Compressive sensing (CS) and sparse representations (SR) for computer vision and image processing

Rama Chellappan
University of Maryland
Compressive sensing: Is it another wave?

- **Computer vision in the grip of many waves**
  - MRFs (1980 – 1995), reemergence in graphics and vision
  - Regularization (From early eighties ..)
  - Invariants (Mid eighties to mid nineties)
  - Particle filters (From mid nineties; has reached the status of FFT in DSP!)
  - Graphical models (From early nineties)
  - Context (From early nineties…)
  - Computational sensing (Started in early nineties, died a bit later, reemerging as computational photography)
  - Manifolds
Why CVPR/PRIP folks should be interested in CS and SR

- Sensor and processing tightly coupled
  - Natural outgrowth of integrated sensing and processing, A2I efforts from DARPA/DSO since the early nineties
  - Already having impact on medical imaging, SAR imaging, face/iris recognition, new camera designs, etc

- Data deluge
  - From not enough data to too much data
  - Optimized data acquisition strategies and sparse representation will allow the full digestion of data
  - Data to action

- Hard sciences and tons of data can and do help each other!
Outline

• Compressive sensing (CS)
  – Basics
  – Restricted isometry property (RIP), Johnson Lindenstrauss (JL) lemma.
  – Recovery algorithms
• Examples
  – Compressive sensing-driven camera
  – Dictionary-based face recognition
  – Reconstruction from gradient fields using $l_1$ optimization
  – Compressive video processing
  – CS and graphics
  – CS for SAR imaging
• Concluding remarks
  – Breakthroughs will come in integrated compressive sensing and processing
  – $l_1$ optimization and sparse representations suggest novel ways of sensing, reconstruction, processing and understanding of signals and images.
Collaborators

• Vishal Patel
• Jai Pillai
• Dikpal Reddy
• Aswin C. Sankaranarayanan (Rice)
• Pavan Turaga
• Rich Baraniuk (Rice) and
• Volkan Cevher (EPFL)
Motivation

- **Shannon/Nyquist sampling theorem**: Sampling rate must be at least twice the message signal bandwidth in order to achieve exact recovery.
- **Compressed sensing** is a new method to capture and represent compressible signals at the rate well below Nyquist’s rate.
  - Employs nonadaptive linear projections (random measurement matrix)
  - Preserves the signal structure (length of the sparse vectors is conserved)
  - Reconstructs the signal from the projections using optimization process (L1 norm)
An early CS paper for signal recovery

- S. Levy and P. K Fullagar, Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution, Geophysics, 46, 1235-1243(1981).
Classical Approach: Transform coding

- **PICTURE**
  - N samples (ALL measurements are taken)
- **ORTHO-GONALIZATION**
  - Full set of projections is found
- **SORTING**
  - K largest coefficients selected
  - N-K coef. dumped
- **CODING**
  - Only K coefficients are coded

**K << N**

Exhaustive Search

Compressed sensing

- **PICTURE**
  - Capture only significant components
  - Only M (K ≈ M) samples (measurements) are taken
- **CODING**

**Signal Reconstruction**

1. Underdetermined system: M < N
2. Reconstructed signal must have N components
3. $L^1$ norm is used to find sparse representation

(a) Measurements must be carefully designed
(b) Original signal (picture) must be sparse

Courtesy: Rich Baraniuk
Sparsity: Motivation

Periodic signal

Sparse in Fourier

Background subtracted images

Images

Wavelet tree

Chirp Signals $\rightarrow$ Sparse in Short Time Fourier Transform
Compressive sampling

- When data is sparse/compressible, can directly acquire a **condensed representation** with no/little information loss through linear **dimensionality reduction**

\[ y = \Phi x \]

\[ M \times 1 \text{ measurements} \quad = \quad \Phi \quad \text{nonzero entries} \]

\[ K < M \ll N \]

From Dr. V. Cevher
How does it work?

- Projection $\Phi$ not full rank…
  $M < N$

  … and so loses information in general

- Infinitely many $x$’s map to the same $y$

- But we are only interested in sparse vectors

- Design $\Phi$ so that each of its $MxK$ submatrices are full rank
  – preserve information in $K$-sparse signals
  – Restricted Isometry Property (RIP) of order $2K$

From Dr. V. Cevher
Restricted Isometry Property (RIP)

A matrix $\Theta$ is said to satisfy the RIP of order $K$ with constants $\delta_K \in (0, 1)$ if

$$(1 - \delta_K) \| v \|_2^2 \leq \| \Theta v \|_2^2 \leq (1 + \delta_K) \| v \|_2^2$$

for any $v$ such that $\| v \|_0 \leq K$. [Candes and Tao, 2005].

