Journal of Computational and Applied Mathematics 346 (2019) 171–183

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

An efficient algorithm for batch images alignment with adaptive rank-correction term

Shuangyue Wang [a], Yunhai Xiao [b,*], Zhengfen Jin [c]

[a] *School of Mathematics and Statistics, Henan University, Kaifeng 475000, China*
[b] *Institute of Applied Mathematics, School of Mathematics and Statistics, Henan University, Kaifeng 475000, China*
[c] *School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, China*

ARTICLE INFO

Article history:
Received 15 January 2018
Received in revised form 19 May 2018

Keywords:
Batch images alignment
Non-smooth convex minimization
Alternating direction method of multipliers
Accelerated proximal gradient algorithm
Symmetric Gauss–Seidel

ABSTRACT

With the appearance of approach named "robust alignment by sparse and low-rank decomposition" (RASL), a number of linearly correlated images can be accurately and robustly aligned despite significant corruptions and occlusions. It has been discovered that this aligning task can be characterized as a sequence of 3-block convex minimization problems which can be solved efficiently by the accelerated proximal gradient method (APG), or alternatively, by the directly extended alternating direction method of multipliers (ADMM). However, the directly extended ADMM may diverge although it often performs well in numerical computations. Ideally, one should find an algorithm which can have both theoretical guarantee and superior numerical efficiency over the directly extended ADMM. We achieve this goal by using the intelligent symmetric Gauss–Seidel iteration based ADMM (sGS-ADMM) which only needs to update one of the variables twice, but surprisingly, it leads to the desired convergence to be guaranteed. The convergence of sGS-ADMM can be followed directly by relating it to the classical 2-block ADMM and with a couple of specially designed semi-proximal terms. Beyond this, we also add a rank-correction term to the model with the purpose of deriving the alignment results with higher accuracy. The numerical experiments over a wide range of realistic misalignments demonstrate that sGS-ADMM is at least two times faster than RASL and APG for the vast majority of the tested problems.

1. Introduction

In the past decades, the increasing popularity of digital cameras has led to a dramatic increase in the amount of visual data. Such data always contains significant illumination variations, partial occlusions, or even misalignments. Particularly, the last difficulty poses steep challenges to existing vision algorithms for many image analyses such as face recognition, and image classification. Therefore, how to design reliable and efficient alignment algorithms for a large amount of images is an urgent and fundamental problem in computer vision.

In the past decade, a lot of work has been done toward aligning images of objects of interest to a fixed canonical template. To seek an alignment, the congealing algorithm [1] minimizes the sum of the pixel-stack entropies at each location in the batch of the aligned images. However, it requires that each row of the matrix be nearly constant if these aligned images are stacked as the column of a large matrix. The least squares congealing approach [2] minimizes the sum of squared distances between image pairs but it demands that each column be nearly constant. However, if the constant conditions are not

* Corresponding author.
E-mail addresses: sywangmath@163.com (S. Wang), yhxiao@henu.edu.cn (Y. Xiao), jinzhengfen@126.com (Z. Jin).

https://doi.org/10.1016/j.cam.2018.07.007
0377-0427/

satisfied, the matrix of the aligned images might have an unknown rank rather than the desired one. The method in [3] uses a log-determinant cost function to be a smooth surrogate for the rank function, and the EM algorithm [4] optimizes a low rank objective function directly with respect to domain transformations drawn from a known group. However, as shown in [5] that, a major drawback of these approaches is that they cannot simultaneously handle the large illumination variations and gross pixel corruptions or partial occlusions that often occur in real images.

Unlike the conventional techniques, the novel approach named "robust alignment by sparse and low-rank decomposition" (RASL) [5] can seek an optimal set of image domain transformations such that the matrix of transformed images can be decomposed as the sum of a sparse matrix of errors and a low-rank matrix of recovered aligned images. More precisely, suppose that we are given n well-aligned and linearly-correlated grayscale images $I_1^0; \ldots; I_n^0 \in \mathbb{R}^{w\times h}$ of an object or a scene, where w and h are the width and length of each image, respectively. Denote $\mathcal{V} : \mathbb{R}^{w\times h} \to \mathbb{R}^m$ be the operator that selects m-pixel region from an image and stacks it as a vector and set $X = [\mathcal{V}(I_1^0), \ldots, \mathcal{V}(I_n^0)] \in \mathbb{R}^{m\times n}$, then X is approximately low-rank. However, suppose that there are n newly arrived images $I_1; I_2; \ldots; I_n$ of the same object but misaligned with respect to each other. Then, it has been shown that there exist domain transformations $\tau_1, \ldots, \tau_n$ such that the transformed images $I_1 \circ \tau_1, \ldots, I_n \circ \tau_n$ are well aligned at the pixel level, or, equivalently, the matrix

$$D \circ \tau = [\mathcal{V}(I_1 \circ \tau_1), \ldots, \mathcal{V}(I_n \circ \tau_n)]$$

has low-rank, where $D = [\mathcal{V}(I_1), \ldots, \mathcal{V}(I_n)]$ denotes the given images, and τ presents the set of the sequence $\tau_1, \ldots, \tau_n$. Therefore, the batch images alignment problem can be expressed to find the transformations τ such that $D \circ \tau$ has low-rank, which amounts to the following minimization problem

$$\min_{\tau,X} \big\{ \mathrm{Rank}(X), \ \text{s.t.}\ D \circ \tau = X \big\}. \tag{1.1}$$

