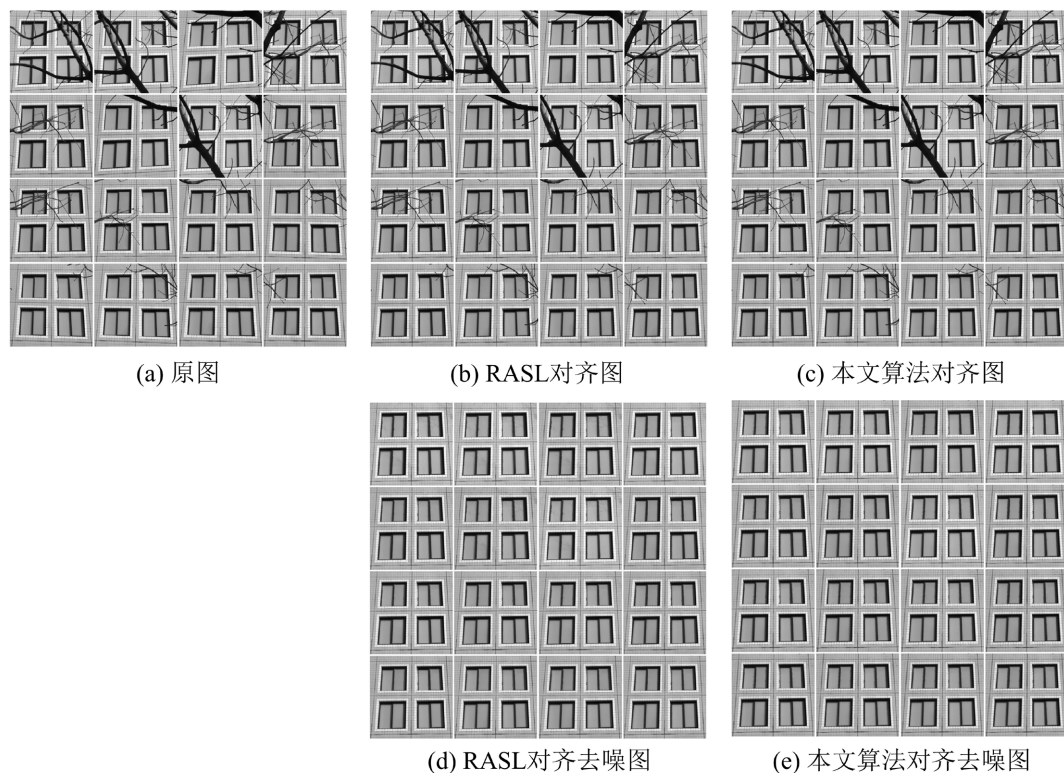


Figure 4. Comparison of the two methods using the Windows building dataset

图 4. 两种方法在 Windows 建筑数据集上的对比图

Table 2. Runtimes of RASL and the algorithm described in this paper on four datasets (in seconds)
表 2. RASL 和本文算法在四个数据集上的运行时间(单位: 秒)

数据集	Digits-3	Gore	LFW	Windows
RASL	372.29	263.62	69.12	135.08
本文算法	78.77	65.14	19.09	42.82

Table 3. Alignment accuracy (MSE) comparison between RASL and the proposed algorithm on four datasets
表 3. RASL 与本文算法在四个数据集上的对齐精度(MSE)对比

数据集	Digits-3	Gore	LFW	Windows
RASL	2.33×10^{-4}	1.20×10^{-6}	6.71×10^{-6}	9.44×10^{-7}
本文算法	2.20×10^{-4}	1.24×10^{-6}	6.35×10^{-6}	9.61×10^{-7}

4.3. 实验分析

从总体上看, 本文算法在视觉上达到了良好的对齐效果, 能够有效分离数字 3 的书写差异、人脸较大的表情变化以及建筑上的大范围遮挡。其中图 2 展示了人脸对齐去噪结果。本文算法与 RASL 算法均能实现较好的对齐, 但在去噪方面, RASL 无法成功分离人脸上变化较大的表情区域(如张大的嘴), 而本文算法能够将人脸的低秩部分与稀疏噪声部分清晰分离。图 4 展示了建筑物窗户对齐与去噪结果。RASL 虽然能够对齐窗户, 但在处理大面积遮挡时仍有残留, 未能完全去除; 相比之下, 本文算法不仅能成功对齐窗户, 还能完全移除遮挡窗户的树枝。这表明即便存在大范围遮挡, 本文算法仍能有效完成对齐与去噪任务, 其性能显著优于 RASL 算法。

