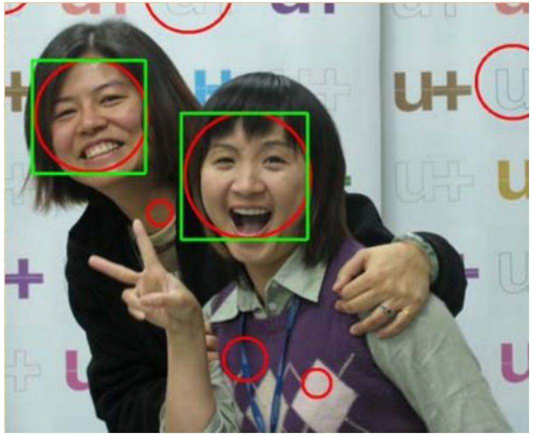
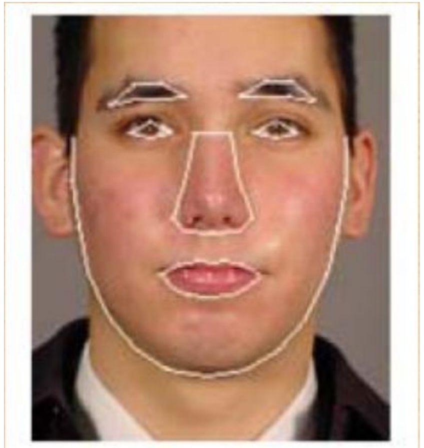
# Face Recognition via sparse representation

Mohsen Nabian

# Problems in Face



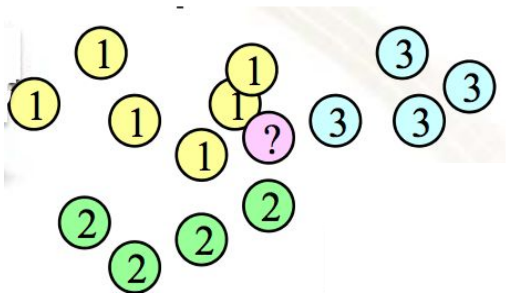


Face Detection

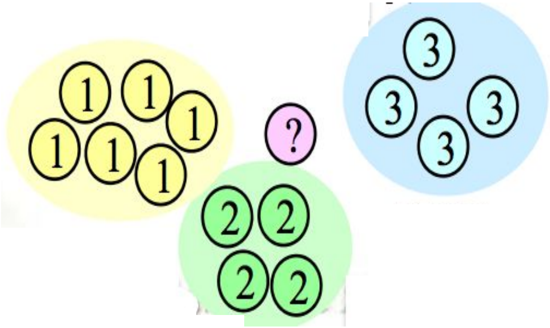
Face Alignment

Face Recognition

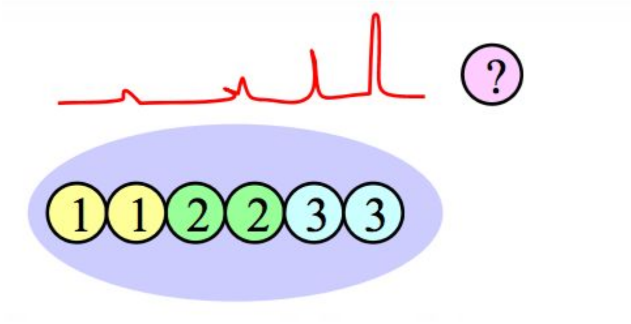
# Approaches



Nearest Neighbor
**(NN)**



Nearest Subspace
**(NS)**



Linear Sparse
Representation
**(LSR)**

# LSR Advantages:

1) Higher Classification Accuracy
2) Detects outliers
3) Not sensitive to Feature Selection
4) Robust to occlusion and corruptions

# 1) Feature Selection

**Traditional Methods (NN,NS,SVM,..) :**

1) Good feature provide more information for classification
2) A lot of research on feature selection (E.g. Eigenface, Fisherface, Laplacianface)
3) Lack of guidelines to decide which feature to use

**Recent Method(LSR) :**

The choice of features is no longer critical

# 2) Robustness to Occlusion

**Traditional Methods (NN,NS,SVM,..) :**

Occlusion is a **significant obstacle** to face recognition.