- RIP is a sufficient condition to find the sparsest solution.
- When RIP holds, $\Theta$ approximately preserves the Euclidean length of $K$-sparse signals.
- All subsets of $K$-columns taken from $\Theta$ are nearly orthogonal.
- RIP has been established for some matrices such as random Gaussian, Hadamard, and Fourier, however, in practice there is no computationally feasible way to check this property for a given matrix, as it is combinatorial in nature.
Direct construction of measurement matrix (Φ)

• Verifying RIP requires \( \binom{N}{K} \) all possible combinations of \( K \) non-zero entries of length \( N \)

• However RIP could be achieved by simply selecting \( Φ \) randomly
  – \( φ_{ij} \) are independent and identically distributed random variables from Gaussian probability density with zero mean and 1/N variance.
  – So measurements \( y = Φx \)

are randomly weighted linear combinations of elements of signal.
For compressible signals

\[
\mathbf{x} \text{ target signal, representation of an image in time or space domain}
\]
\[
x \text{ is in N dimensional domain}
\]
\[
\mathbf{\Psi} \text{ orthonormal basis}
\]
\[
\mathbf{\Psi} = \begin{bmatrix}
\psi_{11} & \cdots & \psi_{1N} \\
\vdots & \ddots & \vdots \\
\psi_{N1} & \cdots & \psi_{NN}
\end{bmatrix}
\]
\[
\mathbf{\psi}_i \text{ orthonormal vector}
\]
\[
\mathbf{\psi}_i = \begin{bmatrix}
\psi_{i1} \\
\vdots \\
\psi_{iN}
\end{bmatrix}
\]

Classical Approach: Find the signal projections on a given basis: 
\[
x = \sum_{i=1}^{N} s_i \mathbf{\psi}_i
\]

\[
S_i \text{ representation of the image in } \mathbf{\Psi} \text{ domain}
\]
\[
S_i = \mathbf{\psi}_i^T \mathbf{x}
\]

\[x \text{ is } K\text{-sparse if only } K \text{ basis vectors } \mathbf{\psi}_i \text{ have } s_i \neq 0 \text{ (theoretical)} \]

\[x \text{ is compressible if it has just a few large } s_i \text{ coefficients and many small (practical)} \]

Compressible signals are well represented by K-sparse representations.

Compressed sensing: Measure only significant components \((M \ll N)\)

\[
\mathbf{y} \text{ measurements}
\]
\[
y \text{ is in M dimensional domain}
\]
\[
\mathbf{\Phi} \text{ measurement matrix}
\]
\[
\mathbf{\Phi} = \begin{bmatrix}
\phi_{11} & \cdots & \phi_{1N} \\
\vdots & \ddots & \vdots \\
\phi_{M1} & \cdots & \phi_{MN}
\end{bmatrix}
\]
\[
y = \mathbf{\Phi} \mathbf{x} = \mathbf{\Phi} \mathbf{\Psi} \mathbf{s} \Rightarrow \text{Goal } y = \mathbf{\Theta} s
\]

Courtesy: Rich Baraniuk
(a) Compressive sensing measurement process with a random Gaussian measurement matrix and discrete cosine transform (DCT) matrix. The vector of coefficients $s$ is sparse with $K = 4$.

$\Phi$ (phi, measurement matrix) $\Psi$ (psi, orthonormal basis) 
$\Theta$ (theta, Compressed Sensing reconstruction matrix)

(b) Measurement process with $\Theta = \Phi \Psi$ There are four columns that correspond to nonzero $s$ coefficients; the measurement vector $y$ is a linear combination of these columns.

From IEEE Signal processing magazine July 2007, R. G. Baraniuk, Compressive sensing
Insight from the 80’s [Kashin, Gluskin, 1984]

• Draw $\Phi$ at random
  – iid Gaussian
  – iid Bernoulli $\pm 1$

• Then $\Phi$ has the RIP with high probability as long as
$$M = O(K \log(N/K)) \ll N$$
  – $M \times 2K$ submatrices are full rank