However, in practice, the low-rank structure of the aligned images is usually violated due to the partial occlusions or corruptions. Suppose that the images $\{I_i \circ \tau_i - e_i\}_{i=1}^n$ are well aligned with a small error e_i corresponding to image I_i, then the formulation (1.1) can be modified correspondingly as

$$\min_{\tau,X,E} \big\{ \mathrm{Rank}(X) + \lambda \|E\|_0, \ \text{s.t.}\ D \circ \tau = X + E \big\}, \tag{1.2}$$

where $\|E\|_0$ denotes the number of non-zero entries in E and $\lambda > 0$ is a parameter that trades off both terms for minimization. However, the rank and l_0-norm minimization is combinatorial and known to be NP-hard, and generally computationally intractable. Therefore, convex relaxation is often used to make the minimization tractable.

The most popular choice is to replace the "rank" term with the nuclear norm, and replace the ℓ_0-norm with the ℓ_1-norm [6], which yields the following convex minimization problem

$$\min_{\tau,X,E} \big\{ \|X\|_* + \lambda \|E\|_1, \quad \text{s.t.} \quad D \circ \tau = X + E \big\}, \tag{1.3}$$

where $\|\cdot\|_*$ is the so-called nuclear norm (also known as Ky Fan norm) defined by the sum of all singular values, and $\|\cdot\|_1$ is defined as the sum of absolute values of all entries. It is worth noting that the objective function in model (1.3) is convex and separable, but the constraint is nonlinear, which leads to the main difficulty of minimization. We assume that $\mathbb{G}$ is some p-parameter group and identify $\tau \in \mathbb{R}^{p\times n}$ with the parameterizations of all of the transformations. To resolve this dilemma, the popular technique in [5,7] is to linearize the term $D \circ \tau$ at $\tau^{(i)}$ as $D \circ (\tau^{(i)} + \Delta\tau) \approx D \circ \tau^{(i)} + \sum_{j=1}^n J_j \Delta\tau_j$, where $\Delta\tau = (\Delta\tau_1, \ldots, \Delta\tau_n) \in \mathbb{R}^{p\times n}$ and $J_j \in \mathbb{R}^{m\times p}$ is the Jacobian of the jth image with respect to the transformation parameter τ_j, i.e.,

$$J_j = \frac{\partial}{\partial \zeta} \left(\frac{\mathcal{V}(I_j \circ \zeta)}{\|\mathcal{V}(I_j \circ \zeta)\|_2} \right) \Big|_{\zeta=\tau_j}, \quad j = 1, \ldots, n. \tag{1.4}$$

Therefore, solving (1.3) leads to solving a sequence of the following convex minimization problem

$$\min_{X,E,\Delta\tau} \Big\{ \|X\|_* + \lambda \|E\|_1, \ \text{s.t.}\ D \circ \tau^{(i)} + \sum_{j=1}^{n} J_j \Delta\tau_j = X + E \Big\}. \tag{1.5}$$

This model (1.5) is convex, separable and non-smooth, so it can be solved efficiently. When the solution $\Delta\bar{\tau}^{(i)}$ is derived, the domain transformation τ is updated immediately as $\tau^{(i+1)} = \tau^{(i)} + \Delta\bar{\tau}^{(i)}$.

The resulting model (1.5) actually has three separable structures in both objective function and constraint, and hence it belongs to the framework of the alternating direction method of multipliers (ADMM). For solving (1.5), the original solver [8] penalized the constraint as an unconstrained minimization problem and then solved accordingly by the accelerated proximal gradient (APG) algorithm. Although the implemented APG algorithm performs efficiently, it only solves an approximated variant, not the original model (1.5) itself. To solve (1.5), Peng et al. [5] directly extended the classical 2-block ADMM to the 3-block case with $X \to \Delta\tau \to E$ Gauss–Seidel order which has been observed to perform well in numerical computations. However, it was shown by Chen et al. [9] that in contrast to the classic 2-block ADMM, the directly extended 3-block ADMM may diverge theoretically. Ideally, one should find a convergent variant which is at least as efficient as the directly extended