**Recent Method(LSR) :**

**robust** to Occlusion

# Symbols

**k** classes (individuals)

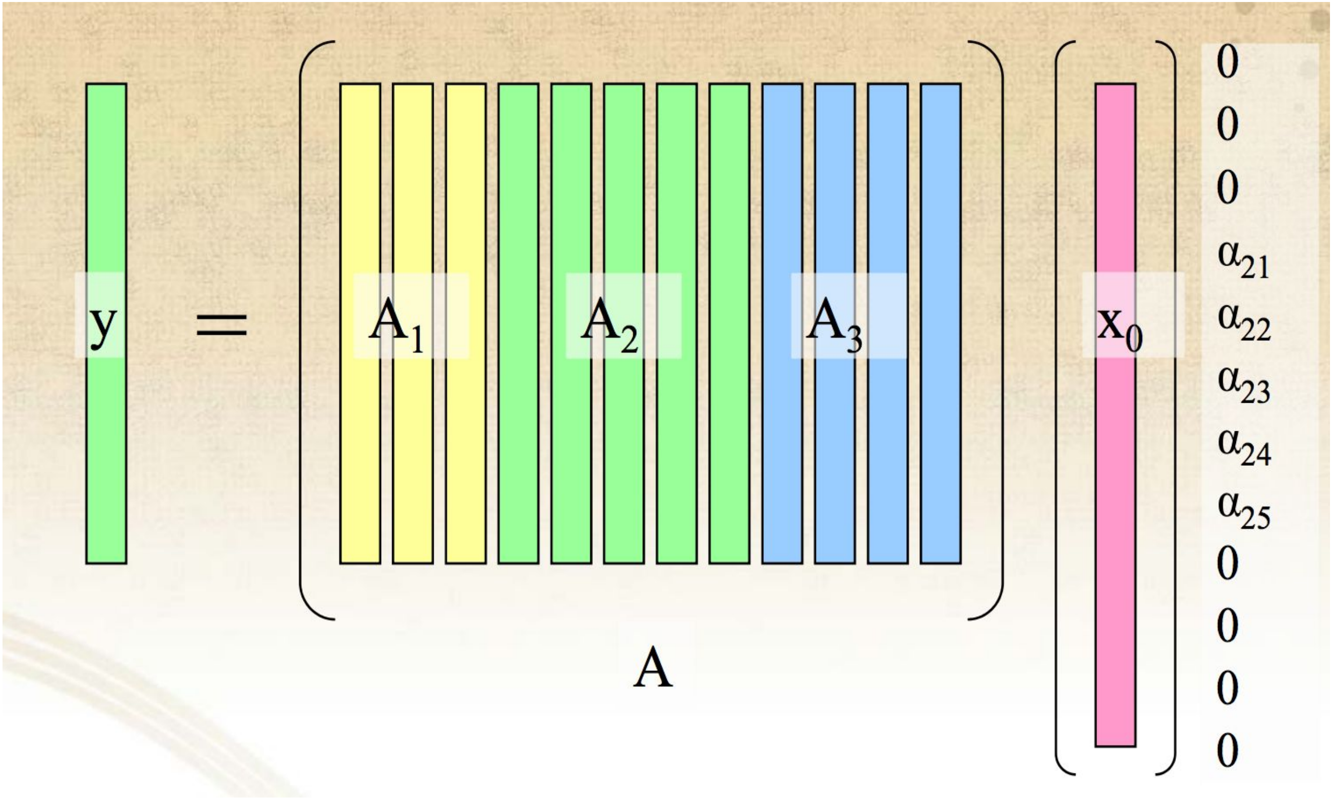
$\mathbf{n_i}$ = number of training images of person i

**n** = total number of images = $n_1+n_2+\ldots+n_i+\ldots+n_k$

Each image = w * h pixels = **m** features

$$v_{ij} \in \mathbb{R}^m \qquad A_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$$

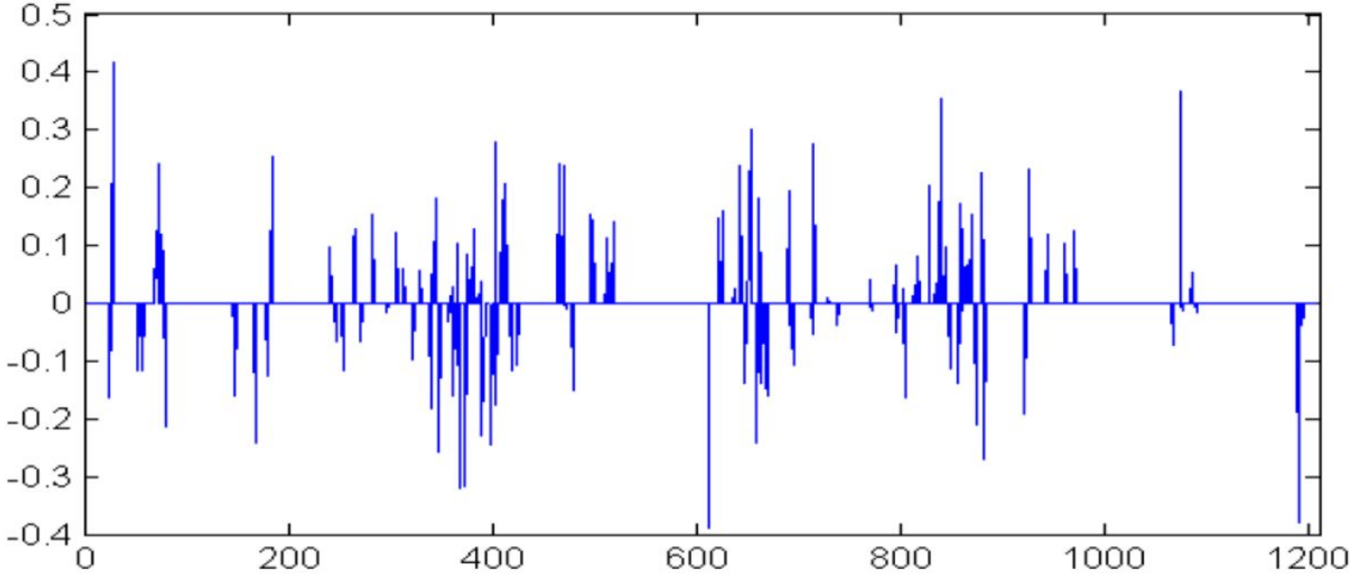
# Model




http://www.cmlab.csie.ntu.edu.tw/~known/known200611/slides

# Solve y= Ax

$$(\ell^2): \qquad \hat{\boldsymbol{x}}_2 = \arg\min \|\boldsymbol{x}\|_2 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y}$$



# Solve y= Ax (summary)

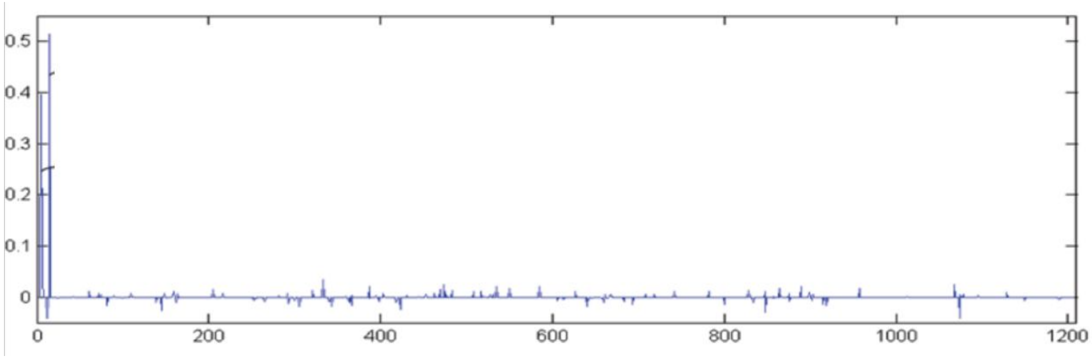
$$(\ell^2): \quad \hat{\boldsymbol{x}}_2 = \arg\min \|\boldsymbol{x}\|_2 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y}$$

$$(\ell^0): \quad \hat{\boldsymbol{x}}_0 = \arg\min \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y},$$

$$\boxed{(\ell^1): \quad \hat{\boldsymbol{x}}_1 = \arg\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y}}$$

$$\boxed{(\ell^1_s): \quad \hat{\boldsymbol{x}}_1 = \arg\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|A\boldsymbol{x} - \boldsymbol{y}\|_2 \le \varepsilon}$$

