SRC :
Sparse Representation Classifier

...@liris.cnrs.fr - http://liris.cnrs.fr/...

Laboratoire d'InfoRmatique en Image et Systèmes d'information

LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon
Université Claude Bernard Lyon 1, bâtiment Nautibus
43, boulevard du 11 novembre 1918 — F-69622 Villeurbanne cedex
http://liris.cnrs.fr

15/01/2018

Outline

- Formulation
- Robust Recognition
- Experiments

FORMULATION

Face recognition as sparse representation

- **Assumption:** the test image, $\boldsymbol{y} \in \mathbb{R}^D$, $D = w \times h$, can be expressed as a linear combination of k training images, say $\{\boldsymbol{y}_i^1, \ldots, \boldsymbol{y}_i^k\}$ of the same subject:

$$\boldsymbol{y} = \boldsymbol{y}_i^1 \alpha_1 + \boldsymbol{y}_i^2 \alpha_2 + \cdots + \boldsymbol{y}_i^k \alpha_k \doteq A_i \vec{\beta}_i, \quad \vec{\beta}_i \in \mathbb{R}^k$$

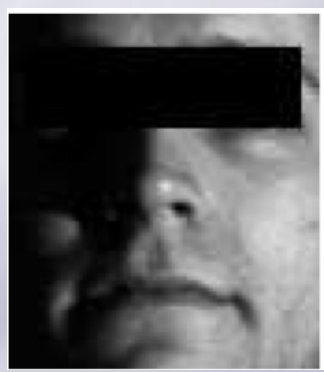
$$\boldsymbol{y} = A_1 \vec{\beta}_1 + A_2 \vec{\beta}_2 + \cdots + A_n \vec{\beta}_n = A\boldsymbol{x}$$

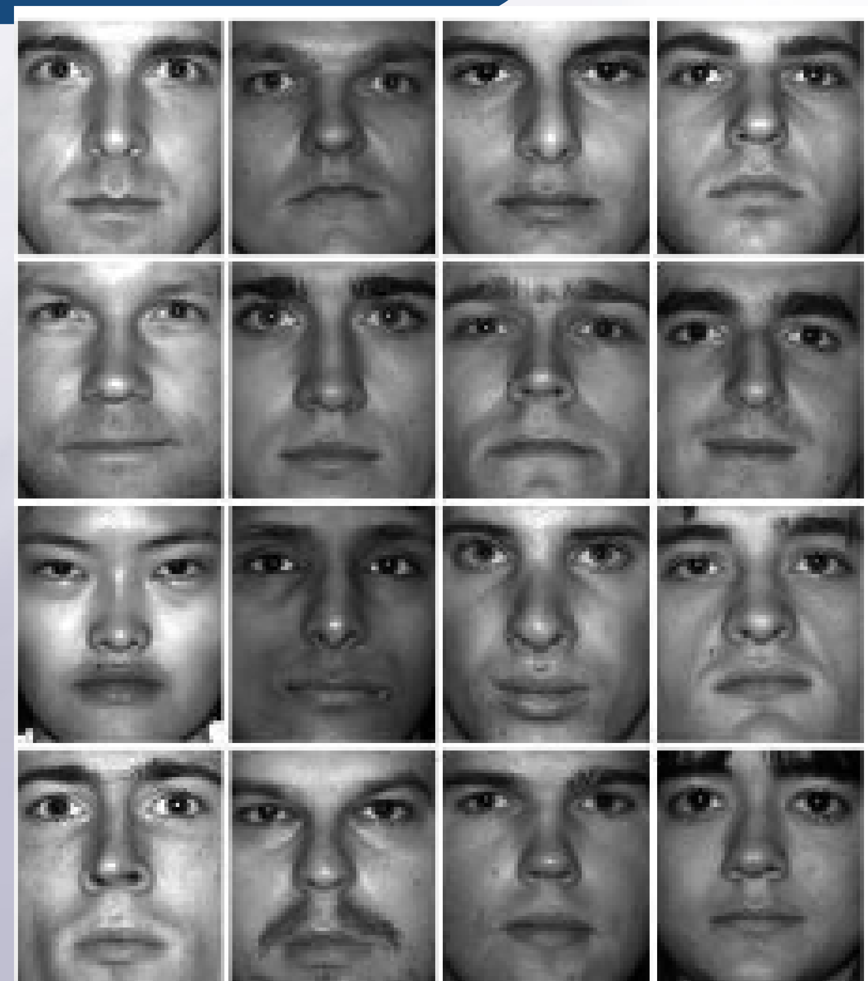
$$A = [A_1, A_2, \ldots, A_n], \quad \boldsymbol{x} = \begin{bmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_{i-1} \\ \vec{\beta}_i \\ \vec{\beta}_{i+1} \\ \vdots \\ \vec{\beta}_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ * \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{nk}$$

