Comparative Study of PPCA and RPCA-EM

Group 11

School of Engineering and Applied Science,

Ahmedabad University

Courses: Machine Learning, Algorithms and Optimization for Big Data

Abstract—As we all know that EM and RPCA can recover the missing data from the given input but they can not completely recover the missing data from this input. They leaves some error after completing the process. And so they creates some holes which are the error of the data. In this report we compare the performance of RPCA algorithm and EM algorithm in terms of time taken and error.

I. INTRODUCTION

We compare the RPCA and EM algorithm for recovery of data from missing and also from the corrupted data. For the missing data we make some part of the data of the image disappear and then perform both algorithm on it to recover the missing value, and for corrupted data we actually try to recover the original data of the image.

II. SYSTEM MODEL AND ALGORITHM

- Input image for missing data we actually disappear some part of the image at some point and then try to recover the data.
- Window size for the both algorithm is 256.
- And for the corrupted image we actually added some text in the image and then try to recover the image which does not have that text on the image as before.
- And then we also calculate the error and time taken to complete the algorithm to finish.

A. PPCA-EM algorithm flow

First of all we take the image of which we want to recover the missing or corrupted data. After taking the image we created the segmentation using the window size. After the segmentation part we compute PCA using the PPCA-EM algorithm with iteration and rank. And then we aggregate the segments to recover the image. And after recovering the image we find the error between original and recovered image. And as output we have recovered image and time taken by the algo and error between the images.

B. RPCA algorithm flow

First of all we take the image of which we want to recover the missing or corrupted data. After taking the image we created the segmentation using the window size. After the segmentation part we compute PCA by using RPCA with respect to iteration and rank. And then we aggregate the segments to recover the image. And after recovering the image we find the error between original and recovered image. And as output we have recovered image and time taken by the algo and error between the images.

III. RESULTS AND OTHER INFORMATION.

Figure 1: using EM algorithm | T=2.8048 s, Error=0.0101

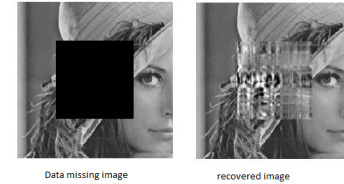Figure 2: from RPCA algorithm| T=56.7897s, Error=0.0291

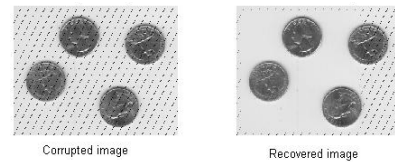Figure 3: using EM algorithm| T=4.5996s, Error=0.0156

Figure 4: from RPCA algorithm| T=66.901s, Error=0.0781

IV. CONCLUSIONS

From, this we can conclude that EM guarantees an improvement in performance as it iterates further. And RPCA guarantees iteration till converges.

REFERENCES

[1] "Robust Principal Component Analysis", E. Candes, 2009
[2] "Principal Component Analysis",I. Jolliffe. Springer-Verlag, 1986.