# Classification



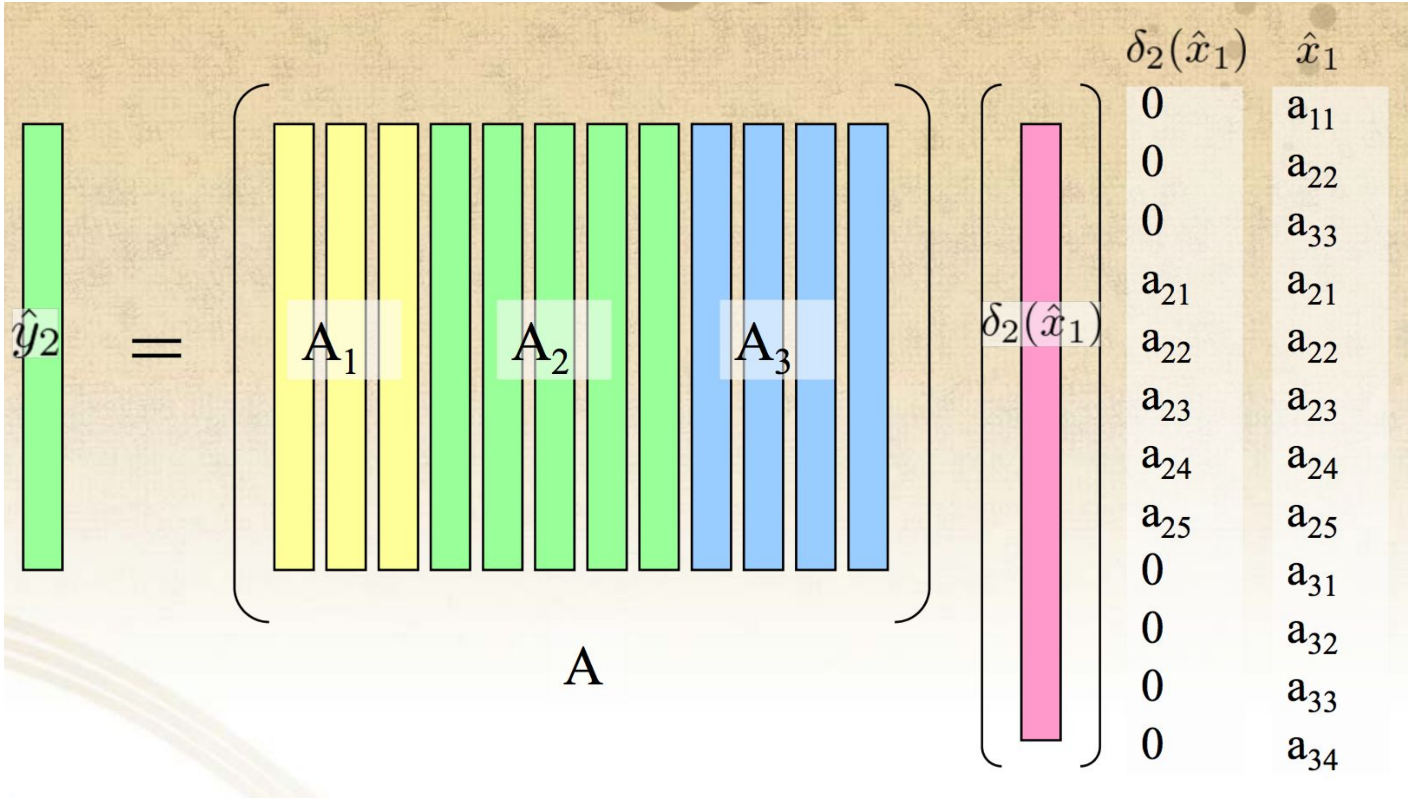

**Heuristic 1 )**

simply assign y to the object class with the single largest entry in $x_1$

**Heuristic 2 )**

$$\min_i \; r_i(\boldsymbol{y}) \doteq \|\boldsymbol{y} - A\,\delta_i(\hat{\boldsymbol{x}}_1)\|_2$$

# Heuristic 2




http://www.cmlab.csie.ntu.edu.tw/~known/known200611/slides

**Algorithm 1. Sparse Representation-based Classification (SRC)**

1: **Input:** a matrix of training samples $A = [A_1, A_2, \ldots, A_k] \in \mathbb{R}^{m \times n}$ for $k$ classes, a test sample $\boldsymbol{y} \in \mathbb{R}^m$, (and an optional error tolerance $\varepsilon > 0$.)

2: Normalize the columns of $A$ to have unit $\ell^2$-norm.

