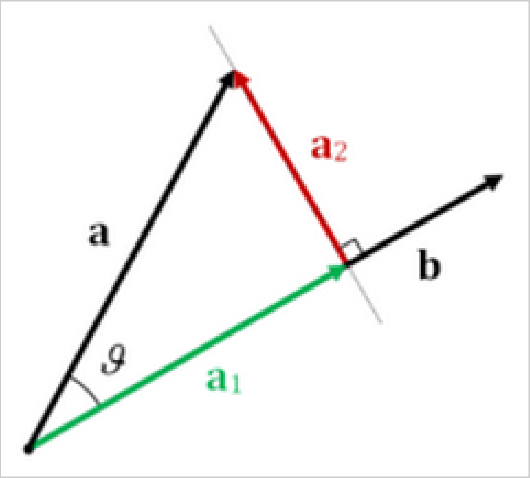
Robust PCA

Yingjun Wu

Preliminary: vector projection



Projection of **a** on **b** ($\mathbf{a}_1$), and rejection of **a** from **b** ($\mathbf{a}_2$).

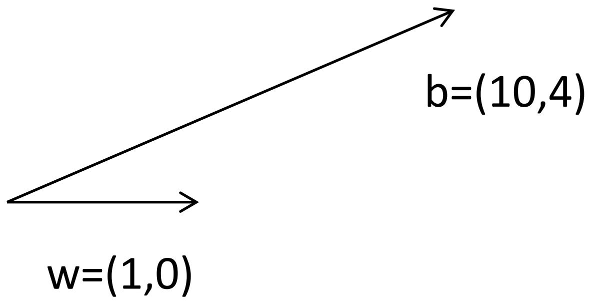
Scalar projection of a onto b:

$$a_1 = |\mathbf{a}| \cos\theta = \mathbf{a} \cdot \hat{\mathbf{b}}$$

a1 could be expressed as:

$$\mathbf{a}_1 = (\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}$$

Example



The vector projection could be calculated as:

$$(b \cdot w)w$$

Preliminary: understanding PCA

Preliminary: methodology in PCA

- Purpose: project a high-dimensional object onto a low-dimensional subspace.
- How-to:
  - Minimize distance;
  - Maximize variation.