From Dr. V. Cevher
Compressive sensing solution

- Measurement matrix must be stable and produce reconstruction of original signal, \( x \), (length \( N \)) from \( M \) measurements
- 1) \( \Psi \) othonormal basis must be selected
- 2) Stable \( \Phi \) (measurement matrix) must be designed (CS measures PROJECTIONS of signal onto a basis) and
- 3) the signal \( x \) has to be reconstructed from the underdetermined system using optimization (norm \( L^1 \)) linear combinations, since the projections (linear combinations are measured) we are extracting the sparse components using optimization algorithm (\( L^1 \)) with restrictions

Courtesy: Rich Baraniuk
Designing a signal reconstruction algorithm

• **Signal reconstruction algorithm aims to find signal’s sparse coefficient vector.**
  
  - **L2 norm** (energy) minimized. Pseudo inverse is closed form to find the solution but it DOES NOT find sparse solution.
  - **L0 norm** counts the number of non-zero elements of s. This optimization can recover K sparse signal exactly with high probability using only $M=K+1$ measurements but solving it is unstable and NP-complete requiring exhaustive enumeration of all $(N \times K)$ possible locations of the non-zero entries in s.
  - **L1 norm** (adding absolute values of all elements) can exactly **recover K sparse signals** and closely approximate compressible signals with high probability using only $M \geq cK \log(N/K)$ iid Gaussian measurements.
  
    • Convex optimization reducing to **linear programming known as Basis pursuit** with the computation complexity about $O(N^3)$
CS signal recovery

- **Recovery:** given \( y = \Phi x \) find \( x \)
  (ill-posed inverse problem)
  (sparse)

- \( \ell_2 \) fast, wrong
- \( \ell_0 \) correct, slow
- \( \ell_1 \) correct, efficient
  mild oversampling

[From Dr. Cevher]

- \[ \hat{x} = \arg \min_{y=\Phi x} \| x \|_2 \]
- \[ \hat{x} = \arg \min_{y=\Phi x} \| x \|_0 \]
- \[ \hat{x} = \arg \min_{y=\Phi x} \| x \|_1 \]

number of measurements required

\[ M = O(K \log(N/K)) \ll N \]

[Candès, Romberg, Tao; Donoho]

linear program
\( \ell_0 / \ell_1 \) equivalence


- Let \( p = M/N \) be a measure of undersampling factor, and \( q = K/M \) be a measure of sparsity. A plot of the pairing of the variables \( p \) and \( q \) describes a two-dimensional phase space.

- \( p \) and \( q \) ranged through 40 equi-spaced points in the interval \([0; 1]\) and \( N = 800\): at each point on the grid, we recorded the mean number of coordinates at which original and reconstruction differed by more than 0.001; averaged over 20 independent realizations (see the paper by Donoho.)
CS recovery algorithms

- Convex optimization
  - noise-free signals
    - Linear programming (Basis pursuit) [Chen et al, 2001, Donoho 2006]
  - noisy signals
    - Basis Pursuit De-Noising (BPDN)
    - Second-Order Cone Programming (SOCP)
- Iterative greedy algorithms
  - Matching Pursuit (MP) [Mallat and Zhang, 1993]
BPDN

- Basis Pursuit Formulation (A Linear program)
  \[ \min \| x \|_1 \quad \text{subject to} \quad y = \Phi x \]

- Noisy measurements
  \[ y = \Phi x + n \]

- Basis Pursuit De-Noising
  \[ \min \| x \|_1 \quad \text{s.t.} \quad \| y - \Phi x \|_2 \leq \varepsilon \]

- Convex, quadratic program

From Dr. Cevher
Recent developments in computer vision/graphics

- Compressive sensing for coded aperture imaging [Veeraraghavan and Reddy, PAMI 2010]
- Sparse representations for face recognition [Wright et al, PAMI 2009, CVPR 2009]
- Robust and secure iris recognition (BTAS 2009, ICASSP2010,PAMI 2011]
- Shape from gradients (CVPR 2009)
- Image reconstruction from gradients
- Compressive Sensing of Reflectance Fields
  - Peers et al, ACM Transactions on Graphics, 2009)
- Compressive sensing for background subtraction [Cevher et al, ECCV 2008]
- Compressive SAR imaging, ICIP 2009, IEEE JI. SASP, 04/10
“Single-pixel” CS camera

From IEEE Signal processing magazine July 2007, R. G. Baraniuk, Compressive sensing

- Flip mirror array $M$ times to acquire $M$ measurements
- Sparsity-based (linear programming) recovery
First image acquisition

Takhar et al, 2006
Duarte et al, 2008

Target
65536 pixels

11000 measurements
(16%)