3: Solve the $\ell^1$-minimization problem:

$$\hat{\boldsymbol{x}}_1 = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \text{ subject to } A\boldsymbol{x} = \boldsymbol{y}. \qquad (13)$$

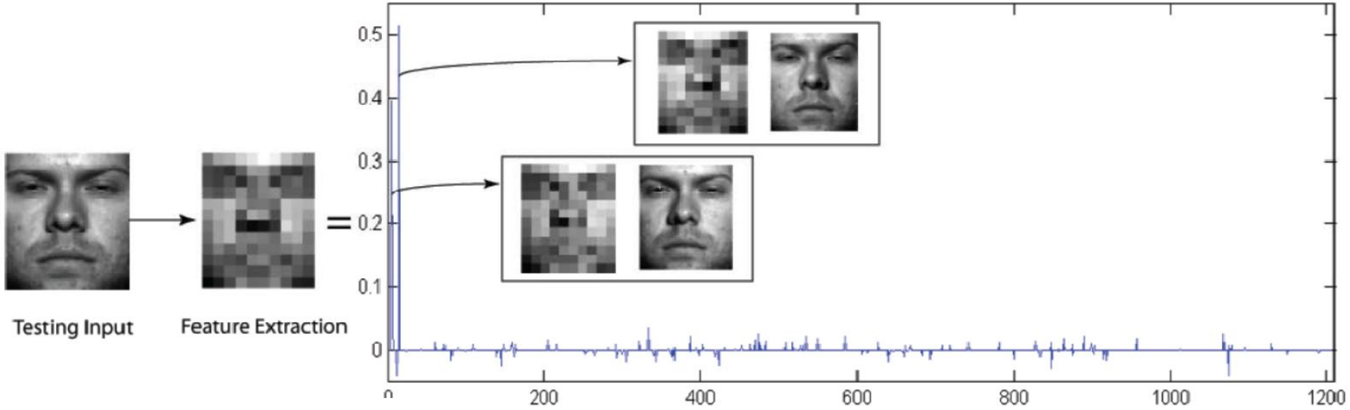
(Or alternatively, solve

$$\hat{\boldsymbol{x}}_1 = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \text{ subject to } \|A\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon.)$$

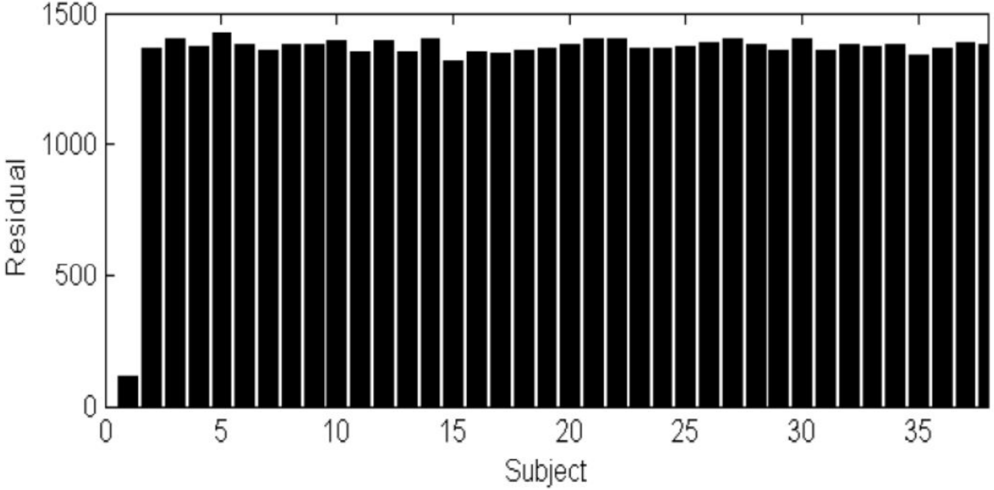
4: Compute the residuals $r_i(\boldsymbol{y}) = \|\boldsymbol{y} - A\,\delta_i(\hat{\boldsymbol{x}}_1)\|_2$ for $i = 1, \ldots, k$.

5: **Output:** $\text{identity}(\boldsymbol{y}) = \arg\min_i r_i(\boldsymbol{y})$.

# Result

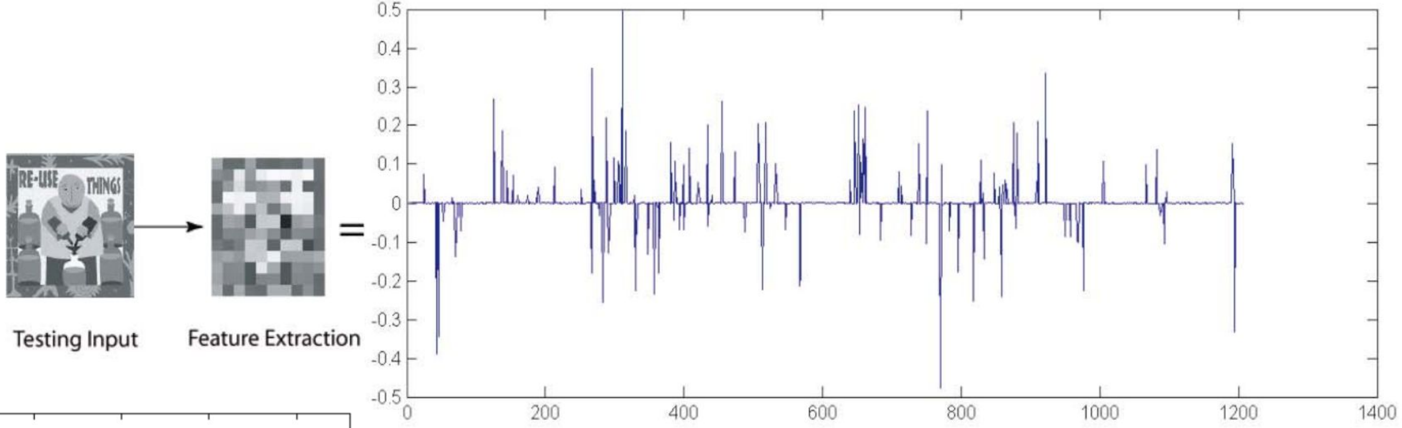


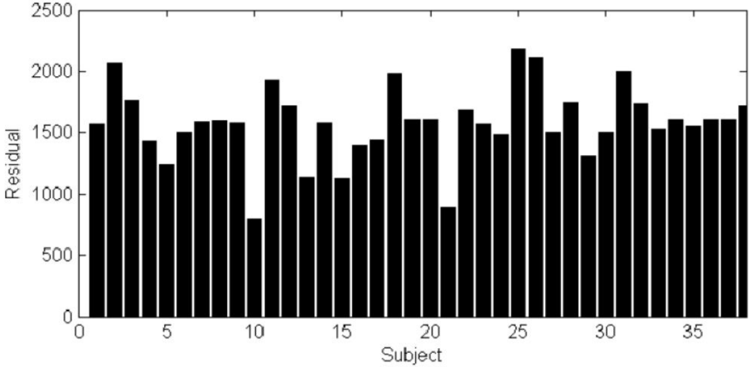
Testing Input
Feature Extraction
=
0.5
0.4
0.3
0.2
0.1
0
200
400
600
800
1000
1200