Preliminary: math in PCA

- Minimize distance

Energy function

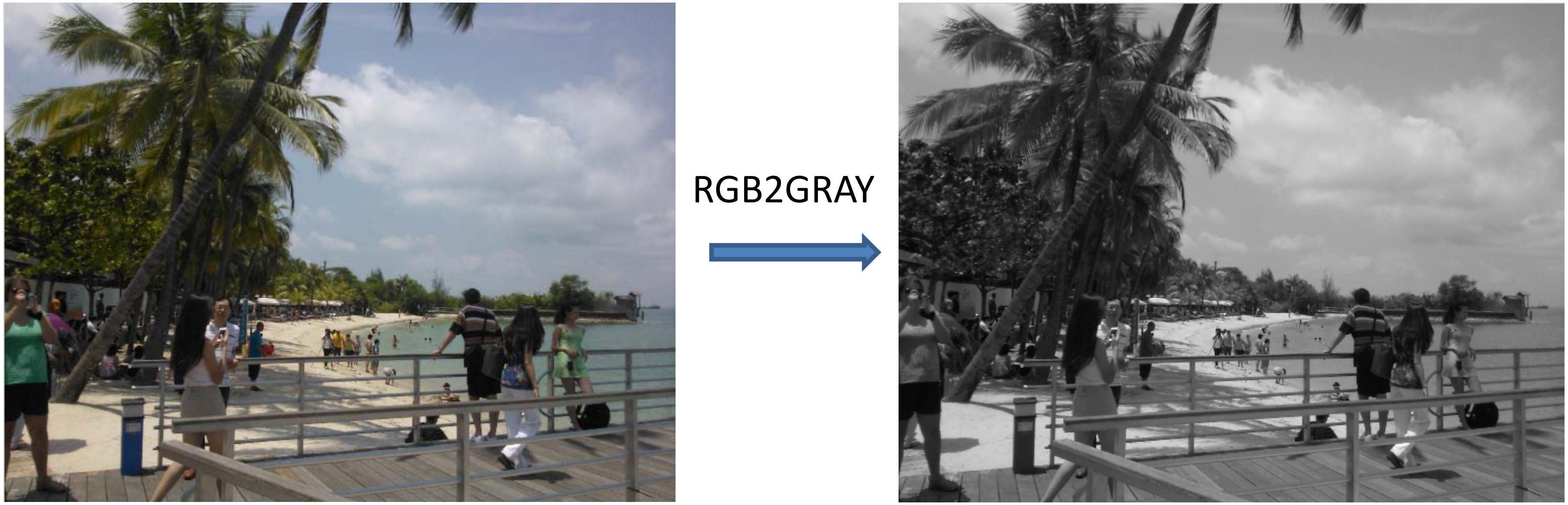
$$\begin{aligned} E_{pca}(\mathbf{B}) &= \sum_{i=1}^{n} e_{pca}(\mathbf{e}_i) = \sum_{i=1}^{n} \|\mathbf{d}_i - \boxed{\mathbf{B}\boxed{\mathbf{B}^{\mathbf{T}}\mathbf{d}_i}}\|_2^2 \\ &= \sum_{i=1}^{n}\sum_{p=1}^{d}(d_{pi} - \sum_{j=1}^{k} b_{pj}c_{ji})^2 \end{aligned}$$

Compress

Recover

Preliminary: PCA example

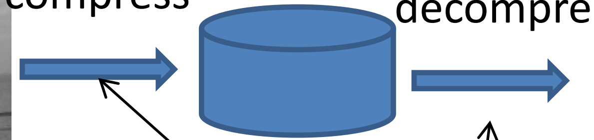
- Original figure



RGB2GRAY

Preliminary: PCA example

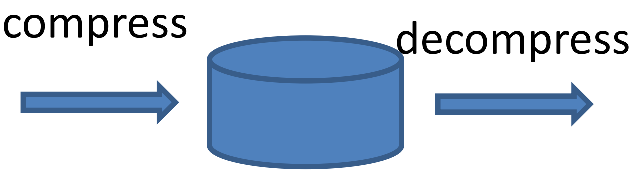
- Do something tricky:



$$
\begin{aligned}
E_{pca}(\mathbf{B}) &= \sum_{i=1}^{n} e_{pca}(\mathbf{e}_i) = \sum_{i=1}^{n} ||\mathbf{d}_i - \boxed{\mathbf{B}}\boxed{\mathbf{B}^{\mathrm{T}}\mathbf{d}_i}||_2^2 \\
&= \sum_{i=1}^{n} \sum_{p=1}^{d} (d_{pi} - \sum_{j=1}^{k} b_{pj} c_{ji})^2
\end{aligned}
$$

Preliminary: PCA example

- Do something tricky:



compress

decompress

Feature#=1900

Feature#=500

Feature#=50

Feature#=10

Preliminary: problem in PCA

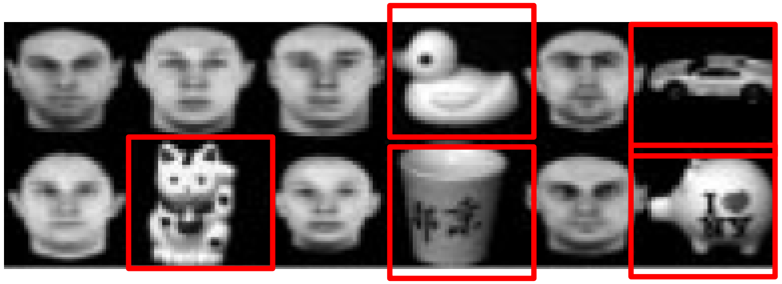
PCA fails to account for outliers.