The solution, $\boldsymbol{x} \in \mathbb{R}^N$, $N = n \times k$, should be a sparse vector — $\frac{n-1}{n}$ of its entries should be zero, except for the ones associated with the correct subject.

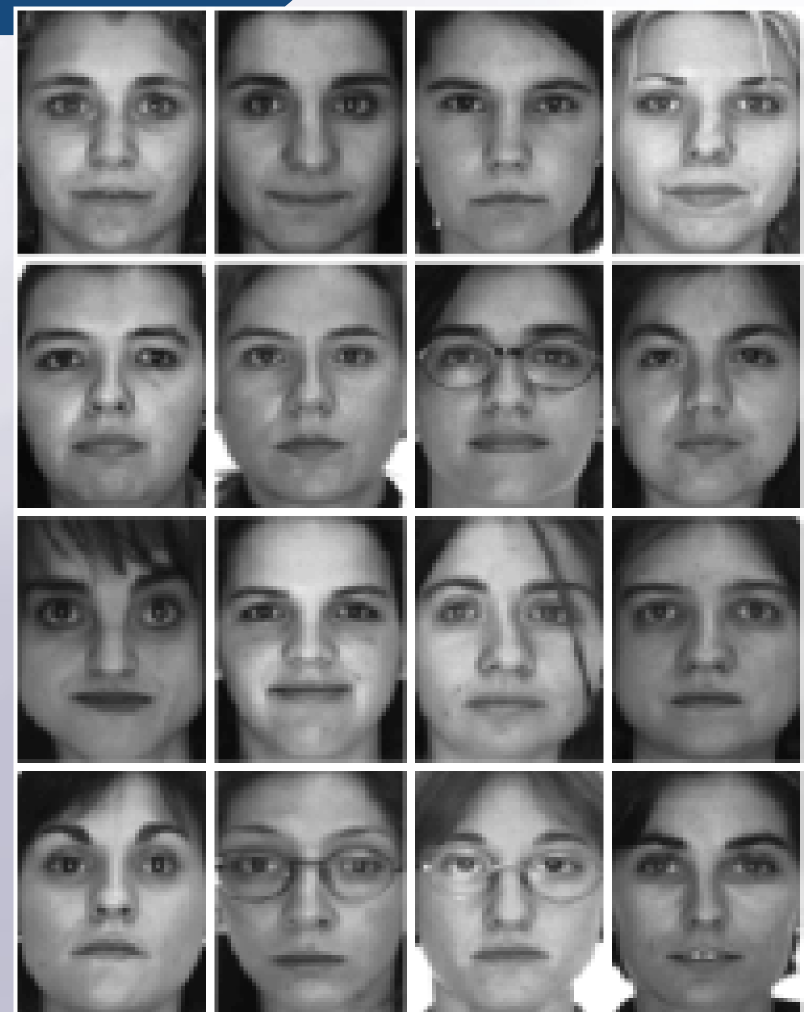
ROBUST RECOGNITION

Occlusion + varying illumination





ROBUST RECOGNITION
Occlusion and Corruption



ROBUST RECOGNITION
Tackling Corruption and Occlusion

- Properties that help to tackle occlusion:
  - Redundancy (essential for error-correcting code)

    But *nothing is more redundant than the original images*
  - Locality (using local features and parts such as ICA and LNMF)

    But *no features or parts are more local than the original pixels*
  - Sparsity (error incurred by occlusion is typically sparse)

    But *sparse representation not been thoroughly exploited in recognition*

ICA : Independent component analysis
LNMF : Local Non-negative Matrix Factorization

