Sparse and Low-Rank Representation for Biometrics – Lecture I: Motivation and Sparse Representation Theory

Allen Y. Yang
Department of EECS, UC Berkeley
yang@eecs.berkeley.edu

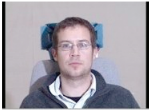
with John Wright and Yi Ma

ICB 2013 Tutorial

Shifting Paradigms in High-Dimensional Pattern Recognition

- Face Recognition


Yale B

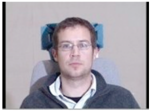

CMU Multi-PIE


Facebook Photo Tagging

Shifting Paradigms in High-Dimensional Pattern Recognition

- Face Recognition

Yale B



CMU Multi-PIE

Facebook Photo Tagging

- Object Recognition

ETHZ Cows vs Cars

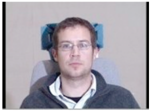
Caltech 101

Amazon Flow

Shifting Paradigms in High-Dimensional Pattern Recognition

- Face Recognition

Yale B



CMU Multi-PIE

Facebook Photo Tagging

- Object Recognition

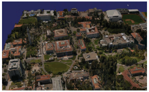
ETHZ Cows vs Cars

Caltech 101

Amazon Flow

- 3D Reconstruction

Oxford Corridor



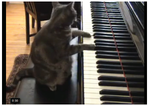
Berkeley Downtown

Google Earth

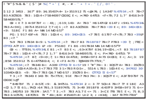
Accurate recognition of HD models presents unique challenges

- Big data vs small training sets: New theory is needed.

average 1M pixels



average 1B voxels



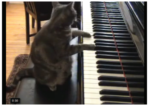
50B webpages

Berkeley
University of California

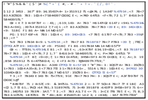
Accurate recognition of HD models presents unique challenges

- Big data vs small training sets: New theory is needed.

average 1M pixels

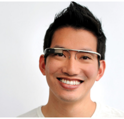


average 1B voxels



50B webpages

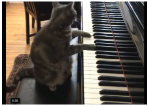
- From desktop to mobile computing: There's an app for everything!



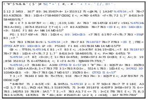
Accurate recognition of HD models presents unique challenges

- Big data vs small training sets: New theory is needed.

average 1M pixels

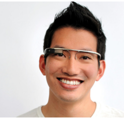


average 1B voxels



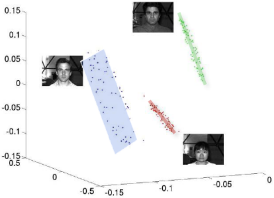
50B webpages

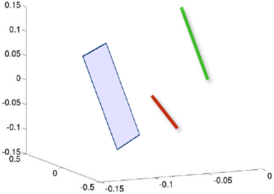
- From desktop to mobile computing: There's an app for everything!



- Real-time performance calls for fast, distributed computing solutions.

Learn low-dimensional structures in high-dimensional data



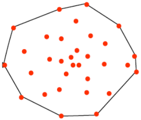


Challenges

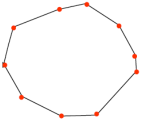
1. **Geometry:** Describe low-dimensional structures in high-dimensional data?
2. **Statistics:** Deal with real data that contain missing observations, corruptions, or even outliers?
3. **Learning:** Handle insufficient data, small sample set problem?
4. **Computation:** Implement a high-dimensional pattern recognition system in real time?

A new regime of geometry, statistics, and computation

- **A sobering message:** Human intuition is severely limited in high-dimensional space.



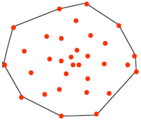
Gaussian samples in 2D



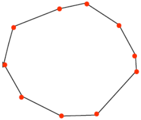
As dimension grows proportionally with the number of samples…

A new regime of geometry, statistics, and computation

- **A sobering message:** Human intuition is severely limited in high-dimensional space.

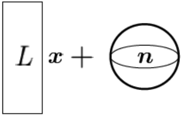


Gaussian samples in 2D

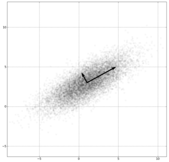


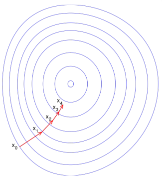
As dimension grows proportionally with the number of samples…

- **Old techniques:** Insufficient to estimate degenerate data model in high-dimension, with limited samples and gross errors.



