

Comparitive Study of RPCA to EM algorithm

Group 8, Team EnvyUs
School of Engineering and Applied Science,
Ahmedabad University
Courses: Machine Learning, Algorithms and Optimization for Big Data
March 20, 2017

Abstract—In this report, we compare the performance of RPCA algorithm to EM Algorithms, in terms of Accuracy and Time Taken.

Index Terms—PPCA, RPCA, EM

I. INTRODUCTION

We compare RPCA and EM algorithms for recovery of data from *missing* and also from *textit*corrupted values.

II. SYSTEM MODEL AND ALGORITHM

- Input image: missing/corrupted entries with $X\%$ of missing entries
- Window size $W_s = 256$ for both algorithms
- On the corrupted matrix, run recovery from RPCA and EM algorithms.
- Note the time taken, accuracy and number of iterations for both algorithms
- Process Flowchart:

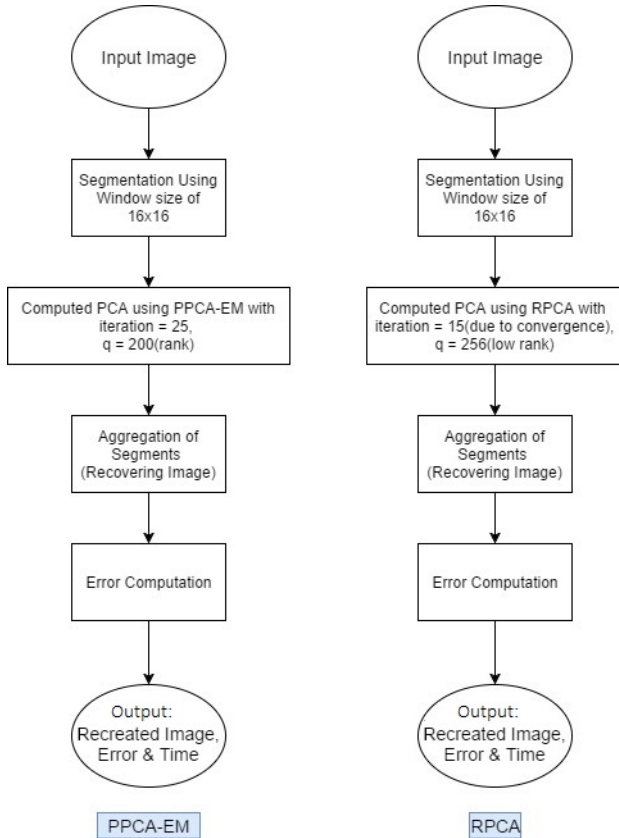


Figure 2.1: Comparison of EM-RPCA processes

III. COMPARING THE ALGORITHMS

For *missing* values;

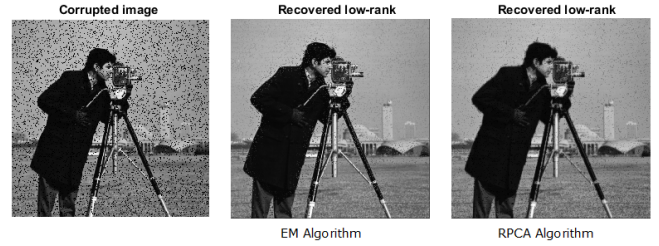
Miss%	RPCA (E_{RMS})	EM (E_{RMS})	RPCA (t_s)	EM (t_s)
1	0.0131	0.0159	1.2021	2.3698
5	0.0316	0.0359	2.3003	3.0371
10	0.0534	0.0534	3.2379	2.9605
20	0.1789	0.0861	4.4058	2.91
25	0.2341	0.1041	2.3487	2.95
50	0.3554	0.2120	1.9874	2.98

For *corrupted* values;

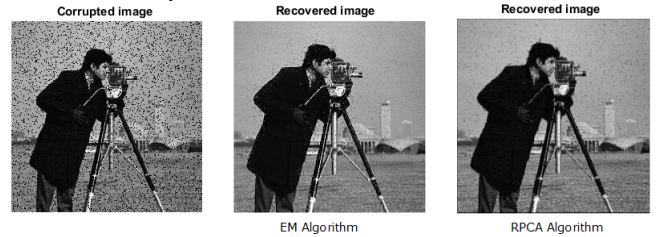
Miss%	RPCA (E_{RMS})	EM (E_{RMS})	RPCA (t_s)	EM (t_s)
1	0.1063	0.1007	3.924	5.823
5	0.1079	0.1038	5.359	5.559
10	0.1156	0.1093	11.360	6.223
20	0.6452	0.1258	7.882	5.778
25	0.7704	0.1376	3.675	8.014
50	0.9625	0.2281	0.2271	7.866

IV. RESULTS

For 10% Missing Entries:



For 10% Corrupt Entries:



V. CONCLUSION

EM guarantees reduction in error regardless of missing % of entries, but RPCA doesn't. The running time of EM is almost always lesser than RPCA algorithm. Hence we can provide that EM is certainly better than RPCA.

REFERENCES

- [1] "Robust Principal Component Analysis", E. Candes, 2009
- [2] Tipping, Michael, and C. Bishop. *Mixtures of probabilistic principal component analyzers.*, Neural computation 11.2 (1999): 443-482.