ROBUST RECOGNITION

Properties of the Occlusion

$y = Ax + e$ $\qquad$ $y_0 = Ax$ $\qquad$ e

- Several characteristics of occlusion :
  - Randomly supported errors (location is unknown and unpredictable)
  - Gross errors (arbitrarily large in magnitude)
  - Sparse errors? (concentrated on relatively small part(s) of the image)

ROBUST RECOGNITION

Problem Formulation

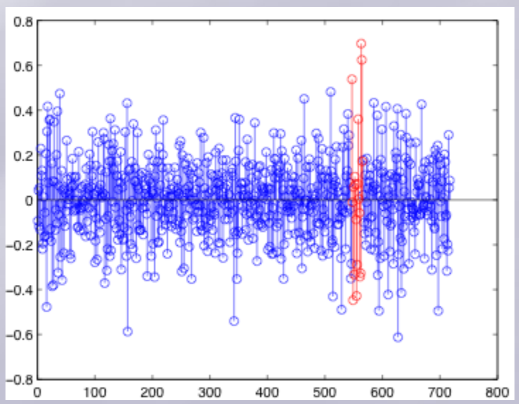
Problem: Find the correct (sparse) solution $\boldsymbol{x}$ from the corrupted and over-determined $(D \gg N)$ system of linear equations:

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{e}, \quad A \in \mathbb{R}^{D \times N}, \ \boldsymbol{y}, \boldsymbol{e} \in \mathbb{R}^{D}.$$

Conventionally, the minimum 2-norm (least squares) solution is used:

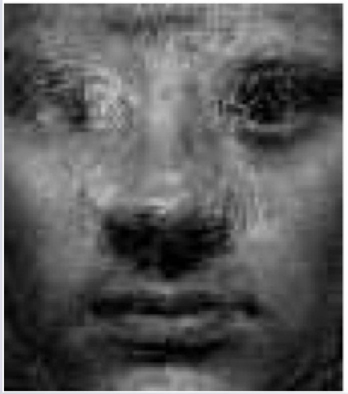
$\boldsymbol{y}$

$\hat{\boldsymbol{x}}_2$



$\hat{\boldsymbol{e}}_2 = \boldsymbol{y} - A\hat{\boldsymbol{x}}_2$

$\hat{\boldsymbol{y}}_0 = A\hat{\boldsymbol{x}}_2$



LIRIS

ROBUST RECOGNITION

Joint Sparsity

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{e} \quad \Longleftrightarrow \quad \boldsymbol{y} = [A\ I]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{e} \end{bmatrix} \doteq B\boldsymbol{q}, \quad B \in \mathbb{R}^{D\times(D+N)}.$$

Thus, we are looking for a sparse solution $\boldsymbol{q}$ to an under-determined $(D < D + N)$ system of linear equations $\boldsymbol{y} = B\boldsymbol{q}$:

$$(P_0) \qquad \hat{\boldsymbol{q}}_0 = \arg\min_{\boldsymbol{q}} \|\boldsymbol{q}\|_0 \quad \text{s.t.} \quad \boldsymbol{y} = B\boldsymbol{q}.$$