Reason: use least squares estimation.

robust PCA

One version of robust PCA: L. Xu et.al’s work.

Mean idea: regard entire data samples as outliers.



Samples are rejected!

robust PCA

Xu’s work modified the energy function slightly and penalty is added.

$$
\begin{aligned}
E_{xu}(\mathbf{B},\mathbf{V}) &= \sum_{i=1}^{n}\left[V_i\,||\mathbf{d}_i - \mathbf{B}\mathbf{B}^{\mathbf{T}}\mathbf{d}_i||_2^2 + \boxed{\eta(1 - V_i)}\right] \\
&= \sum_{i=1}^{n}\left[V_i\Big(\sum_{p=1}^{d}(d_{pi} - \sum_{j=1}^{k} b_{pj}c_{ij})^2\Big) + \eta(1 - V_i)\right]
\end{aligned}
$$

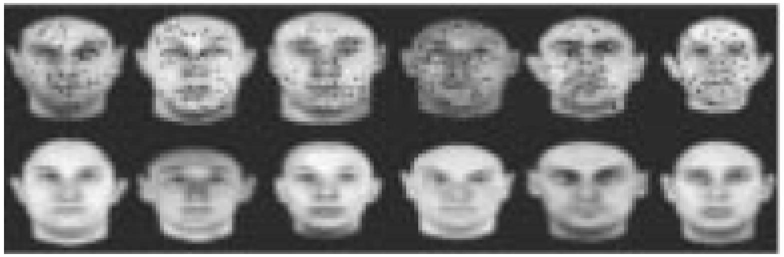
Penalty item

If Vi=1 the sample di is taken into consideration, otherwise it is equivalent to discard the sample.

robust PCA

Another version of robust PCA: Gabriel et.al’s work, or called weighted SVD.

Mean idea: do not regard entire sample as outlier. Assign weight to each feature in each sample. Outlier features could be assigned with less weight.



robust PCA

Weighted SVD also modified the energy function slightly.

$$E_{gz}(\mathbf{B},\mathbf{C}) = \sum_{i=1}^{n}\sum_{p=1}^{d} \underbrace{\boxed{w_{pi}}}_{\text{Weight}} \Big(\underbrace{\boxed{d_{pi}}}_{\text{Original feature}} - \underbrace{\boxed{(\mathbf{b}^{p})^{T}\mathbf{c}_{i}}}_{\text{Decompressed feature}} \Big)^{2}$$