1500
1000
500
0
Residual
0
5
10
15
20
25
30
35
Subject

# Validation (Reject Outlier)



# Sparsity Concentration Index (SCI)

$$\mathrm{SCI}(\boldsymbol{x}) \doteq \frac{k \cdot \max_i \|\delta_i(\boldsymbol{x})\|_1 / \|\boldsymbol{x}\|_1 - 1}{k - 1} \quad \in [0, 1]$$

$$\mathrm{SCI}(\hat{\boldsymbol{x}}) \geq \tau$$

# Feature Selection

$$R \in \mathbb{R}^{d \times m} \text{ with } d \ll m$$

$$\tilde{\boldsymbol{y}} \doteq R\boldsymbol{y} = RA\boldsymbol{x}_0 \quad \in \mathbb{R}^d$$

*reduced* $\ell^1$-minimization problem:

$$(\ell^1_r): \quad \hat{\boldsymbol{x}}_1 = \arg\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|RA\boldsymbol{x} - \tilde{\boldsymbol{y}}\|_2 \leq \varepsilon.$$

# surprising phenomenon

if $x_0$ has $t << n$ nonzeros

$$d \geq 2t \log(n/d)$$

A random linear projection (R) is sufficient for ℓ1 - minimization (12) to recover the correct sparse solution $x_0$

# Feature Transformation Matrix (R)

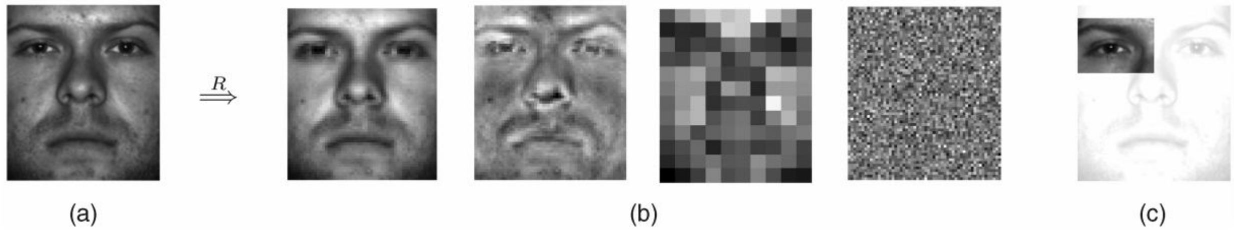



Fig. 6. **Examples of feature extraction.** (a) Original face image. (b) 120D representations in terms of four different features (from left to right): Eigenfaces, Laplacianfaces, downsampled (12 × 10 pixel) image, and random projection. We will demonstrate that all these features contain almost the same information about the identity of the subject and give similarly good recognition performance. (c) The eye is a popular choice of feature for face recognition. In this case, the feature matrix $R$ is simply a binary mask. (a) Original $\boldsymbol{y}$. (b) 120D features $\tilde{\boldsymbol{y}} = R\boldsymbol{y}$. (c) Eye feature $\tilde{\boldsymbol{y}}$.

# Robustness to Occlusion or Corruption

$$\boldsymbol{y} = \boldsymbol{y}_0 + \boldsymbol{e}_0 = A\,\boldsymbol{x}_0 + \boldsymbol{e}_0$$

$$\boldsymbol{y} = [\,A,\;\; I\,]\begin{bmatrix}\boldsymbol{x}_0\\ e_0\end{bmatrix} \doteq B\,\boldsymbol{w}_0$$

$$\hat{\boldsymbol{w}}_1 = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{e}}_1]$$

$$\boldsymbol{y}_r \doteq \boldsymbol{y} - \hat{\boldsymbol{e}}_1$$

# Robustness to Occlusion or Corruption

the $\ell^1$-minimization (22) *cannot* guarantee to correctly recover $\boldsymbol{w}_0 = [\boldsymbol{x}_0,\ e_0]$ if

$$n_i + |\text{support}(e_0)| > d/3.$$

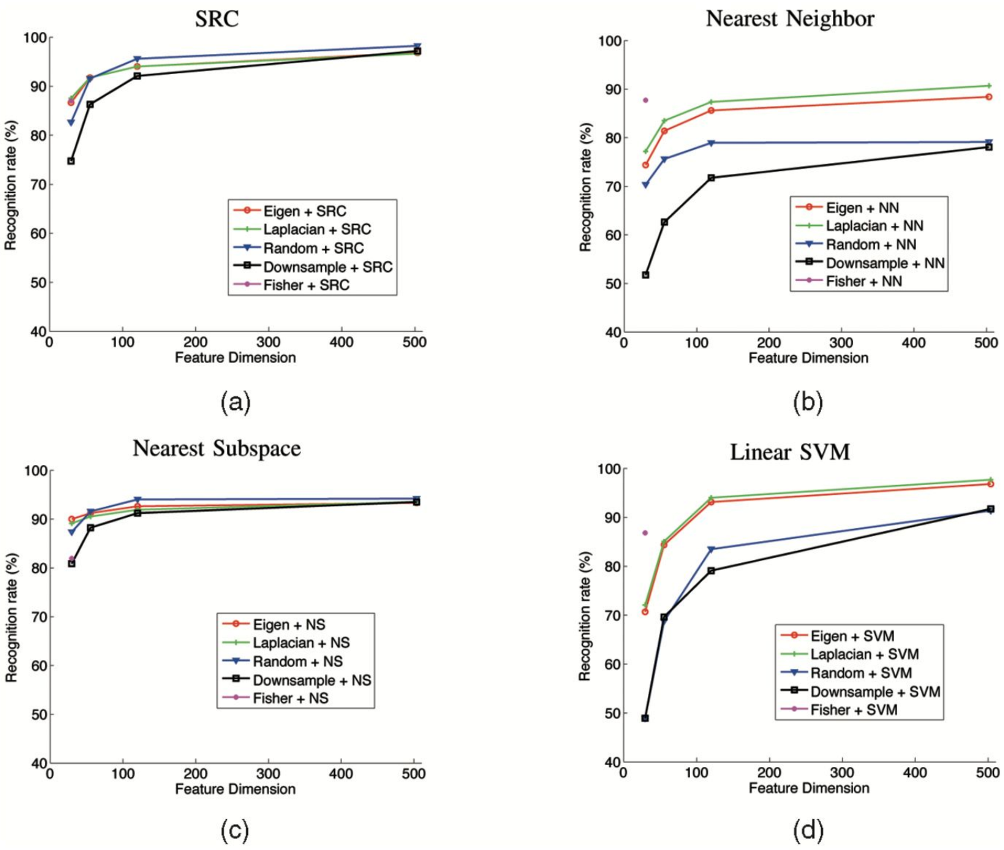
Generally, $d \gg n_i$, so, (8) implies that the largest fraction of occlusion under which we can hope to still achieve perfect reconstruction is 33 percent.