1300 measurements
(2%)
Gradient domain processing: vision and graphics

Our method: Fit surface in 1-norm sense

Agrawal, Raskar
Few applications

Shape from Shading, Photometric Stereo

High Dynamic Range (HDR) Compression

Edge suppression under significant illumination variations
Gradient fields and integrability

Image or surface: \( S(x, y) \)  
Gradients: \( \nabla S = \left\{ \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y} \right\} \)

Divergence: \( \text{Div}(\nabla S) = \nabla \cdot \nabla S \)

Curl: \( \text{Curl}(\nabla S) = \nabla \times \nabla S \)

Integrability: Conservative vector field

For a scalar field \( S(x, y) \)

\[
\nabla \times \nabla S = 0 \\
\text{Curl}(\nabla S) = S_{yx} - S_{xy} = 0
\]

Slides on this topic are from Reddy, Agrawal and Chellappa  
CVPR 2009
Non-integrable gradient fields

• Estimation of gradients
  • E.g. Shape from Shading, Photometric Stereo
  • Noise and outliers in estimation

Input Images → Surface Normals/Gradients

Not-integrable

• Manipulation of integrable gradients
  • Synthesize new gradient field

Image → Gradient Manipulations → New Gradients

Not-integrable
Integrability problem

Desired surface $s$. Corrupted gradient field $g = [g_x^T \ g_y^T]^T$

$$g = Ds + e$$

Curl of the gradient field given by

$$d = Cg = Ce$$

$d = Ce$  Under-determined system of equations

• Similar to error correction.
• Estimate gradient error given the curl values

Poisson equation

$$(P2) \quad \hat{e} = \arg \min \|e\|_2 \text{ s.t. } d = Ce$$
Integrability problem

Sparse recovery:

Sparse errors (arbitrarily large) can be recovered using

(P0) \( \hat{e} = \arg \min \|e\|_0 \) s.t. \( d = Ce \)

(P0) NP-hard. So solve a convex formulation

(P1) \( \hat{e} = \arg \min \|e\|_1 \) s.t. \( d = Ce \)

- (P1) not only corrects outliers, corrects noise as well as (P2)

- Has local error confinement property

- Motivated by recent results on equivalence of (P0) and (P1) in compressed sensing literature
CS and reconstruction from gradient fields

- Using CS we can analyze when recovery is guaranteed.
- For non contiguous errors C has RIP -1.
- How many outliers can $l_1$ minimization fully correct?
- How should they be distributed?
- If large number of outliers then what outliers does $l_1$ find and correct?

$l_1$-minimization

$$\hat{e} = \arg\min_{e} \| e \|_1 \quad \text{s.t.} \quad d = Ce$$

- Corrects outliers. Performs well in noise too.
- Good error confinement property locally.

Ground truth

Least squares

1-norm fit
$l_1$ minimization

Under noise and outliers
$l_1$ minimization

Comparison

Local error confinement
Image recovery: problem setup

An image $x$ is said to be $K$-sparse in gradient if

$$\| TV(x) \|_0 = K,$$

where

$$TV(x)_{n,m} = \sqrt{p(n, m)^2 + q(n, m)^2}$$

with

$$p(n, m) = x(n, m+1) - x(n, m)$$

and

$$q(n, m) = x(n+1, m) - x(n, m)$$

Can we recover?

Figure 1: Shepp-Logan Phantom and its gradient
Basis pursuit (BP) solution

• Find the sparsest solution $x$ that satisfies Fourier constraint:

$$x_{rec} = \arg \min_{x' \in \mathbb{C}^N} \|x\|_0 \quad \text{s. t.} \quad Y = W_\Omega x'$$

• Relax combinatorial problem into a convex optimization one:

$$x_{rec} = \arg \min_{x' \in \mathbb{C}^N} \|x\|_1 \quad \text{s. t.} \quad Y = W_\Omega x'$$

• Theoretical guarantees: RIP, incoherence, ...

Total variation (TV) minimization

• In the spirit of BP, find the sparsest image in the total variation sense that satisfies our Fourier constraint
• In other words, solve (Candes, Romberg, Tao, 2006):
  \[ Y = W_\Omega x \]
• How can we use a greedy pursuit algorithm such as MP and OMP to recover sparse gradient images? (Maleh, Gilbert, Stauss, 2007)

\[
x_{\text{rec}} = \arg \min_{x' \in \mathbb{C}^N} \| TV(x') \|_1 \quad \text{s. t.} \quad W_\Omega x' = W_\Omega x
\]
An observation

- Instead of reconstructing by the TV minimization, recover the image by separately reconstructing the gradients and then solving for the image.
- Requires a far fewer number of measurements than the TV minimization!
Least squares integration method (A. C. Gilbert et al.)