$$\Updownarrow$$

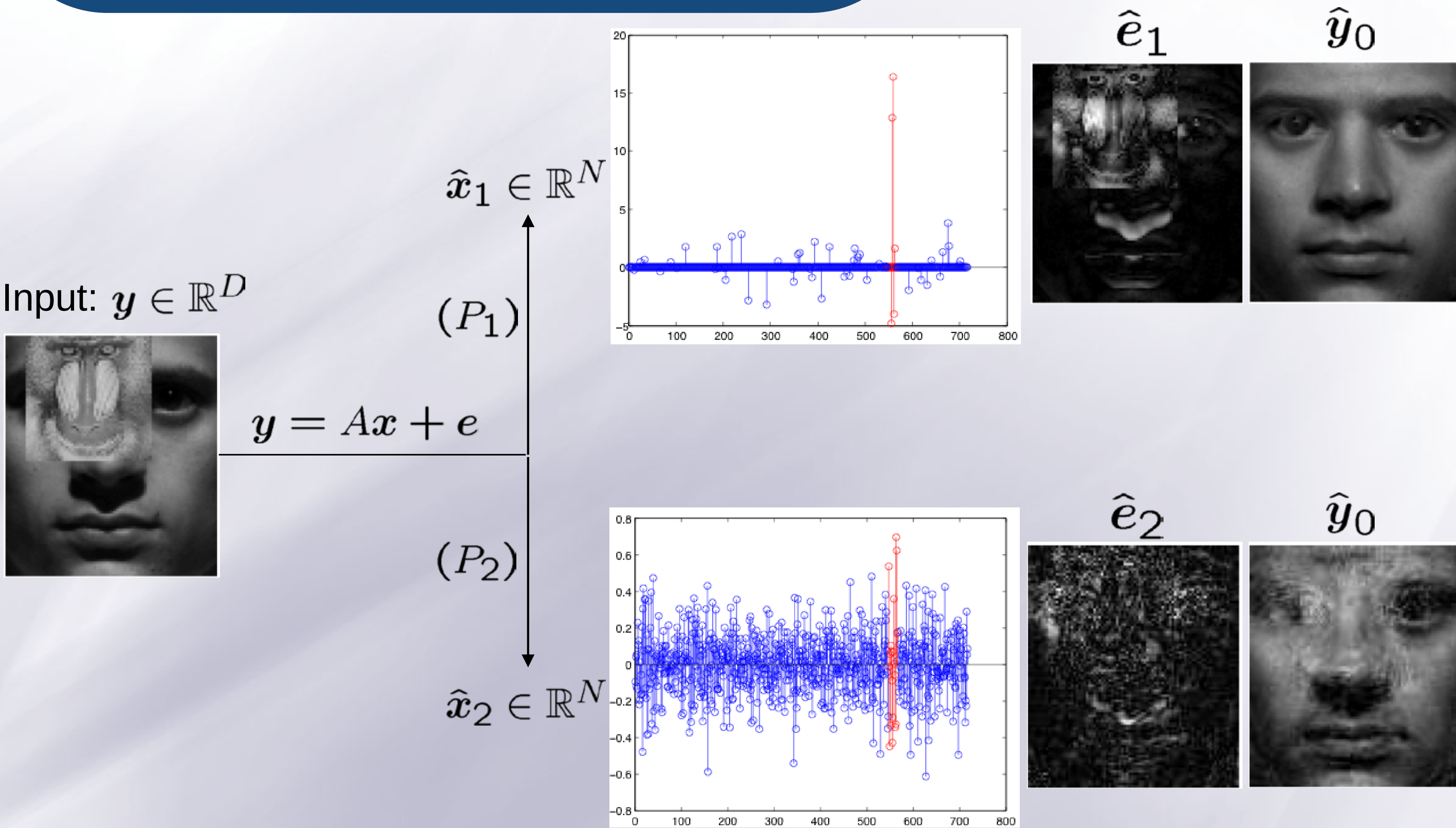
$$(P_1) \qquad \hat{\boldsymbol{q}}_1 = \arg\min_{\boldsymbol{q}} \|\boldsymbol{q}\|_1 \quad \text{s.t.} \quad \boldsymbol{y} = B\boldsymbol{q}.$$

The problem (P_1) can be solved efficiently via Linear Programming, and the solution is stable under moderate noise [Candes & Tao'04, Donoho'04].

The equivalence holds iff $\|\boldsymbol{q}\|_0 = \|\boldsymbol{x}\|_0 + \|\boldsymbol{e}\|_0 \leq \text{EBP}(B)$.

