**BACKGROUND REMOVAL IN VIDEOS WITH LOW-RANKED FACTORIZATION**

Presented by-

Tapas Saha Rahul Chattopadhyay

Senior manager - Analytics, Eureka Forbes

**Abstract:**

Principal component analysis (PCA) is a technique used for dimension reduction purpose. It emphasizes variation and brings out strong patterns within a dataset. It's used to make data easy to explore and visualize. Basically, PCA is applied for exploratory data analysis, dimension reduction and clustering. Although we all know that PCA is most sensitive when outliers are presented in the data set and, thus, is vulnerable to both gross measurement error and adversarial manipulation of the data. In this paper we will discuss about the Robust Principal component analysis (Robust PCA). We decomposed a data matrix in summing form of two matrices, one low-rank and other one sparse component by solving a convex problem called Principal Component Pursuit (RPCA via PCP) by applying simply minimize a weighted combination of the nuclear norm and of the -norm. This Robust statistical method is good enough to deal with corrupted data or missing values. There are many real-life use cases in many fields, in video surveillance, where this methodology allows to breakdown into background and foreground; in face recognition, where we can remove shadows, identifying faces; web data ranking and bioinformatic data.

Keywords: Principal component analysis, Robust PCA, low-rank and sparce matrix, nuclear norm, matrix decomposition

**Introduction**

Any area of a moving or deforming image can be separated from a static or moving background. This is basically a video processing task with many applications. Through this paper, we address this basic task by making maximum use of recent developments in the field of Robust Principal Component Analysis (RPCA). Basically, our work in this paper here is based on the critical breakthrough accomplished by Candes et al. [1], where the authors described a real-life solution for the complex optimization scenario for recovering the sparse parts and low-rank of a large matrix which is the sum of these two entities.

Suppose, we are given a large data matrix , which may be decomposed as , where is low-rank and S is sparse; having both elements are of arbitrary magnitude. We neither know about the low-dimensional column and row space of nor their dimension. Same for S also, we neither know the locations of the nonzero entries nor even how many there are. Now can we construct the low-rank and sparse components both correctly and significantly?

If we stack all the data points as column vectors of a matrix , this matrix may be or not contain low-rank. mathematically, , where is low-rank and N is a small perturbation matrix. Classical Principal Component Analysis gives us the rank- estimate of by solving

minimize |||| subject to rank () ≤

One can easily solve this problem scenario via using singular value decomposition (SVD) and one can get a number of optimal properties where the noise N is small and i.i.d. on Gaussian.

**Robust PCA**: Nowadays, PCA is debatably the most widely used statistical tool for exploratory data analysis and dimensionality reduction. However, it is vulnerable to both gross measurement error and adversarial manipulation of the data. Thus, presence of a single grossly corrupted entry within M could redact the estimated arbitrarily far from the true . Here Robust PCA comes, where we can recover a low-rank matrix from highly corrupted data matrix . Where instead of the small noise term N in classical PCA, the elements in S may be arbitrarily large in magnitude, and supposed to be sparse but unknown.

Suppose, a 2-dimensional matrix M of size m × n, which stores pixel information from a video sequence , j = 1, . . . , n, or simply a set of images by concatenating each frame as a column in . Then, the background part of the video sequence is stored within the low-rank matrix and the locally deforming parts embedded in the sparse matrix. The above optimization model using the and nuclear norms with the convex optimization problem named Principal Component Pursuit (PCP) which provides us a feasible solution of (1) that can recover the low-rank (background) and the sparse (foreground) parts of the original matrix.

argmin || ||∗ + ||S||1 subject to (1)

where L and S are matrices of the same size as M, || · ||∗ and || · ||1 denotes nuclear norm ( -norm of the singular values) and -norm respectively. is arbitrary balanced parameter (according to above problem, value chosen as follow ). PCP plays an important role in image processing and video processing field. Even though this above formula leads to a computationally feasible solution, but the complexity is still high due to involving the calculation of many Singular Value Decompositions (SVD) for a very large matrix.

**Objective and Assumption**

**Objective**

* Through this paper, we bring in a modified approximated RPCA algorithm that can deal with moving cameras, also gives importance on the block sparse structure of the pixels that corresponds to the moving objects. We propose a SVD-free algorithm for background that outsails present state-of-the-art methods with computational cost/time and performance. Finally, we have demonstrated the experiments and numerical results by evaluating the proposed methods.