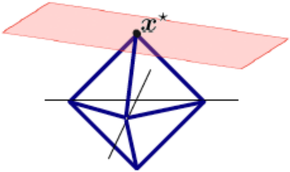
over-determined
+ dense, Gaussian

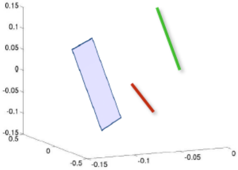


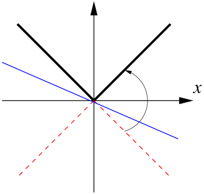


A new regime of geometry, statistics, and computation

- **Theory:** high-dimensional geometry, statistics, combinatorics, coding theory ...
- **Algorithms:** large scale convex optimization, parallel and distributed computing ...
- **Applications:** massive data driven methods, image denoising, super resolution, bioinformatics, image classification, recognition ...



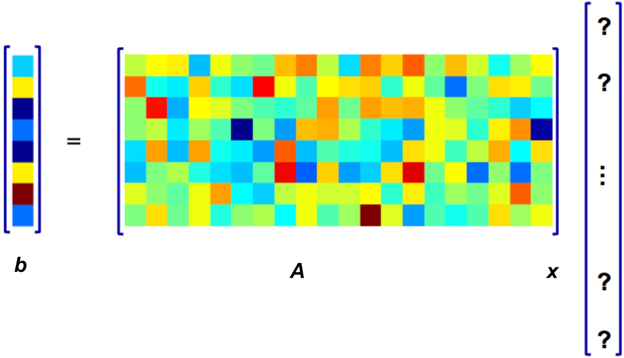




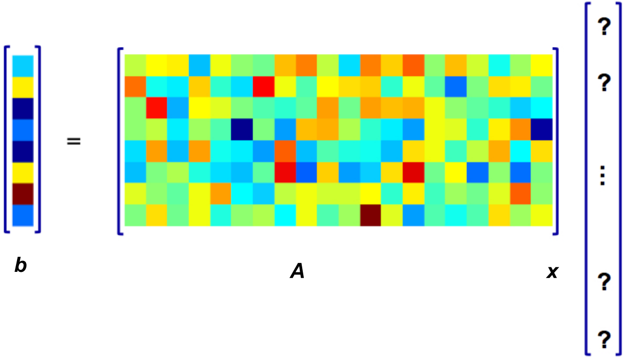
Today's Plan

- **Lecture I:** Motivation and Sparse Representation Theory.
- **Lecture II:** Low-Rank Representation and Applications.
- **Lecture III:** Efficient Sparse Optimization Algorithms.
- **Discussion:** Trends in Wearable Computing

Under-determined Linear Systems

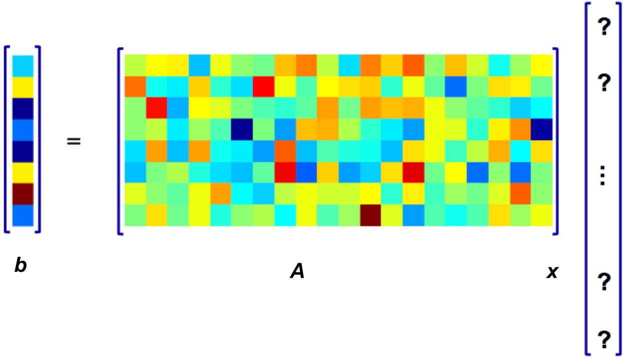


Under-determined Linear Systems



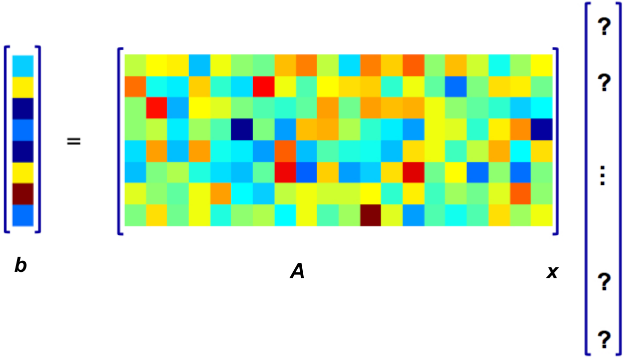
- Does under-determined system always have a solution?

Under-determined Linear Systems



- Does under-determined system always have a solution?
- Is the solution unique?

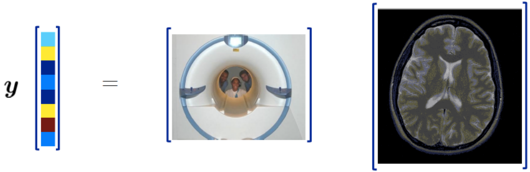
Under-determined Linear Systems



- Does under-determined system always have a solution?
- Is the solution unique?
- Does it have a sparse solution? Is the sparse solution unique?