Generate a reconstruction \( \hat{X} \) of \( X \) given approximations \( \hat{\partial_x X} \) and \( \hat{\partial_y X} \) of \( \partial_x X \) and \( \partial_y X \), respectively. Here, \( \partial_x X \) and \( \partial_y X \) are the gradients in the horizontal and vertical directions, respectively.

\[
\hat{X} = \underset{Y}{\text{argmin}} \left\| \partial_x Y - \hat{\partial_x X} \right\|_2^2 + \left\| \partial_y Y - \hat{\partial_y X} \right\|_2^2 + \beta \left\| \partial_x Y \right\|_2^2 + \beta \left\| \partial_y Y \right\|_2^2 + \lambda \left\| \mathcal{F}_\Omega Y - \mathcal{F}_\Omega X \right\|_2^2
\]

Similar to enforcing integrability using Fourier Basis while reconstructing depth form gradients. 
Frankot and Chellappa, IEEE Trans. PAMI, 1988)

In the Fourier domain

\[
\hat{X} = \underset{\hat{Y}}{\text{argmin}} \left\| (1 - e^{-2\pi i \omega_1 / N}) \hat{Y} - \hat{\partial_x X} \right\|_2^2 + \left\| (1 - e^{-2\pi i \omega_2 / N}) \hat{Y} - \hat{\partial_y X} \right\|_2^2 + \beta \left\| (1 - e^{-2\pi i \omega_1 / N}) \hat{Y} \right\|_2^2 + \beta \left\| (1 - e^{-2\pi i \omega_2 / N}) \hat{Y} \right\|_2^2 + \lambda \left\| \hat{Y} - \hat{X} \right\|_2^2.
\]

(3.9)

\[
(3.10) \quad \hat{X}_{\omega_1, \omega_2} = \frac{(1 - e^{2\pi i \omega_1 / N}) \hat{\partial_x X}_{\omega_1, \omega_2} + (1 - e^{2\pi i \omega_2 / N}) \hat{\partial_y X}_{\omega_1, \omega_2} + \lambda \hat{X}_{\omega_1, \omega_2} 1_\Omega}{(1 + \beta) \left( \left| 1 - e^{-2\pi i \omega_1 / N} \right|^2 + \left| 1 - e^{-2\pi i \omega_2 / N} \right|^2 \right) + \lambda 1_\Omega}.
\]

**Proposition III.1.** The least squares problem (3.9) can be solved element-wise by the following formula:
Sparse gradient-based image reconstruction

\[ p(n, m) = x(n, m + 1) - x(n, m) \quad \quad \quad \quad q(n, m) = x(n + 1, m) - x(n, m) \]

\[ Y = W_\Omega x \]

\[ P = (e^{i\frac{2\pi k_2}{l}} - 1)Y \]

\[ P = W_\Omega p \]

\[ Q = (e^{i\frac{2\pi k_1}{l}} - 1)Y \]

\[ Q = W_\Omega q \]

\[ p_{rec} = \arg \min_{p'} \| p' \|_1 \quad \text{s. t.} \quad \| W_\Omega p' - P \|_2 \leq \varepsilon \]

\[ q_{rec} = \arg \min_{q'} \| q' \|_1 \quad \text{s. t.} \quad \| W_\Omega q' - Q \|_2 \leq \varepsilon \]

Integrate gradients

Poisson Solver

\[ x_{rec} \]
The gradient field \((p, q)\) of an image should be integrable (i.e. the integral along any closed curve should be equal to zero).

This is often not the case when inherent noise during the estimation process contaminates \((p, q)\).

Have to enforce integrability.

Numerical examples

Figure 2: (a) $512 \times 512$ Shepp-Logan Phantom example. Reconstructed by (b) zero-filling, (c) GradientOMP, (d) TV method, and (e) our proposed method.  

94.6% random undersampling

Figure 3: (a) Original MRI image. (b) Radial $k$-space trajectories. Reconstructed by (c) zero-filling, (d) GradientOMP, and (e) our method. 

68.5% random undersampling
Experiments

Figure 4: (a) Noisy $128 \times 128$ phantom image (SNR 20dB). (b) Sampling pattern. Reconstructed image by (c) GradientOMP and (d) by our method.