**Assumption**

* There is a critical issues and limitations of RPCA that estimates global motion parameters simultaneously with the foreground and background separation task. Here we are considering matrix block-sparsity instead of generic matrix sparsity as an inherited feature in video processing applications. It provides an extremely efficient algorithm to solve the low-rank and sparse decomposition problem.
* We assume that the original matrix is decomposable, which means that the original matrix can be written as the sum of two matrices mainly a low-rank matrix with singular vectors that are smooth, and a sparse matrix which has a uniform and random pattern of sparsity.
* In other words, the non-zero entries, within the sparse matrix, are not randomly distributed but they formed small blocks within the matrix. Thus, the final solution led RPCA approaches are computationally expensive.

**Closely Related Research Work**

Various algorithmic procedures have been proposed to address and solve the problems in video processing as well as image processing also, viz

* Principal Component Analysis (PCA) (Oliver et al. (1999))
* RPCA via Robust Subspace Learning (RSL) (Torre & Black (2001); Torre & Black (2003))
* RPCA via on Graphs (RPCAG) (Shahid et al. (2015)) and Fast Robust PCA on Graphs (FRPCAG) (Shahid et al. (2016))
* RPCA via Templates for First-Order Conic Solvers (TFOCS) (Becker et al. (2011))
* RPCA via Inexact Augmented Lagrange Multiplier (IALM) (Lin et al. (2009))
* RPCA via Bayesian Framework (BRPCA) (Ding et al. (2011))

Torre & Black (2003) proposed a RPCA method called Robust Subspace Learning (RSL) which is a batch robust PCA method that focus how to recover a good low-rank approximation which is the best fit of majority of data. RSL solves a nonconvex optimization with the different minimized problem based on the concept of soft-detecting, with down-weighting the outliers.

Becker et al. (2011) stated the same opinion as Candes et al. (2009) that composes of some matrix which can be broken down into two matrices , where is low-rank matrix and is sparse component. Here the objective function remains same as PCP only difference in constraints portion, instead of the constraints , the constraints is

Ding et al. (2011) raised a hierarchical Bayesian framework that deliberated for decomposing an original matrix into a low-rank matrix , a sparse matrix along with a noise matrix . Moreover, the Bayesian framework grants exploitation of additional structural form within the matrix. Markov dependency is incorporated between successive rows in the matrix that implies an appropriate temporal dependency. As moving objects are strongly correlated across successive frames.

Zhou et al. introduced a method called DECOLOR where we can segment out moving parts from an image sequence by organizing a Markov Random Fields (MRF) framework along with solving a non-convex penalty. Whereas DECOLOR minimizes a non-convex energy through a different optimization. It converges to a local optimum with results which depends on the initialization of the foreground support, where PCP always minimizes its energy globally.

In other work to come across this time addressing the complexity issue, Zhou and Tao proposed the approximated RPCA, called GoDec. The intention of GoDec is to provide a probable solution for the low-rank and sparse decomposition in presence of noise. It converges to a local minimum, when the exact and unique decomposition does not exist for many real scenarios.

**Methodology**

Candes et al. [1] solved Robust PCA by decomposing , where and be the low-rank matrix and sparse matrix respectively. The direct method is to use norm to minimize the energy function

(1)

with λ is arbitrary balanced parameter. The above problem is known as principal component pursuit (PCP). Here neither nor are convex hence (1) is an intractable convex problem. To break this more freely, the easiest way is to go with norm that presents an approximate minimization convex problem

(2)

|| · ||∗ denotes nuclear norm i.e., the sum of singular values. By choosing these small assumptions, the principal component pursuit solution entirely regains the low-rank and the sparse entities, provided that these two matrices are bounded by the following inequality

,

where, and are positive numerical constants. m and n are the size of the original matrix . After further discussion, is chosen as and by choosing this value of the solution of (2) approximately converges to the solution of (1).

For this paper, we have chosen augmented Lagrange multiplier (ALM) algorithm to solve the convex PCP problem [2,3]. For this ALM method, operates on the augmented Lagrangian as follows

we solve PCP by repeatedly updating, and next updating the Lagrange multiplier by .

For our matrix decomposition problem, we can solve a sequence of convex programs by identifying and both are having efficient and simple solutions. Suppose denote the shrinkage operator where , and extending to matrices by applying it towards every entity, i.e.

.

Also, for matrices , let which denotes the singular value thresholding operator where , and is any singular value decomposition, i.e., in simple word it become

.

Hence, our practical strategy is to minimize with respect to by fixing first, next minimize with respect to by fixing , and then at last upgrade the Lagrange multiplier matrix based on the residual .