Example: High-Dimensional Data Acquisition

- MRI imaging



$\boldsymbol{z}$

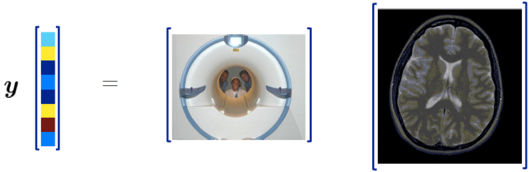
Image to be sensed

$$y_i = \int_{\boldsymbol{u}} \boldsymbol{z}(u) \exp(-2\pi j \boldsymbol{k}(t_i)^* \boldsymbol{u}) d\boldsymbol{u}$$

Observations are Fourier coefficients!

Example: High-Dimensional Data Acquisition

- MRI imaging



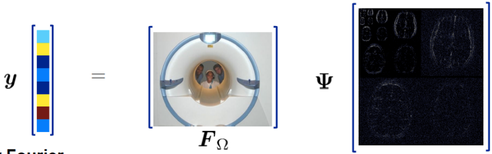
$\boldsymbol{z}$

Image to be sensed

$y_i = \int_{\boldsymbol{u}} \boldsymbol{z}(u) \exp(-2\pi j \boldsymbol{k}(t_i)^* \boldsymbol{u}) d\boldsymbol{u}$

Observations are Fourier coefficients!

- Sparse representation under proper basis



A few Fourier coefficients

$\boldsymbol{F}_\Omega$

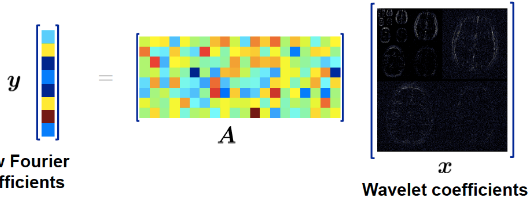
$\boldsymbol{x}$

Wavelet coefficients: $\boldsymbol{z} = \boldsymbol{\Psi x}$

Example: High-Dimensional Data Acquisition

- Modeling MRI as underdetermined system

$$A \doteq F \cdot \Psi$$



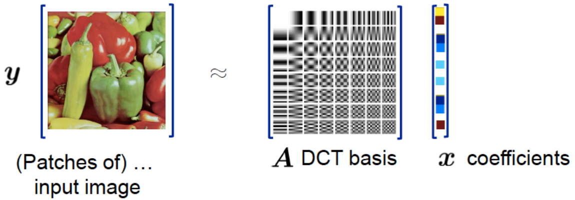
Data Acquisition via Sparse Representation

1. Formulate the data acquisition model: $\mathbf{y} = F\mathbf{z}$
2. Find a sparsifying basis of the data: $\mathbf{z} = \Psi\mathbf{x}$
3. Form overcomplete dictionary and corresponding underdetermined linear system.

$$\mathbf{y} = F\Psi\mathbf{x} \doteq A\mathbf{x}$$

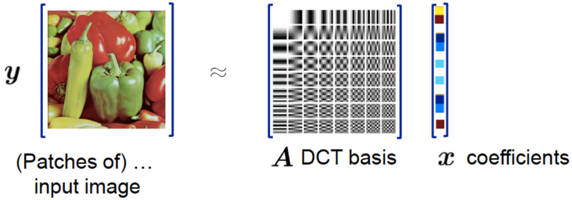
Example: Data Compression

- JPEG image compression

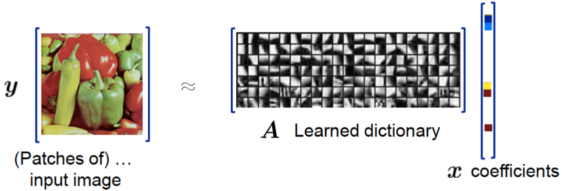


Example: Data Compression

- JPEG image compression

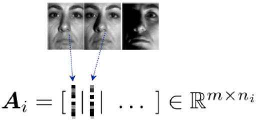


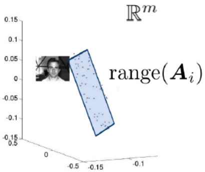
- Sparse representation using learned dictionary



Example: Sparse Representation-based Classification

- Data-driven face subspace model

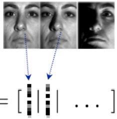




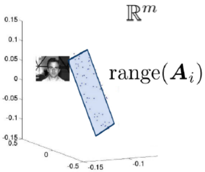
Linear subspace model for images of **same face** under **varying lighting.**

Example: Sparse Representation-based Classification

- Data-driven face subspace model



$$\boldsymbol{A}_i = [\ \cdots\ ] \in \mathbb{R}^{m \times n_i}$$



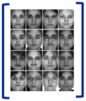
Linear subspace model for images of **same face** under **varying lighting.**

- Sparse representation using learned dictionary

$\boldsymbol{y} \in \mathbb{R}^m$

Test image

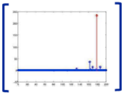
=



$\boldsymbol{A} = [\boldsymbol{A}_1 \mid \boldsymbol{A}_2 \mid \cdots \mid \boldsymbol{A}_k]$