# Experimental Results :

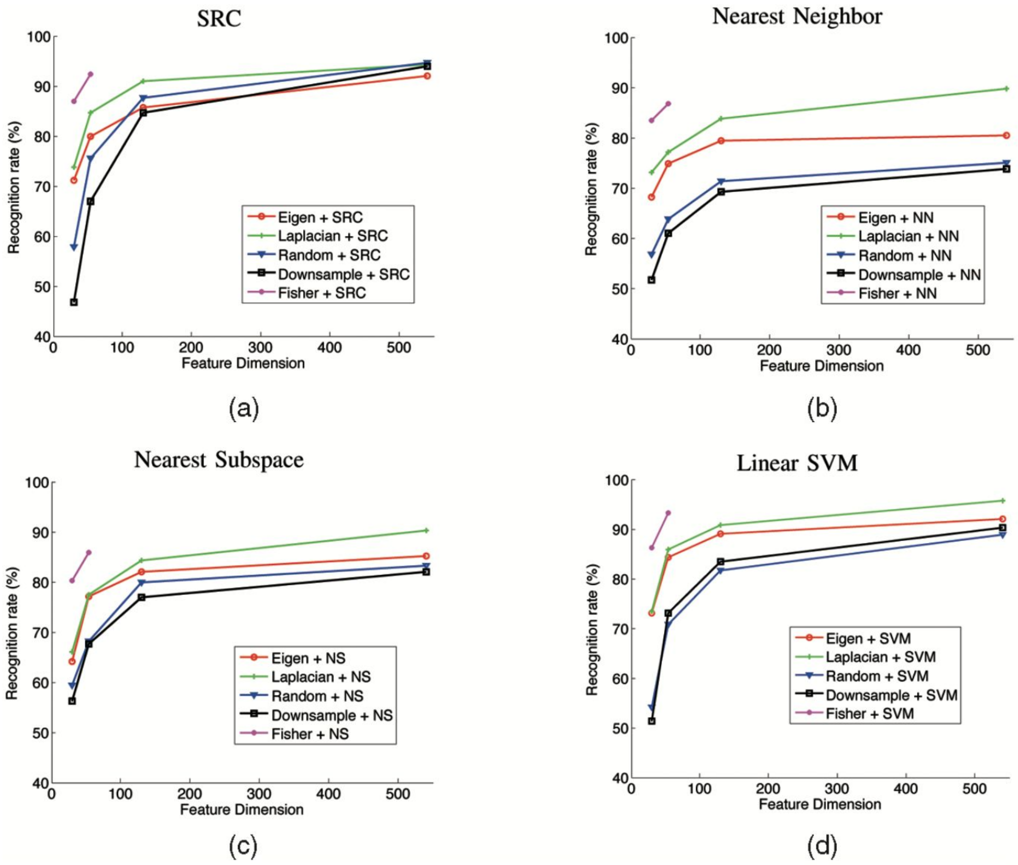
## Databases :

- **Extend Yale B** – 2,414 frontal-face images of 38 individuals – Images are cropped to 192x168 pixels – Captured under various laboratory controlled lighting conditions

- **AR Face DataBase** – Over 4,000 frontal-face images of 126 individuals including illumination change, different expressions and facial disguise – Cropped to 165x120 pixels various feature transformations and classifiers :

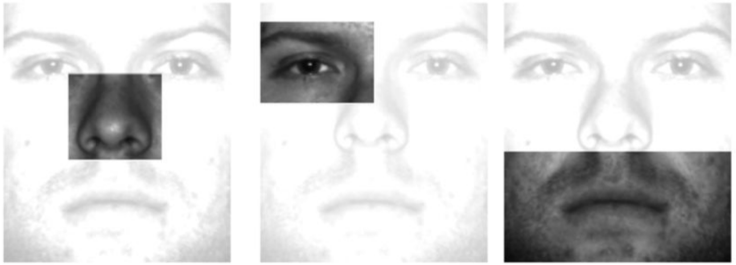


various feature transformations and classifiers :



# Recognition with partial face

(a)