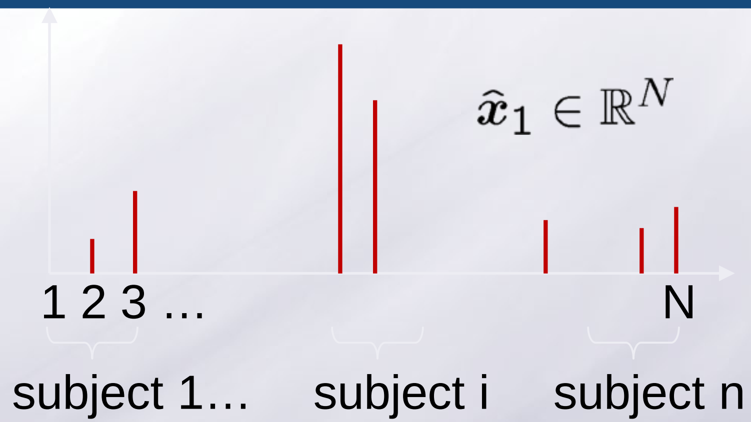
ROBUST RECOGNITION

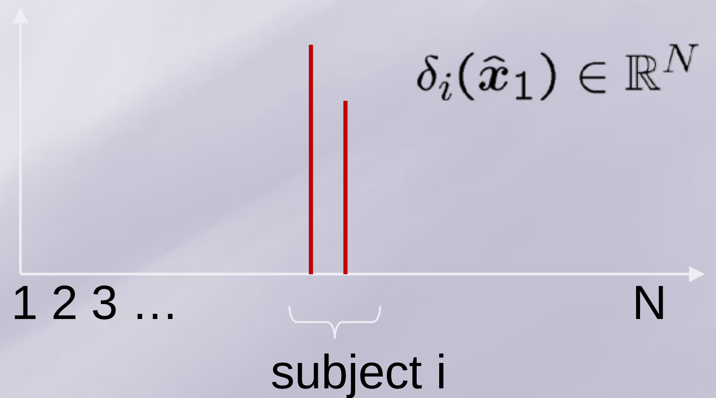
L_1 versus L_2 Solution



ROBUST RECOGNITION

Classification from Coefficients





$$r_i = \|y - A\delta_i(\hat{\boldsymbol{x}}_1) - \hat{\boldsymbol{e}}_1\|_2$$

Classification criterion: assign to the class with the smallest residual.

ROBUST RECOGNITION

Algorithm Summary

Input: $N = n \times k$ training images of size $D = w \times h$ pixels, partitioned into n classes, $A_1, \ldots, A_n$, and an (occluded) test sample $\boldsymbol{y}$.

Set $B = [A_1, A_2, \ldots, A_n, \; I] \in \mathbb{R}^{D\times(D+N)}$.

Solve the linear programming problem:

$$\hat{\boldsymbol{q}}_1 = \arg \min_{\boldsymbol{q}=[\boldsymbol{x}\,\boldsymbol{e}]} \|\boldsymbol{q}\|_1 \;\; \text{s.t.} \;\; \boldsymbol{y} = B\boldsymbol{q}$$

for $i = 1 : n$

Compute the reconstruction $\hat{\boldsymbol{y}}_0 = \boldsymbol{y} - \hat{\boldsymbol{e}}_1$.

Compute the residual $r_i = \|\hat{\boldsymbol{y}}_0 - A\delta_i(\hat{\boldsymbol{x}}_1)\|_2$.

end

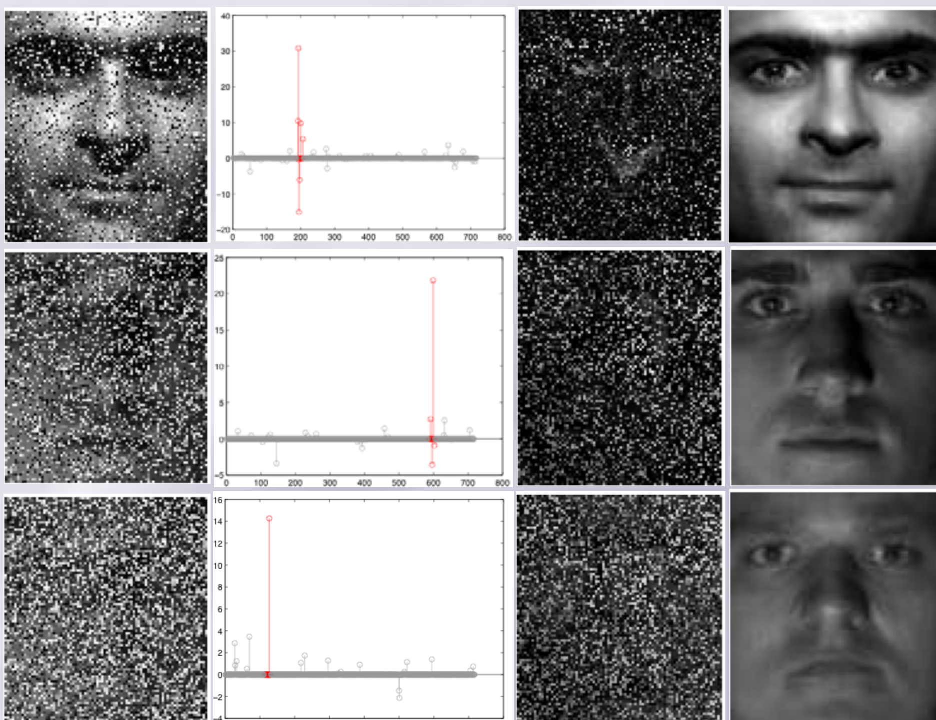
Output: $\text{id}(\boldsymbol{y}) = \arg\min_{i=1,\ldots,n} r_i$