Combined training dictionary

×



$\boldsymbol{x} \in \mathbb{R}^n$

coefficients

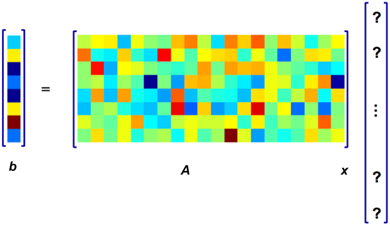
+

$\boldsymbol{e} \in \mathbb{R}^m$

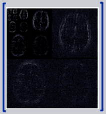
corruption, occlusion

Observation: Solution is not unique

- In all of these examples, #observations ≪ #unknowns ($d \ll n$).

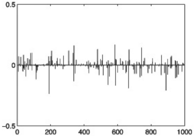


- Can we find a well-defined unique solution that is also the sparsest solution?

Signal acquisition	Image compression	Face Recognition
 $x^\star$ contains **just a few** significant wavelet coefficients.	 $x^\star$ uses **just a few** dictionary elements.	 $x^\star$ uses **just a few** training faces.  $e^\star$ corrects **a few** gross errors.

Sparsity

- A vector $\mathbf{x} \in \mathbb{R}^n$ is **sparse** if only a few entries are nonzero.



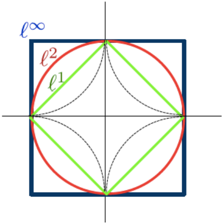
- The number of nonzeros is called ℓ_0-**norm** of $\mathbf{x}$

$$\|\mathbf{x}\|_0 \doteq \#\{i | x_i \neq 0\}.$$

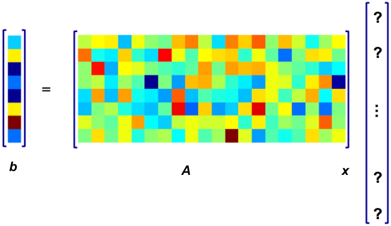
Geometrically

$$\|\boldsymbol{x}\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$$

$$\|\boldsymbol{x}\|_0 = \lim_{p \searrow 0} \|\boldsymbol{x}\|_p^p.$$



The Sparsest Solution



- Optimization program that finds the sparsest $\mathbf{x}$ that satisfies our linear model

$$\ell_0\text{-minimization:}\quad \min \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{b}$$

Theorem: Uniqueness [Gorodnitsky & Rao '97]

- Suppose $\mathbf{b} = A\mathbf{x}_0$, and sparsity $k \doteq \|\mathbf{x}_0\|_0$. If Null($A$) contains no $2k$-sparse vectors, then $\mathbf{x}_0$ is the **unique optimal solution** of ℓ_0-minimization.
- **Spark condition:** For a matrix A, the spark is the smallest number n such that there exists a set of n columns in A that are linearly dependent:

$$\text{spark}(A) = \min_{\mathbf{x} \neq 0} |\mathbf{x}\|_0 \quad \text{s.t.} \quad A\mathbf{x} = 0.$$

Convex Relaxation

- ℓ_0-min is not tractable [Natarjan '95, Amaldi&Kann '97]

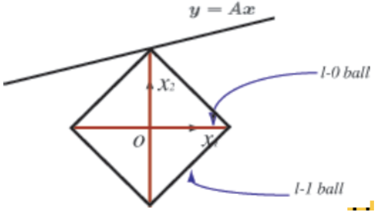
$$\min \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

Convex Relaxation

- ℓ_0-min is not tractable [Natarjan '95, Amaldi&Kann '97]

$$\min \|\mathbf{x}\|_0 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

- Relax nonconvex function using its convex envelope



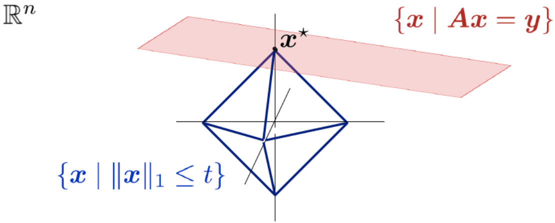
$$\ell_1\text{-min:} \quad \min \|\mathbf{x}\|_1 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

where $\|\mathbf{x}\|_1 = \sum_i |x_i|$.

- Lecture III: Efficient ℓ_1-min solvers.

Does the relaxation lose any solutions?

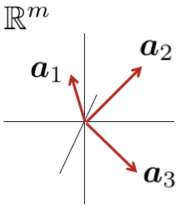
- Geometric intuition



- When ℓ_0-min and ℓ_1-min are equivalent?