We summarized the above strategy as below:

Step 1: Start with

Step 2:

Step 3: output:

This above algorithm is a special case of augmented Lagrange multiplier algorithms which known as alternating directions methods [3]. Also, is a small constant > 0 and is the number of parameters which tells about the global motion model.

It is required to compute singular vectors of where we found corresponding singular values exceed the threshold . So, the important part for above algorithm is choice of properly and the terminating criterion. For this work, we choose , as suggested in [3]. We use converging condition for the algorithm as follow , with .

**Application**

There are many real-life business areas where we can model the original data as a low-rank plus a sparse component. This RPCA model can be properly fitted for the many statistical applications. Here we discuss some examples from computer science domain, and the objective purpose of the applications depends on either the low-rank part or the sparse part.

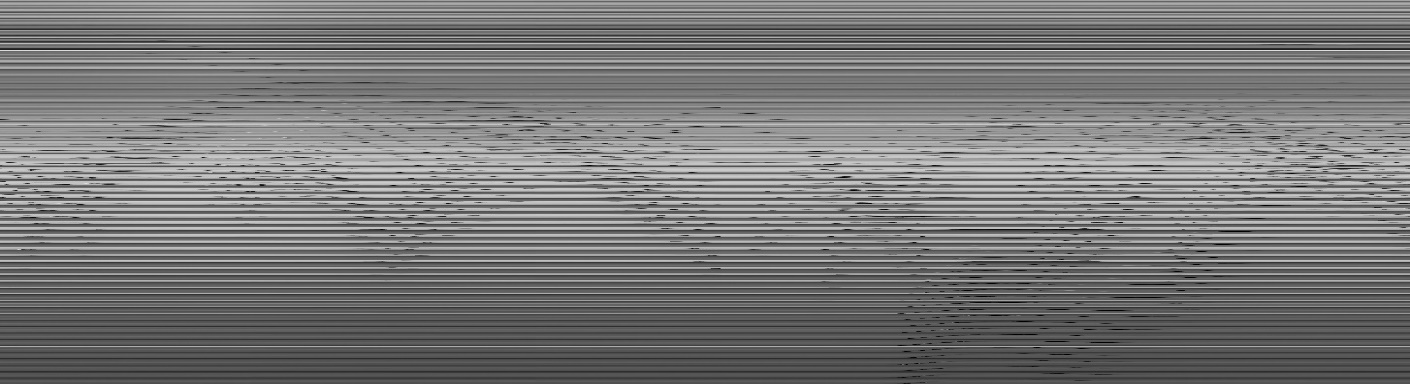
* Video Surveillance: We can track any activity from a background by providing a sequential stack of surveillance video frames. In this case we just need to do stack all the video frame within the original matrix , afterwards the low-ranked matrix and the sparse component provide us the background part and moving objects in the foreground respectively. Though video segment consists of hundreds or thousands of frames and each image frame also contains many thousands of pixels. We can decompose the data matrix if we provide a scalable solution to this problem. We will show some results on video decomposition in result section.
* Face Recognition: We know that images with belong to a convex or Lambertian surface span a low-dimensional subspace. Thats why low-dimensional models are mostly effective for image processing data. We can fairly analyze images of a human’s face through a low-dimensional subspace. To retrieve this proper subspace is fateful for many applications mainly face recognition and alignment. Often the decomposition of a realistic face images tends to be a difficult task due to presence of sufficiency in brightness, specularities, heavy noise or self-shadowing within the image. In result section, we will show some results on image decomposition.
* Latent Semantic Indexing: The content of a large corpus of documents frequently needs to be analyzed and indexed for web search engines. For this commonly used method is the Latent Semantic Indexing. The main idea is to assemble a document-versus-term within the original matrix where entries will be encoded the relevant term or word to a document such as how frequently it appears within the document. To decompose the matrix as a low-rank matrix with a residual part, which need not to be a sparse, we can use either PCA or SVD technique. If we break down the matrix M into a low-rank matrix and a sparse component , then will catch the common keywords used in all the documents and seizes few keywords whose are the best characterize each document from others.
* Ranking and Collaborative Filtering: The problem of anticipating user tastes is gaining increasing importance in online commerce and advertisement. Now a days companies collect user rankings, feedbacks after all activities on various products like commodities, books, movies, games, web tools etc. For example, the netflix prize for movie ranking is one of them. In this scenario, the purpose of the problem is to provide best preference for a user on any of the products after predicting the incomplete ranking data provided by the user. This problem acts like a low-rank matrix completion problem. However, there is frequent scarcity of control while we pass through the data gathering process. As a result, a small portion of data matrix may be noisy and tampered. Here we need to complete the matrix as well as correct the errors simultaneously. Means, from an incomplete and corrupted data matrix we need to deduce a low-rank matrix .