robust PCA

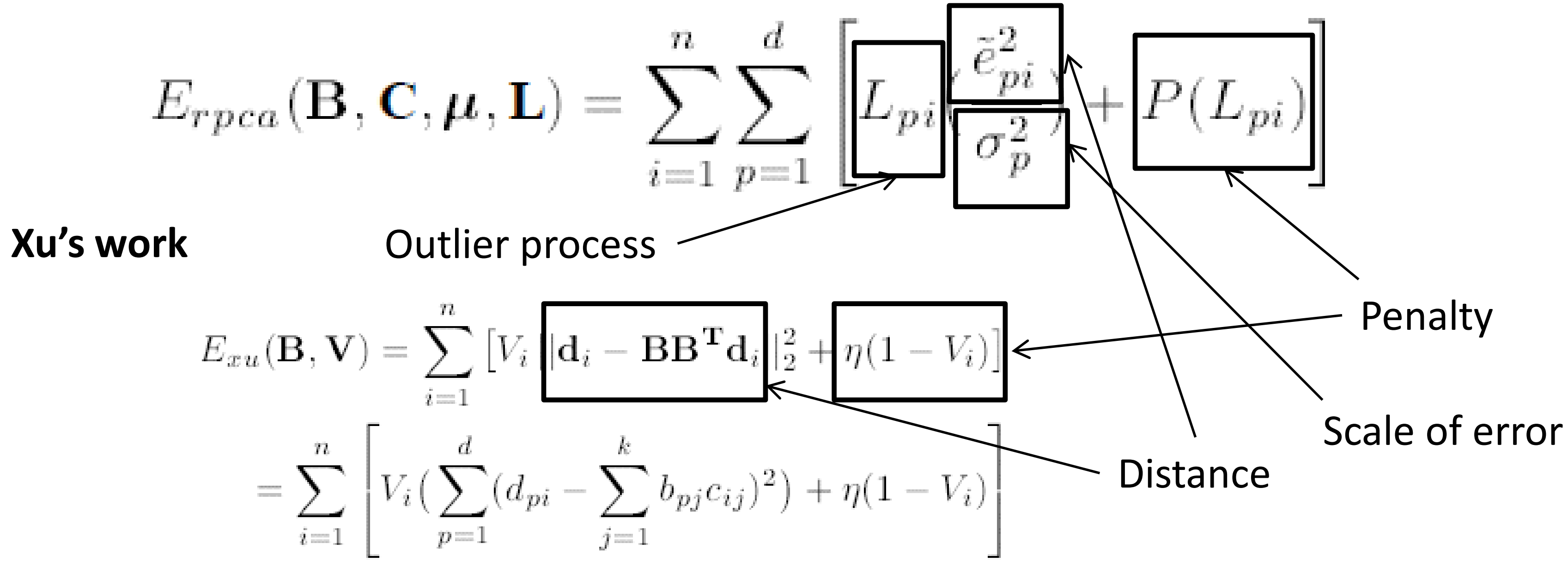
Flaw of Gabriel’s work: cannot scale to very high dimensional data such as images.

Flaw of Xu’s work: useful information in flawed samples is ignored; least squares projection cannot overcome the problem of outlier.

robust PCA

To handle the problem in the two methods, a new version of robust PCA is proposed.

Still try to modify the energy function of PCA…



$$E_{rpca}(\mathbf{B},\mathbf{C},\boldsymbol{\mu},\mathbf{L}) = \sum_{i=1}^{n}\sum_{p=1}^{d}\left[\boxed{L_{pi}}\left(\frac{\boxed{\tilde{e}_{pi}^{2}}}{\boxed{\sigma_p^{2}}}\right)+\boxed{P(L_{pi})}\right]$$

Xu's work

Outlier process

$$E_{xu}(\mathbf{B},\mathbf{V}) = \sum_{i=1}^{n}\left[V_i\,\boxed{\|\mathbf{d}_i-\mathbf{B}\mathbf{B}^{\mathbf{T}}\mathbf{d}_i\|_2^2}+\boxed{\eta(1-V_i)}\right]$$

$$= \sum_{i=1}^{n}\left[V_i\left(\sum_{p=1}^{d}(d_{pi}-\sum_{j=1}^{k}b_{pj}c_{ij})^2\right)+\eta(1-V_i)\right]$$