Mutual Coherence

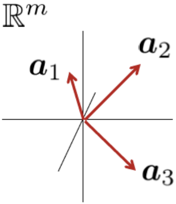
- We see that $\mathbf{y} = A\mathbf{x} = \sum_{i \in \mathrm{supp}(\mathbf{x})} \mathbf{a}_i x_i$.



- **Intuition:** Recoverying $\mathbf{x}$ is easier if the elements $\mathbf{a}_i$ are not similar ...

Mutual Coherence

- We see that $\mathbf{y} = A\mathbf{x} = \sum_{i \in \mathrm{supp}(\mathbf{x})} \mathbf{a}_i x_i$.



- **Intuition:** Recoverying $\mathbf{x}$ is easier if the elements $\mathbf{a}_i$ are not similar ...
- **Mutual coherence:** Assume column vectors are normalized $\|\mathbf{a}_i\| = 1$:

$$\mu(A) \doteq \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|$$

Mutual Coherence Condition

Theorem: Mutual Coherence [Gribonval&Nielsen '03, Donoho&Elad '03]

Suppose $\mathbf{y} = A\mathbf{x}_0$ with

$$\|\mathbf{x}_0\|_0 < \frac{1}{2}(1 + 1/\mu(A)).$$

Then $\mathbf{x}_0$ is the unique optimal solution to ℓ_1-min.

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

Mutual Coherence Condition

Theorem: Mutual Coherence [Gribonval&Nielsen '03, Donoho&Elad '03]

Suppose $\mathbf{y} = A\mathbf{x}_0$ with

$$\|\mathbf{x}_0\|_0 < \frac{1}{2}(1 + 1/\mu(A)).$$

Then $\mathbf{x}_0$ is the unique optimal solution to ℓ_1-min.

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}$$

1. **LHS:** Source signal $\mathbf{x}_0$ is sufficiently sparse.
2. **RHS:** Matrix A is sufficiently incoherent, so it preserves the structure of $\mathbf{x}_0$ after projection.

Limitations of Coherence

- **Lower-bound:** For an $m \times n$ matrix A, $\mu(A) \geq \sqrt{\frac{n-m}{m(n-1)}}$.
  Therefore, the MC theorem suggests equivalence condition when

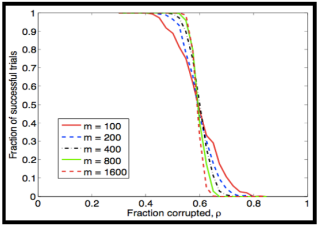
$$\|\mathbf{x}_0\|_0 < \frac{1}{2}(1 + 1/\mu(A)) = O(\sqrt{m}).$$

Limitations of Coherence

- **Lower-bound:** For an $m \times n$ matrix A, $\mu(A) \geq \sqrt{\frac{n-m}{m(n-1)}}$.
  Therefore, the MC theorem suggests equivalence condition when

$$\|\mathbf{x}_0\|_0 < \frac{1}{2}(1 + 1/\mu(A)) = O(\sqrt{m}).$$

- In practice, the phase transition point is often much higher:



Plot: Fraction of correct recovery vs. fraction of nonzeros $\|x_0\|_0/m$

$$\|\mathbf{x}_0\|_0 = \alpha \cdot m.$$

Strengthen the Bound – the RIP Condition

- MC condition only considers the pairwise spread $\langle \mathbf{a}_i, \mathbf{a}_j \rangle$.
- **Restricted isometry property** generalizes to subsets of size k:
  If A_I is well-spread for all index sets I of size k, then

$$\Rightarrow \text{For all } k\text{-sparse } \mathbf{x},\ \|A\mathbf{x}\|_2 \approx \|\mathbf{x}\|_2.$$

Strengthen the Bound – the RIP Condition

- MC condition only considers the pairwise spread $\langle \mathbf{a}_i, \mathbf{a}_j \rangle$.
- **Restricted isometry property** generalizes to subsets of size k:
  If A_I is well-spread for all index sets I of size k, then

$$\Rightarrow \text{For all } k\text{-sparse } \mathbf{x},\ \|A\mathbf{x}\|_2 \approx \|\mathbf{x}\|_2.$$

Definition: Restricted Isometry Property

A satisfies the RIP of order k with constant $\delta < 1$ if for all k-sparse $\mathbf{x}$:

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|A\mathbf{x}\|_2^2 \leq (1+\delta)\|\mathbf{x}\|_2^2,$$

- If δ is small, it means A is well-spread for all k-sparse signals.