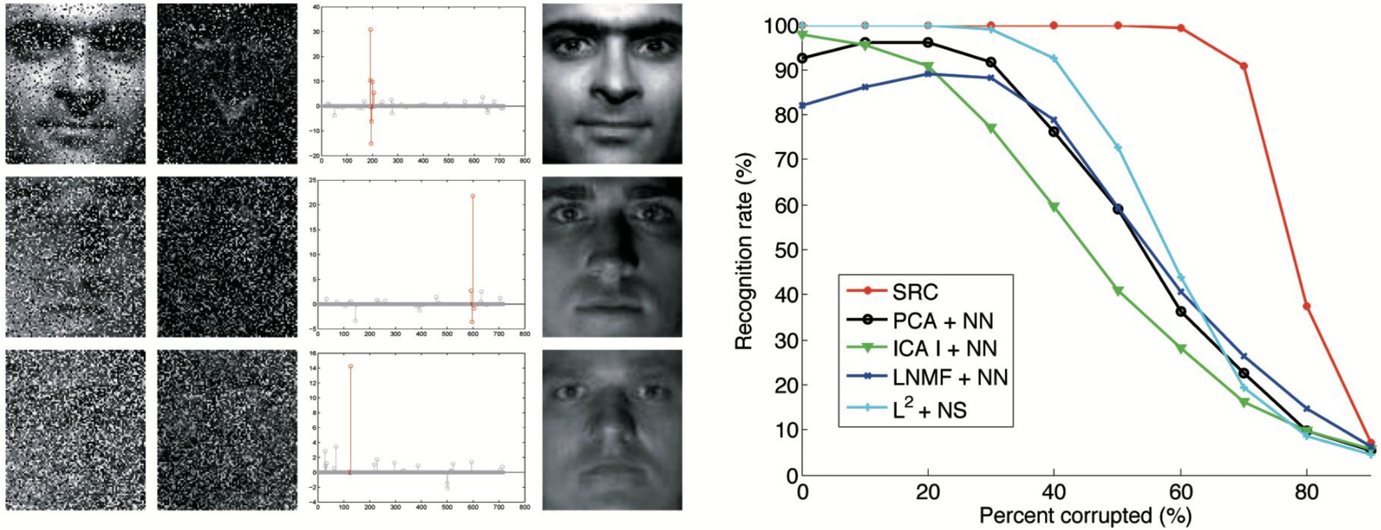


(b)

| Features | Nose | Right Eye | Mouth & Chin |
|---|---|---|---|
| Dimension ($d$) | 4,270 | 5,040 | 12,936 |
| SRC | **87.3%** | **93.7%** | **98.3%** |
| NN | 49.2% | 68.8% | 72.7% |
| NS | 83.7% | 78.6% | 94.4% |
| SVM | 70.8% | 85.8% | 95.3% |

Fig. 10. **Recognition with partial face features.** (a) Example features. (b) Recognition rates of SRC, NN, NS, and SVM on the Extended Yale B database.

# Recognition under random corruption

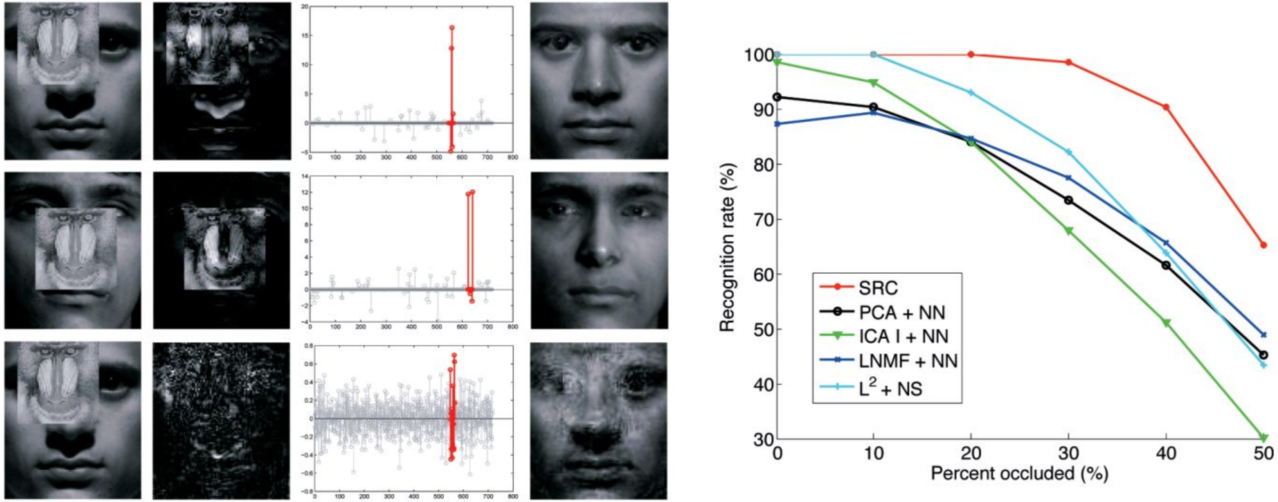



(a) (b) (c) (d) (e)

| Percent corrupted | 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% |
|---|---|---|---|---|---|---|---|---|---|---|
| Recognition rate | 100% | 100% | 100% | 100% | 100% | 100% | 99.3% | 90.7% | 37.5% | 7.1% |

Fig. 11. **Recognition under random corruption.** (a) Test images $y$ from Extended Yale B, with random corruption. Top row: 30 percent of pixels are corrupted. Middle row: 50 percent corrupted. Bottom row: 70 percent corrupted. (b) Estimated errors $\hat{e}_1$. (c) Estimated sparse coefficients $\hat{x}_1$. (d) Reconstructed images $y_r$. SRC correctly identifies all three corrupted face images. (e) The recognition rate across the entire range of corruption for various algorithms. SRC (red curve) significantly outperforms others, performing almost perfectly upto 60 percent random corruption (see table below).

# Recognition under varying Degree of contiguous corruption




(a) (b) (c) (d) (e)

| Percent occluded | 0% | 10% | 20% | 30% | 40% | 50% |
|---|---|---|---|---|---|---|
| Recognition rate | 100% | 100% | 99.8% | 98.5% | 90.3% | 65.3% |

Fig. 12. **Recognition under varying level of contiguous occlusion.** Left, top two rows: (a) 30 percent occluded test face images $y$ from Extended Yale B. (b) Estimated sparse errors, $\hat{e}_1$. (c) Estimated sparse coefficients, $\hat{x}_1$, red (darker) entries correspond to training images of the same person. (d) Reconstructed images, $y_r$. SRC correctly identifies both occluded faces. For comparison, the bottom row shows the same test case, with the result given by least squares (overdetermined $\ell^2$-minimization). (e) The recognition rate across the entire range of corruption for various algorithms. SRC (red curve) significantly outperforms others, performing almost perfectly up to 30 percent contiguous occlusion (see table below).