65% random undersampling
Face recognition via sparse representations

- Represent the test face as a sparse linear combination of the training faces.
- Estimate the class of the test image from the sparse coefficients.
- Alignment is handled by finding the best transformation between the training and test images.
- Illumination variation is taken care of by including training images under varying illuminations - Wagner et al. CVPR 2009
Dictionary-based face recognition

• Can work with only one training image per class.
• Sparse representation is found by learning a dictionary for each class using KSVD.
• Uses a relighting approach to deal with the illumination problem.
Dictionary-based face recognition

Classification result among several classes
**Objective:** Find the best dictionary to represent the samples \( B = [x_1, \cdots, x_m] \) as sparse compositions, by solving the following optimization problem:

\[
\arg\min_{D, \Gamma} \|B - D\Gamma\|_F^2 \quad \text{subject to} \quad \forall i, \|\gamma_i\|_0 \leq T_0.
\]

**Input:** Initial dictionary \( D^{(0)} \in \mathbb{R}^{N \times P} \), with normalized columns, signal matrix \( B = [x_1, \cdots, x_m] \) and sparsity level \( T_0 \).

**Output:** Trained dictionary \( D \) and sparse representation matrix \( \Gamma \).

**Procedure:**
Set \( J = 1 \). Repeat until convergence:

- **Sparse coding stage:** Use any pursuit algorithm to compute the sparse representation vectors \( \gamma_i \) for each signal \( [x_1, \cdots, x_m] \).
- **Dictionary update stage:** For each column \( k = 1, \cdots, P \) in \( D^{(J-1)} \) update by:
  - Define the group of examples that use this atom, \( \omega_k = \{i|1 \leq i \leq P, \gamma_i^k(\cdot) \neq 0\} \).
  - Compute the overall representation error matrix, \( E_k \), by
    \[
    E_k = B - \sum_{j \neq k} c_j \gamma_j^k.
    \]
  - Restrict \( E_k \) by choosing only the columns corresponding to \( \omega_k \) and obtain \( E_k^R \).
  - Apply SVD decomposition \( E_k^R = U\Delta V^T \). Select the updated dictionary column \( \hat{c}_k \), to be the first column of \( U \). Update the coefficient vector \( \hat{\gamma}_k^R \) to be the first column of \( V \) multiplied by \( \Delta(1,1) \).
- Set \( J = J + 1 \).

---

Given a set of training examples, \( B = [x_1, \cdots, x_m] \), find the dictionary that leads to the best representation for each member in this set, under strict sparsity constraints.

Outlier rejection

**Figure:** An invalid test image and the resulting residuals.

**Figure:** The score values corresponding to a valid sample (left) and an invalid sample (right).

Define score \( s_{yi} \) of the test image \( y \) to the \( j^{th} \) class as

\[
s_{yi} = \frac{1}{\| r'(y) \|_2^2}.
\]

Sort score values in decreasing order such that

\[
s_{y1}' \geq s_{y2}' \geq \cdots \geq s_{yC}'.
\]

Use \( \lambda_y = \frac{s_{y1}'}{s_{y2}'} \) to reject outliers.
Dealing with the illumination problem

• **Robust albedo estimation** (Biswas et al. PAMI 2009)
  
  – Estimate albedo
  
  – Relight images with different light source direction
  
  – Use relighted images for training

\[
I = \rho \max(n^T s, 0)
\]

Lambert’s Cosine Law

Fig. 4. Examples of the original images (first column) and the corresponding relighted images with different light source directions from the PIE data set.
Experimental results (Yale B)

- **DFR** – 99.17%
- **SRC** – 98.1%
- **CDPCA** – 98.83%

Results on other data sets not shown due to lack of time.

Fig. 5. Performance comparison on the Extended Yale B database with various features, feature dimensions and methods. (a) Our method (DFR) (b) SRC [5] (c) NN [5] (d) NS [5] (e) SVM [5] (f) CDPCA.

- DFR – 99.17%
- SRC – 98.1%
- CDPCA – 98.83%
Experimental results

![Graph showing recognition rate vs. number of dictionary atoms on Extended Yale B and PIE data sets.]

Fig. 8. Recognition rate vs. number of dictionary atoms on Extended Yale B and PIE data sets.