RIP for Sparse Recovery

Theorem [Candès&Tao '05, Candès '08]

Suppose $\mathbf{y} = A\mathbf{x}_0$ with $\|\mathbf{x}_0\|_0 = k$ and A satisfies RIP condition with

$$\delta_{2k} \leq \sqrt{2} - 1,$$

then $\mathbf{x}_0$ is the unique optimal solution to ℓ_1-min:

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad A\mathbf{x} = \mathbf{y}.$$

- Again, if $\mathbf{x}_0$ is sufficiently sparse (structured) and A is "nice", then a convex program, ℓ_1-min, can exactly recover $\mathbf{x}_0$.

Face Recognition via Sparse Representation

1. Face-subspace model [Belhumeur et al. '97, Basri & Jacobs '03]
   Assume **b** belongs to Class i from K classes.

$$\begin{aligned} \mathbf{b} &= \alpha_{i,1}\mathbf{a}_{i,1} + \alpha_{i,2}\mathbf{a}_{i,2} + \cdots + \alpha_{i,n_1}\mathbf{a}_{i,n_i}, \\ &= A_i\alpha_i. \end{aligned}$$

   Dimension of **b** can be thousands, but the subspace model is a few tens.

Face Recognition via Sparse Representation

1. Face-subspace model [Belhumeur et al. '97, Basri & Jacobs '03]
   Assume **b** belongs to Class i from K classes.

$$\begin{aligned}\mathbf{b} &= \alpha_{i,1}\mathbf{a}_{i,1} + \alpha_{i,2}\mathbf{a}_{i,2} + \cdots + \alpha_{i,n_1}\mathbf{a}_{i,n_i},\\ &= A_i\alpha_i.\end{aligned}$$

   Dimension of **b** can be thousands, but the subspace model is a few tens.
2. Nevertheless, Class i is the **unknown label** we need to solve:

Sparse representation $\quad \mathbf{b} = [A_1, A_2, \cdots, A_K]\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_K\end{bmatrix} = A\mathbf{x}.$

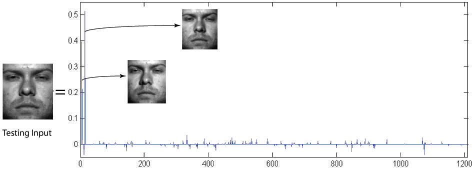
Face Recognition via Sparse Representation

1. Face-subspace model [Belhumeur et al. '97, Basri & Jacobs '03]
   Assume **b** belongs to Class i from K classes.

$$\begin{aligned}\mathbf{b} &= \alpha_{i,1}\mathbf{a}_{i,1} + \alpha_{i,2}\mathbf{a}_{i,2} + \cdots + \alpha_{i,n_1}\mathbf{a}_{i,n_i},\\ &= A_i\alpha_i.\end{aligned}$$

   Dimension of **b** can be thousands, but the subspace model is a few tens.
2. Nevertheless, Class i is the **unknown label** we need to solve:

   Sparse representation $\quad \mathbf{b} = [A_1, A_2, \cdots, A_K]\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_K\end{bmatrix} = A\mathbf{x}.$
3. $\mathbf{x}^* = [0 \cdots 0\ \alpha_i^T\ 0 \cdots 0]^T \in \mathbb{R}^n.$



Sparse representation $\mathbf{x}^*$ encodes membership through its nonzero coefficients!

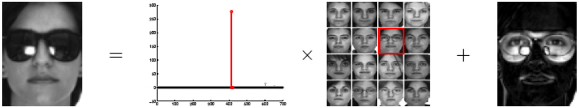
Reference:
Wright, AY, Sastry, Ma, *Robust face recognition via sparse representation*. **IEEE PAMI**, 2009.

Image Occlusion, Corruption, and Disguise

Image Occlusion, Corruption, and Disguise

1. Sparse representation + sparse error

$$\mathbf{b} = A\mathbf{x} + \mathbf{e}$$



2. Cross-and-bouquet model [Wright et al. '09, '11]

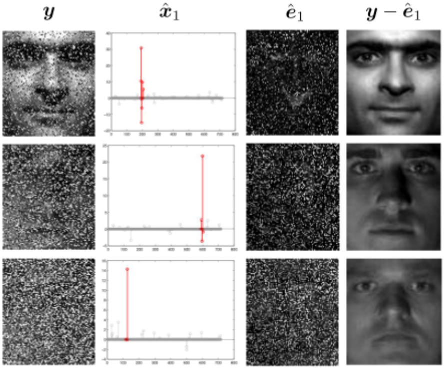
$$\min_{\mathbf{x},\mathbf{e}} \|\mathbf{x}\|_1 + \|\mathbf{e}\|_1 \quad \text{subj. to} \quad \mathbf{b} = A\mathbf{x} + \mathbf{e}$$

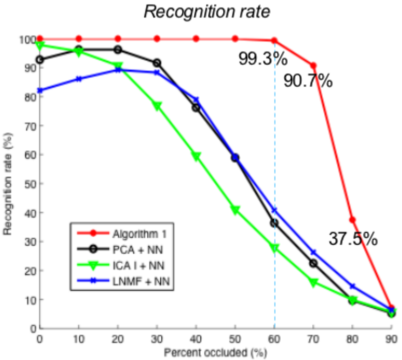
When size of A grows proportionally with the sparsity in $\mathbf{x}$, asymptotically CAB can correct 100% noise in $\mathbf{e}$.