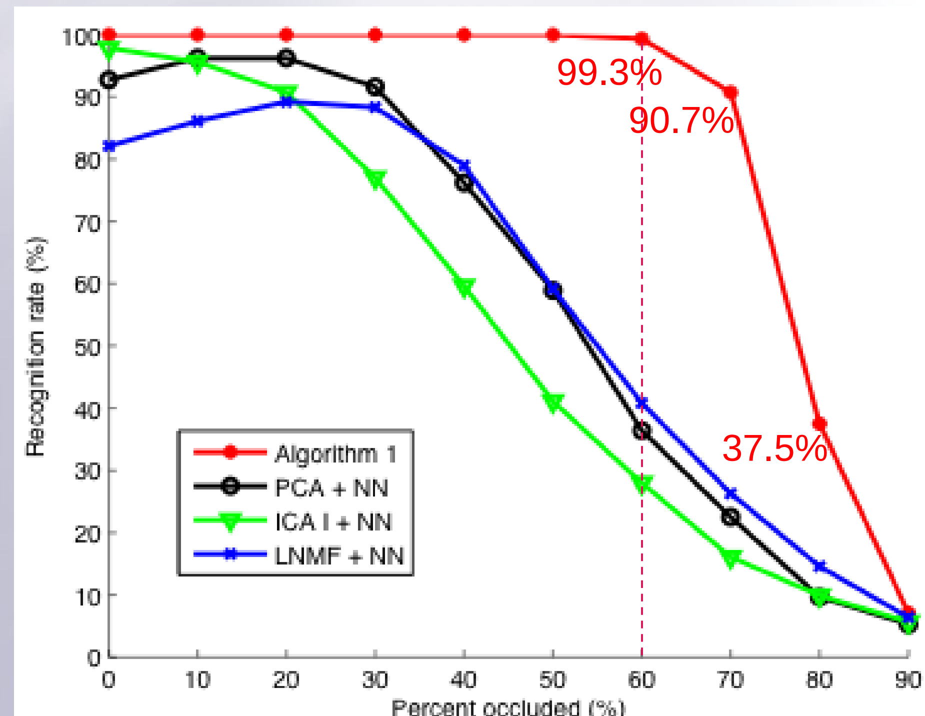
EXPERIMENTS
Varying Level of Random Corruption

Extended Yale B Database (38 subjects)

Training: subsets 1 and 2 (717 images)
Testing: subset 3 (453 images)