Penalty

Scale of error

Distance

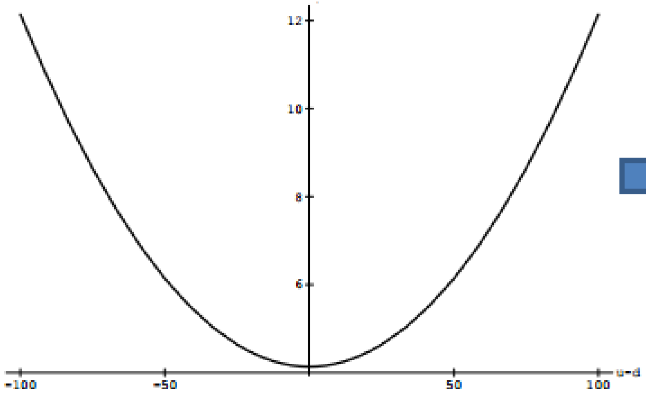
robust PCA

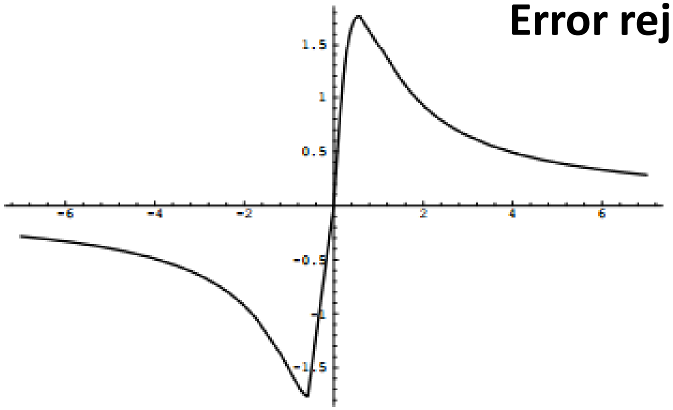
To handle the problem in the two methods, a new version of robust PCA is proposed.

Still try to modify the energy function of PCA…

$$E_{rpca}(\mathbf{B},\mathbf{C},\boldsymbol{\mu},\mathbf{L}) = \sum_{i=1}^{n}\sum_{p=1}^{d}\left[L_{pi}\left(\frac{\tilde{e}_{pi}^{2}}{\sigma_{p}^{2}}\right) + P(L_{pi})\right]$$

Increase without bound!

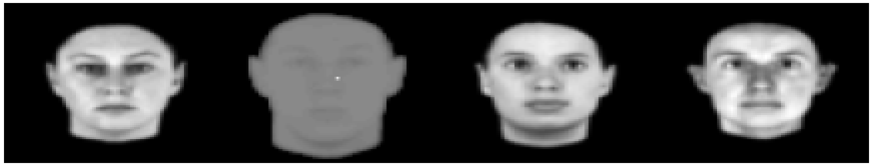




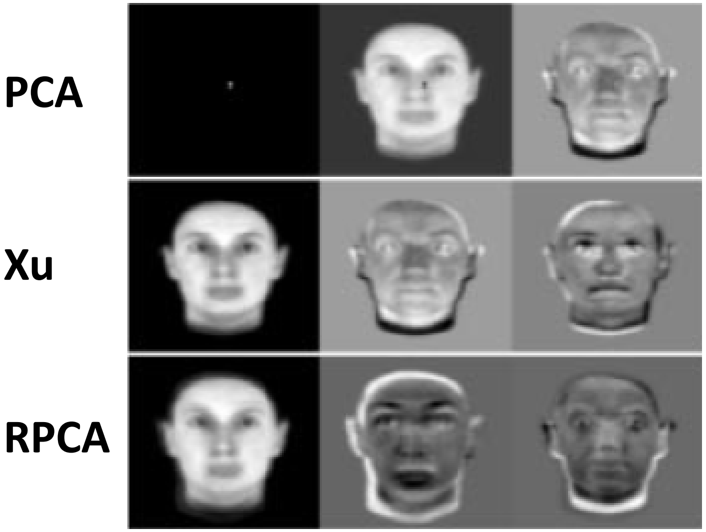
Error rejected!