Reference:

Wright and Ma, *Dense Error Correction via ℓ_1 Minimization*, IEEE Trans. IT, 2011.

Performance on the YaleB database

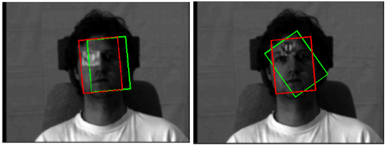




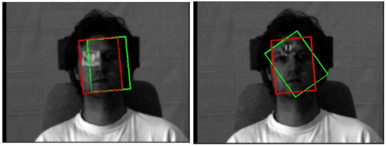
Reference:
Wright, AY, Sastry, Ma, *Robust face recognition via sparse representation*. **IEEE PAMI**, 2009.

Top 25 IEEExplore Downloads since 2010. 1600+ citations on Google.

Face Alignment Problem: Misalignment violates linear subspace model



Face Alignment Problem: Misalignment violates linear subspace model

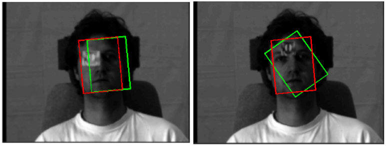


1. **Find an image transformation** τ (2-D function that transforms image coordinates)

$$\min \|\mathbf{e}\|_1 \quad \text{subj. to} \quad \mathbf{b} \circ \tau_i = A_i \mathbf{x} + \mathbf{e}$$

per each class A_i, while minimize the alignment error $\mathbf{e}$.

Face Alignment Problem: Misalignment violates linear subspace model



1. **Find an image transformation** τ (2-D function that transforms image coordinates)

$$\min \|\mathbf{e}\|_1 \quad \text{subj. to} \quad \mathbf{b} \circ \tau_i = A_i \mathbf{x} + \mathbf{e}$$

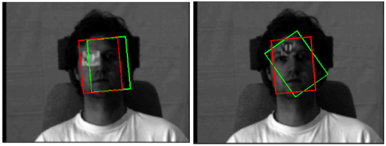
per each class A_i, while minimize the alignment error $\mathbf{e}$.

2. **Iterative linear approximation** [Lucas & Kanade '81, Hager & Belhumeur '98]:

$$\mathbf{b} \circ \tau_i + \nabla_\tau (\mathbf{b} \circ \tau_i) \cdot \Delta\tau_i \approx A_i \mathbf{x} + \mathbf{e}.$$

Convert to a linear sparse optimization constraint.

Face Alignment Problem: Misalignment violates linear subspace model



1. **Find an image transformation** τ (2-D function that transforms image coordinates)

$$\min \|\mathbf{e}\|_1 \quad \text{subj. to} \quad \mathbf{b} \circ \tau_i = A_i \mathbf{x} + \mathbf{e}$$

per each class A_i, while minimize the alignment error $\mathbf{e}$.

2. **Iterative linear approximation** [Lucas & Kanade '81, Hager & Belhumeur '98]:

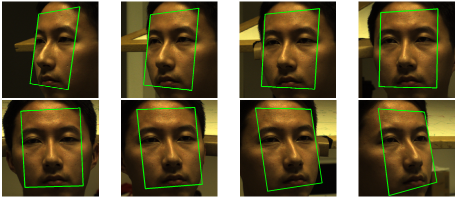
$$\mathbf{b} \circ \tau_i + \nabla_\tau (\mathbf{b} \circ \tau_i) \cdot \Delta\tau_i \approx A_i \mathbf{x} + \mathbf{e}.$$

Convert to a linear sparse optimization constraint.

3. Compensated training images are fed back to the sparse representation model:

$$\mathbf{b} = [\tau_1^{-1}(A_1), \cdots, \tau_K^{-1}(A_K)]\mathbf{x} + \mathbf{e}.$$

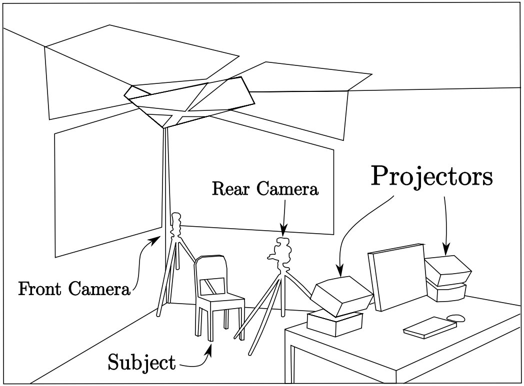
Region of Convergence for 2-D Alignment



Using affine transformations, ℓ_1-min approach can compensate 3-D rotation up to 30 degree and 2-D translation up to 10 pixels.