**TABLE IV**

**PERFORMANCE COMPARISON OF DIFFERENT METHODS WITH RESPECT TO THE NUMBER OF TRAINING SAMPLES PER SUBJECT.**

<table>
<thead>
<tr>
<th>No. of training images</th>
<th>DFR</th>
<th>SRC</th>
<th>NS</th>
<th>CDPCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75.89%</td>
<td>42.37%</td>
<td>36.13%</td>
<td>5.52%</td>
</tr>
<tr>
<td>2</td>
<td>84.71%</td>
<td>37.20%</td>
<td>46.36%</td>
<td>26.22%</td>
</tr>
<tr>
<td>3</td>
<td>85.18%</td>
<td>37.45%</td>
<td>52.40%</td>
<td>30.25%</td>
</tr>
</tbody>
</table>
Iris recognition

• Recognize a person from the texture features on his/hers iris image [Pillai et al, BTAS 2009]
• Existing algorithms [Daugman 93] give high recognition rates.
• Iris images acquired from a partially cooperating user often suffer from
  – Specular reflections, segmentation error, occlusion & blur.
• We developed a sparse representation-based algorithm for iris image selection and recognition - To appear in PAMI
Results

Image Selection – ROC curves

 SCI Variations with distortions

Pillai et al, BTAS 2009

<table>
<thead>
<tr>
<th>Image Quality</th>
<th>NN</th>
<th>Masek’s Implementation</th>
<th>Our Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>98.33</td>
<td>97.5</td>
<td>99.17</td>
</tr>
<tr>
<td>Blurred</td>
<td>95.42</td>
<td>96.01</td>
<td>96.28</td>
</tr>
<tr>
<td>Occluded</td>
<td>85.03</td>
<td>89.54</td>
<td>90.30</td>
</tr>
<tr>
<td>Seg. Error</td>
<td>78.57</td>
<td>82.09</td>
<td>91.36</td>
</tr>
</tbody>
</table>
Sectored random projections For cancelable iris biometrics

• Need For Cancelability
  – Iris patterns are unique to each person.
  – Iris patterns cannot be re-issued if stolen.
  – Different patterns required for different applications.

• Cancelable Biometrics – Apply a revocable and non-invertible transformation on the original one [Ratha et al, 2001, Teoh et al, 2006, Hao et al, 2006]

• Requirements
  – Performance should be retained.
  – Should be non-invertible and revocable.
  – Different codes for different applications.

• Pillai et al, ICASSP 2010, PAMI 2011.
Random projections for cancelability

- Random Projections (RP) can be utilized due to Johnson Lindenstrauss (JL) lemma [1984].

**Lemma 1. (Johnson-Lindenstrauss)** Let $\epsilon \in (0, 1)$ be given. For every set $S$ of $\#(S)$ points in $\mathbb{R}^N$, if $n$ is a positive integer such that $n > n_0 = O \left( \frac{\ln(\#(S))}{\epsilon^2} \right)$, there exists a Lipschitz mapping $f : \mathbb{R}^N \rightarrow \mathbb{R}^n$ such that

\[
(1 - \epsilon)\|u - v\|^2 \leq \|f(u) - f(v)\|^2 \leq (1 + \epsilon)\|u - v\|^2 \quad (1)
\]

for all $u, v \in S$.

- One such mapping is projection using a Random Matrix.
Random projections

- N dimensional Gabor features of the iris image $g$ is projected randomly onto a subspace of dimension $n$ as follows
  $$y = \Phi g \in \mathbb{R}^n$$

- Recognition is performed on the vector $y$.
- Direct Random Projections on iris images give poor results due to:
  - Occlusion due to eyelids act as outliers corrupting the whole data after the linear transform.
  - Combines both the good and bad regions of the iris image.
Results – Recognition performance

Performance Comparison

Effect Of Dimension Reduction

Observations
• SRP performs close to the original system without random projections.
• Only minor drop in performance due to dimension reduction upto 30% of the original dimension.
• Dimension reduction further improves non invertibility due to the non zero dimension of the null space.
• Pillai et al, ICASSP 2010.
Other works not discussed

• CS for video interpretation
• Compressive wireless arrays for bearing estimation of sparse sources in angle domain (Cevher et al, ICASSP 2008).
• Compressed sensing for multi-view tracking and 3-D voxel reconstruction (Reddy et al, ICIP 2008).
• Compressive SAR imaging (Patel, et al, 2008)
Remarks

• Compressive sensing cannot and will not solve all computer vision problems.
• May provide better solutions in many cases.
• Renewed interest in $l^{-1}$-based methods!
• Compressive sensor design will lead the way.
  – Coded aperture imaging
  – Algorithms have to be integrated with sensing to reap the full benefits.
  – Tempting to replace all $l^{-2}$ regularization methods with $l^{-1}$.
• Ask if you can do without CS before attempting to do with CS!
• Check the Rice University website
Compressive sensing for video processing

• Background subtraction and tracking

• Registration

• Acquisition of dynamic scenes
LDS model

- Video modeled as first order Gauss Markov process

\[
y(t) = C x(t) + w(t) \\
x(t+1) = A x(t) + v(t)
\]

- \( x(t) \) \( d \times 1 \) hidden state vector
- \( A \): \( d \times d \) transition matrix
- \( C \): \( N \times d \) observation matrix

\[X(0) \rightarrow A,C,Q,R \rightarrow y(0), y(1), y(2)\ldots\]
Why LDS models?