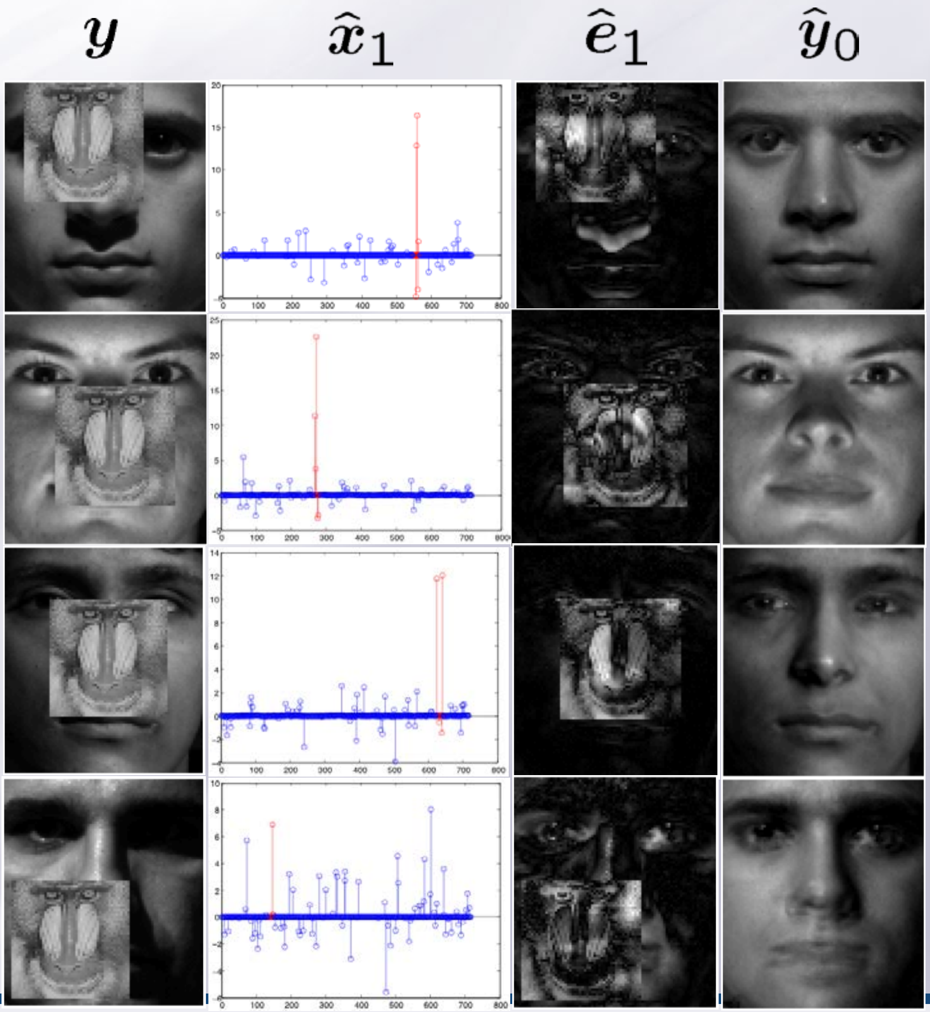
$\boldsymbol{y}$ $\hat{\boldsymbol{x}}_1$ $\hat{\boldsymbol{e}}_1$ $\hat{\boldsymbol{y}}_0$





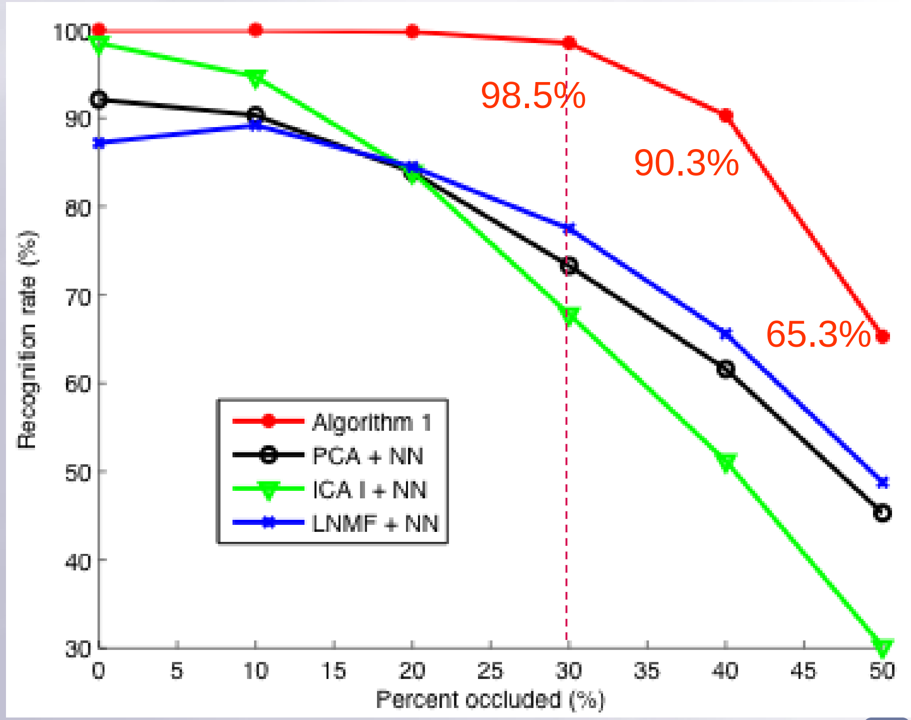
EXPERIMENTS
Varying Levels of Contiguous Occlusion

