

Comparision of Robust Principal Component Analysis and Probabilistic Principal Component Analysis

Ashutosh Kakadiya, Harsh Patel, Nisarg Tike, Prerak Raja, Riddhesh Sanghvi

School of Engineering and Applied Science
Ahmedabad University

Abstract—Principal component analysis (PCA) is a technique used to emphasize variation and bring out strong patterns in a dataset. It is used for dimensionality reduction and provides compact feature for representing set of data. Here, the main purpose is to generate original image from corrupted and missing entry image. Two methods are proposed which are Robust Principal Component Analysis and Probabilistic Principal Component Analysis. We have done comparision of both on the basis percent of missing and corrupted values and listed down error values for both the methods. **Keywords:** RPCA, Factor Analysis, PPCA, Expectation Maximization (EM).

I. ROBUST PCA

Suppose we have a data matrix, which is the superposition of a low-rank component and a sparse component, it is possible to recover both the low-rank and the sparse components exactly by solving a convex program called Principal Component Pursuit.

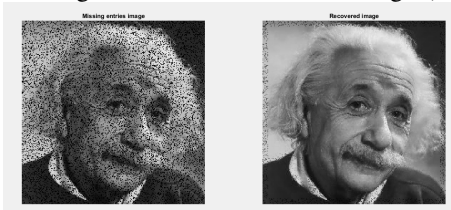
II. PROBABILISTIC PCA

Probabilistic Principal Component Analysis addresses limitations of classical PCA. PPCA can be used as a general Gaussian density model in addition to reducing dimensions. Maximum-likelihood estimates can be computed for elements associated with principal components and it captures dominant correlations with few parameters, multiple PCA models can be combined as a probabilistic mixture. Expectation Maximization is one of the algorithm in Probabilistic Principal Component Analysis for which we have generated the results.

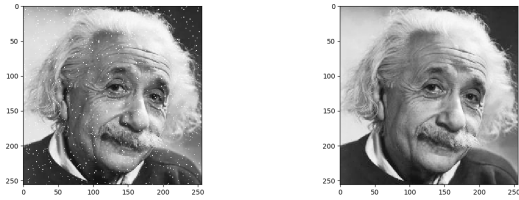
III. COMPARISON

We focused on 2 parameters, accuracy and time required for recovery of the image. These parameters are compared on both matrix corruption and matrix completion.

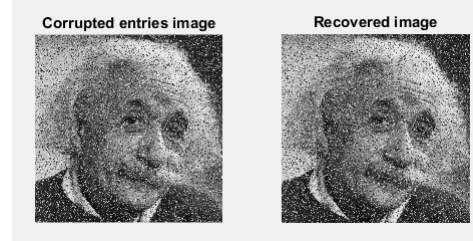
20% Missing Entries → Recovered Image (RPCA)



20% Missing Entries → Recovered Image (PPCA)



20% Corrupt Entries → Recovered Image (RPCA)



20% Corrupt Entries → Recovered Image (PPCA)



Corrupted Entries	RPCA Err %	PPCA Err %		
5%	17.03	8		
10%	23.92	11		
20%	33.33	17		
25%	37.01	19		
50%	52.35	27		
Missing Entries	RPCA Err %	PPCA Err %	RPCA time	PPCA time
5%	4.29	0.36	10.68s	8.42s
10%	6.05	0.4	11.40s	9.1s
20%	9.83	0.6	11.72s	9.6s
25%	11.37	0.7	12.00s	9.8s
50%	68.27	1	10.03s	11.23s

REFERENCES

- [1] E.J Candes, Xiaodong Li, Yi Ma, and John Wright. Robust Principal Component Analysis.
- [2] Tipping, M. E., and C. M. Bishop. Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society. Series B (Statistical Methodology), Vol. 61, No.3, 1999, pp. 611622.
- [3] Josep Porta, Jakob Verbeek, and Ben Krose. Active appearance-based robot localization using stereo vision. IVI - Instituut voor Informatica