• Useful for modeling video sequences
  – Soatto
• Have been used for unsupervised video clustering
  – Turaga, et al CVIU 2008
• Have been used for human activity recognition
• Orthonomral C matrices enable geodesic distances for classification
Measurement strategy

• Take two sets of measurements at each time-instant – one common, another that varies with time

\[ z_t = \begin{pmatrix} \tilde{z}_t \\ \tilde{z}_t \end{pmatrix} = \begin{bmatrix} \tilde{\Phi} \\ \tilde{\Phi}_t \end{bmatrix} y_t = \Phi_t y_t \]

Common measurements: Same measurement matrix
Model driven sensing of dynamic scenes

Scene

CS camera

Common measurements

Innovation measurements

Estimate state sequence using SVD

Model based CoSaMP for recovery of observation matrix
Solve for the state-sequence and the observation matrix

• The state-sequence can be estimated up to a linear transform by SVD of common measurements.

\[
[\tilde{\mathbf{z}}]_{1:T} = \begin{bmatrix} \tilde{z}_1 & \tilde{z}_2 & \cdots & \tilde{z}_T \end{bmatrix} = \tilde{\Phi} C \begin{bmatrix} x_1 & x_2 & \cdots & x_T \end{bmatrix} = \tilde{\Phi} C [x]_{1:T}
\]

• Assume columns of \( C \) are sparse in a basis \( \psi \), then solve for the following problem by structured CoSAMP.

\[
\min \sum_{k=1}^{d} \| \psi^T c_k \|_1, \text{ subject to } \| z_t - \Phi_t C \hat{x}_t \|_2 \leq \epsilon, \forall t
\]
Reconstruction results

Candle sequence
- 1024 frames captured at 1000 fps
- Resolution 64 x 64
- Measurement rate: 1.2 %
- State dimension: 15

LED sequence
- 500 frames
- Resolution 64 x 64
- Measurement rate: 2.1 %
- State dimension: 7
Video classification results

• UCSD dataset consists of 254 videos of length 50 frames capturing traffic of three types – light, moderate, heavy. We performed a classification experiment of the videos into these three categories.
• 4 different train-test scenarios provided with the dataset.
• Classification is performed using the subspace-angles based metric with a nearest-neighbor classifier on the LDS parameters.
• The experiment was performed using the parameters estimated directly without reconstructing the frames. For comparison, we also perform the same experiments with fitting the LDS model on the original frames (oracle LDS).
Classification results

Classification results on the traffic databases for two different values of state space dimension d. Results are over a database of 254 videos, each of length 50 frames at a resolution of 64 x 64 pixels under a measurement rate of 4%.

<table>
<thead>
<tr>
<th></th>
<th>Expt. 1</th>
<th>Expt. 2</th>
<th>Expt. 3</th>
<th>Expt. 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>d=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle LDS</td>
<td>85.71</td>
<td>85.93</td>
<td>87.5</td>
<td>92.06</td>
</tr>
<tr>
<td>CS-LDS</td>
<td>84.12</td>
<td>87.5</td>
<td>89.06</td>
<td>85.71</td>
</tr>
<tr>
<td>d=10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle LDS</td>
<td>77.77</td>
<td>82.81</td>
<td>92.18</td>
<td>80.95</td>
</tr>
<tr>
<td>CS-LDS</td>
<td>85.71</td>
<td>73.43</td>
<td>78.10</td>
<td>76.10</td>
</tr>
</tbody>
</table>
Rotating Mill Tool captured by Dragonfly2

Mill tool rotating at 50Hz

Reconstructed Video at 2000fps

Normal Video: 25fps

Coded Strobing Video: 25fps

Mill tool rotating at 150Hz

Reconstructed Video at 2000fps

Mill tool rotating at 100Hz

Reconstructed Video at 2000fps

Mill tool rotating at 200Hz

Reconstructed Video at 2000fps