Extended Yale B Database (38 subjects)

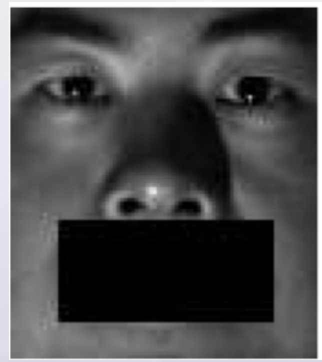
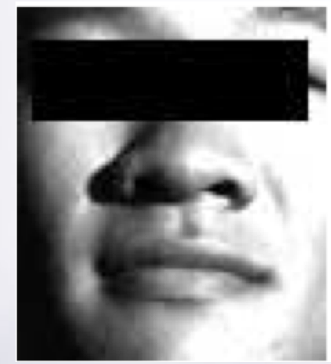


Training: subsets 1 and 2 (717 images), Error Back Projection ~ 13.3%.

Testing: subset 3 (453 images)



EXPERIMENTS
Recognition with Face Parts Occluded

Occluded	Rec. rate
Nose	98.7%
Mouth	97.1%
Eyes	95.6%

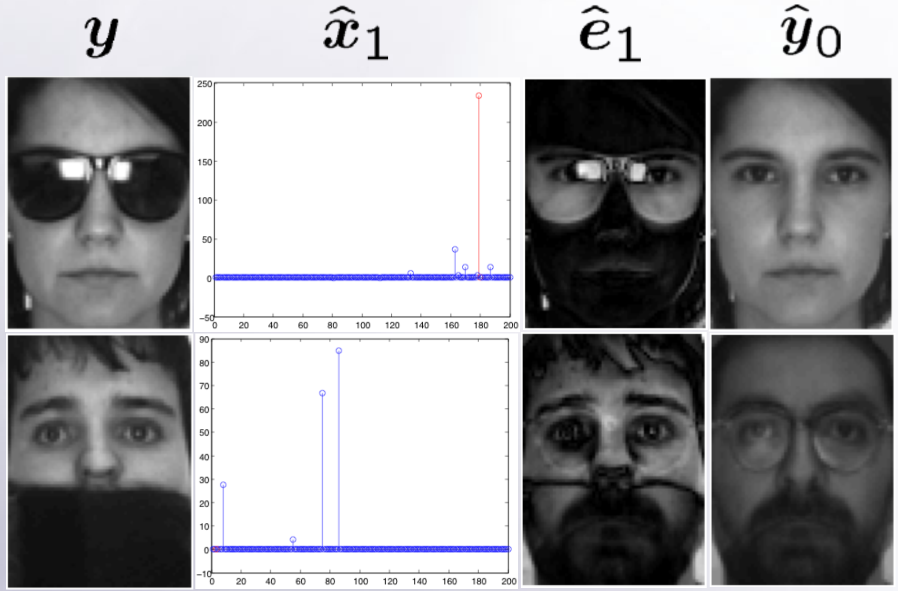
Results corroborate findings in human vision: the eyebrow or eye region is most informative for recognition [Sinha'06].

However, the difference is less significant for our algorithm than for humans.

EXPERIMENTS

Recognition with Disguises

The AR Database (100 subjects)
Training: 799 images (un-occluded)
Error Back Projection = 11.6%.
Testing: 200 images (with glasses)
200 images (with scarf)



Cases	Rec. rate	Cases	Rec. rate
Sunglasses	97.5%	Scarves	93.5%
Men	97.5%	Women	93.5%
Men, sunglasses	100%	Women, sunglasses	95.0%
Men, scarves	95.0%	Women, scarves	92.0%