We can solve many other problems in graphical model learning, bioinformatic data, linear system identification, and coherence decomposition in optical systems by solving this decomposition problem.

**Results**

* Video

video is a combination of photos where lots of photo passes within a second. That’s the reason we see the movement of the objects. Hence, we can store a video in a matrix form. Matrix dimension is very high. Suppose a video of resolution m × n. So, for 1 frame of photo the matrix size will be 1 × m\*n. Let the duration of the video is t and frame rate (fps) is f, Then the total size of the matrix will be t\*f × m\*n. The original input matrix looks like as bellow.



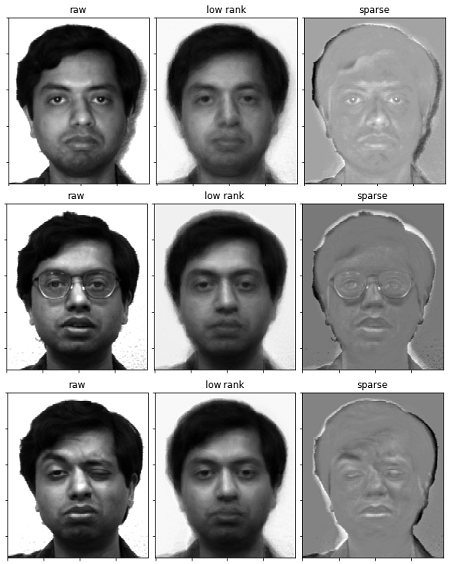
Here the wavy black lines mean movement of moving object and horizontal line represents

Here we checking different frame of our input video then easily we can differentiate background (low rank) and foreground(sparse) matrix, which looks like as follow



* Image

We used some noisy data as an original matrix then low rank matrix depicts the original matrix approximately and noise part stored in sparse matrix. Here we figured out some examples below



**Conclusion**

This paper shows some evident facts, one can detach the sparse and low-rank components exactly by convex programming. This may work under very broad conditions that are much bigger than those provided by the best outcomes. To add on, our analysis has shown rather close relationships between matrix recovery (from sparse errors) and matrix completion. Further, our results even generalize to the fact when there are both corrupted and incomplete entries. Also, there is no free parameter in Principal Component Pursuit and we can get results by simple optimization algorithms with good accuracy and efficiency. It is very important to understand that, our results may point to a wide spectrum of new algorithmic and theoretical issues together with new well-known applications that can now be analyzed systematically.

**References**

[1] Emmanuel J. Candes, Xiaodong Li, Yi Ma, and John Wright, ` “Robust principal component analysis?,” J. ACM, vol. 58, no. 3, pp. 11:1–11:37, June 2011. 1, 2, 3

[2] Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of a corrupted low-rank matrices. Mathematical Programming, submitted, 2009.

[3] X. Yuan and J. Yang. Sparse and low-rank matrix decomposition via alternating direction methods. preprint, 2009.

[4] Erfanian Ebadi, S; Guerra Ones, V; Izquierdo, E (2015); Efficient Background Subtraction with Low-rank and Sparse Matrix Decomposition, IEEE International Conference on image processing.

[5] Torre, F. D. L. & Black, M. (2003). A framework for robust subspace learning, International Journal on Computer Vision pp. 117–142.

[6] Becker, S., Candes, E. & Grant, M. (2011). Tfocs: Flexible first-order methods for rank minimization, Low-rank Matrix Optimization Symposium, SIAM Conf. on Optimization.

[7] Ding, X., He, L. & Carin, L. (2011). Bayesian robust principal component analysis, IEEE Transaction on Image Processing.

[8] Xiaowei Zhou, Can Yang, and Weichuan Yu, “DECOLOR: Moving object detection by detecting contiguous outliers in the low-rank representation.,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 3, pp. 597–610, 2013. 1

[9] Tianyi Zhou and Dacheng Tao, “GoDec: Randomized lowrank and sparse matrix decomposition in noisy case,” in Proceedings of the 28th International Conference on Machine Learning (ICML-11), Lise Getoor and Tobias Scheffer, Eds. June 2011, ICML ’11, pp. 33–40, ACM. 1, 2, 3

[10] Oliver, N., Rosario, B. & Pentland, A. (1999). A bayesian computer vision system for modeling human interactions, International Conference on Vision Systems, ICVS 1998.

[11] Get the dataset here

Image - <http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html>

Video - <http://backgroundmodelschallenge.eu/>