References:
Ganesh, Ma, Wagner, Wright, AY, Zhou, *Face recognition by sparse representation*, **Cambridge University Press**, 2011.
Ma, AY, Wright, Wagner, *Recognition via High-Dimensional Data Classification*, **US Patent**, 2013.

A Photo Booth for Face Recognition



Most existing face recognition solutions require large numbers of training images:

$$A = [A_1, \ldots, A_C]$$

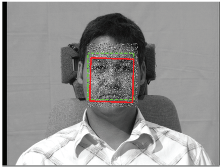
Question: What if subjects of interest only have limited training images, e.g., one image per class?

Single-Sample Face Recognition via Sparse Illumination Transfer

- Alignment Stage: Illumination Compensation + Misalignment + Pixel Corruption

Single Training Image

Test image align with pixel corruption

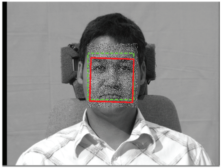


Single-Sample Face Recognition via Sparse Illumination Transfer

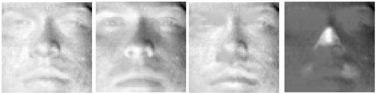
- Alignment Stage: Illumination Compensation + Misalignment + Pixel Corruption

Single Training Image

Test image align with pixel corruption



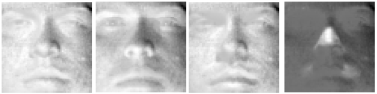
- **Sparse Illumination Transfer (SIT)** via Transfer Learning Approach



Given new face illumination examples from additional irrelevant subjects $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$:

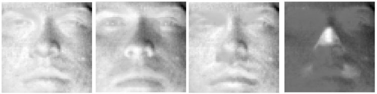
$$C = [\cdots, \mathbf{c}_i - \mathbf{c}_j, \cdots]_{i \neq j, (i,j)} \text{ belong to the same subject}$$

C is called a SIT dictionary independent of the training dictionary A, constructed offline.



1. **Alignment Stage:**

$$\begin{array}{rcl}\hat{\tau}_i & = & \arg\min_{\tau_i, x_i, \mathbf{y}_i, \mathbf{e}} \|\mathbf{y}_i\|_1 + \|\mathbf{e}\|_1, \\ \text{subj. to} & & \mathbf{b} \circ \tau_i - \mathbf{a}_i x_i = C\mathbf{y}_i + \mathbf{e}\end{array}$$



1. **Alignment Stage:**

$$\begin{aligned} \hat{\tau}_i &= \arg\min_{\tau_i, x_i, \mathbf{y}_i, \mathbf{e}} \|\mathbf{y}_i\|_1 + \|\mathbf{e}\|_1, \\ \text{subj. to} &\quad \mathbf{b} \circ \tau_i - \mathbf{a}_i x_i = C\mathbf{y}_i + \mathbf{e} \end{aligned}$$

2. **Recognition Stage:** Illumination and Pose Transfer

$$\tilde{\mathbf{a}}_i \doteq (\mathbf{a}_i x_i + C\mathbf{y}_i) \circ \tau_i^{-1}.$$

Figure : **Left:** Testing **b**. **Mid Left:** Training $\mathbf{a}_i$. **Mid Right:** $C\mathbf{y}_i$. **Right:** Warped $\tilde{\mathbf{a}}_i$.

Results and Comparison: Single-Sample Face Recognition

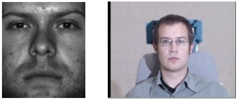


Figure : SIT on YaleB dataset, train on Multi-PIE Session I, test on Session I & II (166 subjects).

1. SIT improves existing face recognition solutions (alignment provided manually)

Method	Session 1 (%)	Session 2 (%)
SRC '09	88.0	53.6
ESRC '12	89.6	56.6
SRC + SIT	91.6	59.0
ESRC + SIT	**93.6**	**59.3**

2. Full pipeline (alignment + recognition) with added pixel corruption

Corruption	10%	20%	30%	40%
DSRC '09	32.9%	31.7%	28.9%	24.1%
MRR '12	24.9%	14.5%	11.7%	9.2%
SIT	**74.3%**	**70.3%**	**67.1%**	**55.8%**

References:
Zhuang, AY, Zhang, Sastry, Ma, *Single-Sample Face Recognition via Sparse Illumination Transfer*, **CVPR**, 2013.
US patent application filed by UC Berkeley, 2013.