Experiments

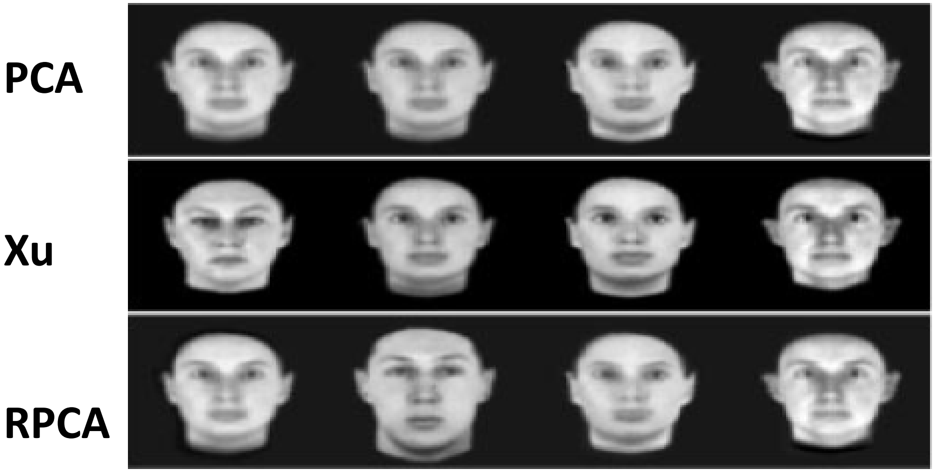
Four faces, the second face is contaminated.



Learned basis images.



Reconstructed faces.



Experiments

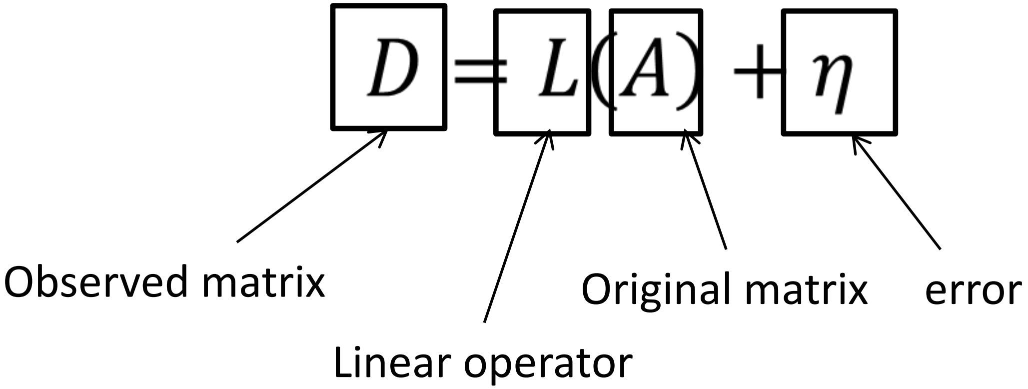
Original video

PCA

RPCA

Recent works

- John Wright et.al proposed a new version of RPCA.
- Problem: assume a matrix A is corrupted by error or noise, if we observed D, how to recover A?



$$\boxed{D} = \boxed{L}(\boxed{A}) + \boxed{\eta}$$

Observed matrix

Original matrix

error

Linear operator

Recent works

- Matrix completion
  - Try to recover the missing entries from an incomplete matrix.

$$D = L(A) + \eta$$

  - Consider L(A) as a subset of all entries of A, and η is zero.
  - Try to minimize:

$$\min_X rank\ (X) \text{ subject to } L(X)=D$$

Recent works

- Robust PCA
  - Using the idea of matrix completion.

$$D = L(A) + \eta$$

  - Consider L(A) as a identity operator, and η is a sparse matrix.
  - Try to minimize:

$$\min_{X,E} rank\ (X) + \gamma \left|\left|E\right|\right|_0$$

subject to $D=X+E$

Robust PCA demo

References

- De la Torre, F. et.al, *Robust principal component analysis for computer vision*, ICCV 2001
- M. Black et.al, *On the unification of line processes, outlier rejection, and robust statistics with applications in early vision*, IJCV 1996
- D. Geiger et.al, *The outlier process*, IEEE workshop on NNSP, 1991
- L. Xu et.al, *Robust principal component analysis by self-organizing rules based on statistical physics approach*, IEEE trans. Neural Networks, 1995
- John Wright et. al, *Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization*, NIPS 2009
- Emmanuel Candes et.al, *Robust Principal Component Analysis?,* Journal of ACM, 2011

Thank you!