表 2 汇总了 RASL 和本文算法在四个数据集上的运行时间, 表 2 表明了: 在计算效率方面, 我们的方法相比原 RASL 方法平均减少了约 75.3% 的计算时间, 显著提升了计算效率。

从表 3 可以看出: 在对齐精度方面, 本文算法与原始 RASL 算法在四个数据集上的最终对齐精度基本处于同一水平, 具体而言, 在 LFW 数据集和数字 3 数据集上, 本文算法实现了略优于原始 RASL 的对齐精度; 而在 Gore 视频数据集和 Windows 建筑数据集上, 原始 RASL 的 MSE 则略低一点。这说明本文算法在实现加速 RASL 算法的同时, 基本保持了实际应用场景所需的对齐精度。

5. 结论

本文针对 RASL 算法求解核范数正则化时需多次计算大矩阵 SVD、计算复杂度高且对齐去噪效果的不足, 提出了基于矩阵分解与列采样的快速对齐去噪算法。主要改进包括: (1) 将低秩矩阵分解为两个矩阵之积, 并用它们的 Frobenius 范数约束替代原核范数约束, 显著提升了计算效率与去噪性能; (2) 引入列采样策略, 进一步降低了子问题的计算开销。在四个数据集上的实验结果表明, 本算法在保持较好对齐精度的同时, 计算效率相比 RASL 有显著优势。

参考文献

- [1] Learned-Miller, E.G. (2006) Data Driven Image Models through Continuous Joint Alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**, 236-250. <https://doi.org/10.1109/tpami.2006.34>
- [2] Huang, G.B., Jain, V. and Learned-Miller, E. (2007) Unsupervised Joint Alignment of Complex Images. 2007 *IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, 14-21 October 2007, 1-8. <https://doi.org/10.1109/iccv.2007.4408858>
- [3] Cox, M., Sridharan, S., Lucey, S. and Cohn, J. (2008) Least Squares Congealing for Unsupervised Alignment of Images. 2008 *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 23-28 June 2008, 1-8.

- <https://doi.org/10.1109/cvpr.2008.4587573>
- [4] Cox, M., Sridharan, S., Lucey, S. and Cohn, J. (2009) Least-Squares Congealing for Large Numbers of Images. 2009 *IEEE 12th International Conference on Computer Vision*, Kyoto, 29 September-2 October 2009, 1949-1956. <https://doi.org/10.1109/iccv.2009.5459430>
 - [5] Vedaldi, A., Guidi, G. and Soatto, S. (2008) Joint Data Alignment up to (Lossy) Transformations. 2008 *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 23-28 June 2008, 1-8. <https://doi.org/10.1109/cvpr.2008.4587781>
 - [6] Peng, Y.G., Ganesh, A., Wright, J., Xu, W.L. and Ma, Y. (2012) RASL: Robust Alignment by Sparse and Low-Rank Decomposition for Linearly Correlated Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 2233-2246. <https://doi.org/10.1109/tpami.2011.282>
 - [7] Weimer, M., Karatzoglou, A. and Smola, A. (2008) Improving Maximum Margin Matrix Factorization. *Machine Learning*, **72**, 263-276. <https://doi.org/10.1007/s10994-008-5073-7>
 - [8] Deng, L. (2012) The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]. *IEEE Signal Processing Magazine*, **29**, 141-142. <https://doi.org/10.1109/msp.2012.2211477>
 - [9] Recht, B., Fazel, M. and Parrilo, P.A. (2010) Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *SIAM Review*, **52**, 471-501. <https://doi.org/10.1137/070697835>
 - [10] Wright, J., Peng, Y., Ma, Y., Ganesh, A. and Rao, S. (2009) Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization. *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, Vancouver, 7-12 December 2009, 2080-2088.
 - [11] Huang, G.B., Ramesh, M., Berg, T. and Learned-Miller, E. (2008) Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Technical Report, University of Massachusetts.
 - [12] Zhang, Z., Liang, X., Ganesh, A. and Ma, Y. (2011) TILT: Transform Invariant Low-Rank Textures. In: *Lecture Notes in Computer Science*, Springer, 314-328. https://doi.org/10.1007/978-3-642-19318-7_25