# outlier rejection

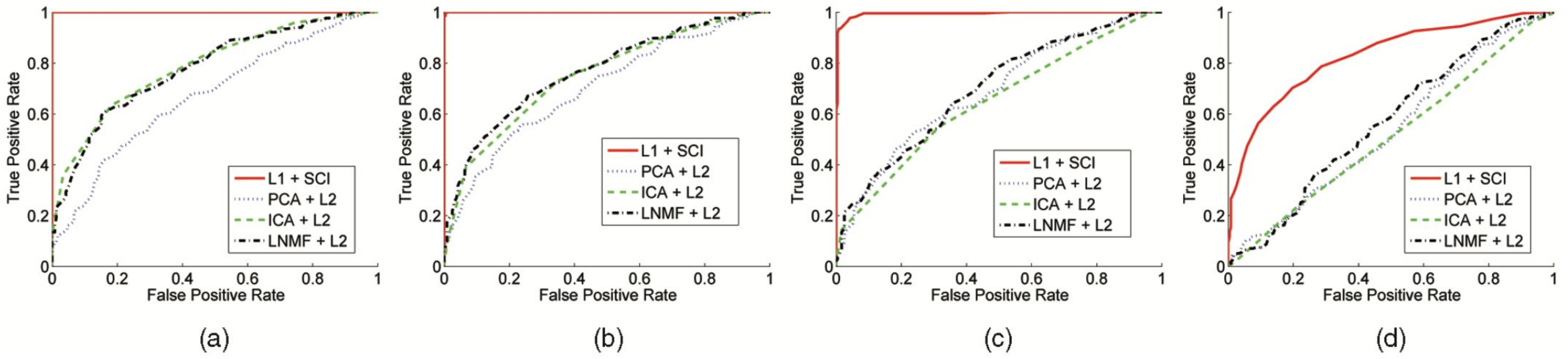



Fig. 14. **ROC curves for outlier rejection.** Vertical axis: true positive rate. Horizontal axis: false positive rate. The solid red curve is generated by SRC with outliers rejected based on (15). The SCI-based validation and SRC classification together perform almost perfectly for upto 30 percent occlusion. (a) No occlusion. (b) Ten percent occlusion. (c) Thirty percent. (d) Fifty percent.

# Summary

- a high performance classification using sparse representation.
- choice of feature is no longer critical
- Robust to Occlusion and corruption
- Accurate in identifying outliers
- Yet having Same Running Time as traditional methods

# Thank You

# Symbols

**k** classes  (individuals)

$\mathbf{n_i}$ = number of training images of person i

**n**  = total number of images = $n_1+n_2+\ldots+n_i+\ldots+n_k$

Each image = w * h pixels = **m** features

$$v_{ij} \in \mathbb{R}^m \qquad A_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$$

# Statement

Given sufficient training samples of the $i$th object class, $A_i = [\boldsymbol{v}_{i,1}, \boldsymbol{v}_{i,2}, \ldots, \boldsymbol{v}_{i,n_i}] \in \mathbb{R}^{m \times n_i}$, any new (test) sample $\boldsymbol{y} \in \mathbb{R}^m$ from the same class will approximately lie in the linear span of the training samples[5] associated with object $i$:

$$\boldsymbol{y} = \alpha_{i,1}\boldsymbol{v}_{i,1} + \alpha_{i,2}\boldsymbol{v}_{i,2} + \cdots + \alpha_{i,n_i}\boldsymbol{v}_{i,n_i}, \tag{1}$$

for some scalars, $\alpha_{i,j} \in \mathbb{R}$, $j = 1, 2, \ldots, n_i$.

- membership *i* of the test sample is initially unknown.

$$A \doteq [A_1, A_2, \ldots, A_k] = [\boldsymbol{v}_{1,1}, \boldsymbol{v}_{1,2}, \ldots, \boldsymbol{v}_{k,n_k}]$$

$$\boldsymbol{y} = A\boldsymbol{x}_0 \quad \in \mathbb{R}^m,$$

where $\boldsymbol{x}_0 = [0, \cdots, 0, \alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n_i}, 0, \ldots, 0]^T \in \mathbb{R}^n$ is a coefficient vector whose entries are zero except those associated with the $i$th class.

# Solve y= Ax

$A_{m*n}$

If m > n :  overdetermined → Unique Solution

If m < n :  underdetermined → Many Solutions

# Solve y= Ax

$$(\ell^0): \qquad \hat{\boldsymbol{x}}_0 = \arg\min \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y},$$

# Solve y= Ax

$$(\ell^0): \qquad \hat{\boldsymbol{x}}_0 = \arg\min \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y},$$

NP hard

# Solve y= Ax

$$(\ell^1): \qquad \hat{\boldsymbol{x}}_1 = \arg\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y}$$

Polynomial Time

# Dealing with small dense noise

$$\boldsymbol{y} = A\boldsymbol{x}_0 + \boldsymbol{z}$$

$$(\ell_s^1): \quad \hat{\boldsymbol{x}}_1 = \arg\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|A\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon$$

# Solve y= Ax

So As long as the number of nonzero entries of $X_0$ is a small fraction of the dimension m, *ℓ1-minimizaion* will recover $X_0$.

# Robust training set design

Subset 1 Subset 2 Subset 3 Subset 4

| Training set | 1 | 2 | 3 | **4** |
|---|---|---|---|---|
| Neighborliness | 1,124 | 1,122 | 1,190 | **1,330** |

(a)

Neutral (N) Happy (H) Angry (A) Screaming (S)

| Training set | **N+H** | N+A | N+S | H+A | H+S | A+S |
|---|---|---|---|---|---|---|
| Neighborliness | **585** | 421 | 545 | 490 | 550 | 510 |

(b)

Fig. 15. **Robust training set design.** (a) Varying illumination. Top left: four subsets of Extended Yale B, containing increasingly extreme lighting conditions. Bottom left: estimated neighborliness of the polytope $\mathrm{conv}(\pm B)$ for each subset. (b) Varying expression. Top right: four facial expressions in the AR database. Bottom right: estimated neighborliness of $\mathrm{conv}(\pm B)$ when taking the training set from different pairs of expressions.