# Generalized Principal Component Analysis for Image Representation & Segmentation

**Yi Ma**

Control & Decision, Coordinated Science Laboratory
Image Formation & Processing Group, Beckman
Department of Electrical & Computer Engineering

University of Illinois at Urbana-Champaign

INTRODUCTION

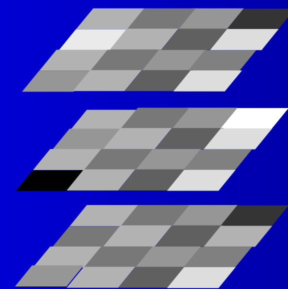
GPCA FOR LOSSY IMAGE REPRESENTATION

IMAGE SEGMENTATION VIA LOSSY COMPRESSION

OTHER APPLICATIONS

CONCLUSIONS AND FUTURE DIRECTIONS

# Introduction – Image Representation via Linear Transformations



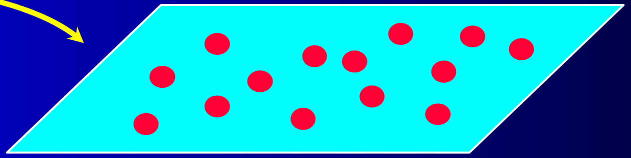
pixel-based representation

three matrixes of RGB-values

better representations?

linear transformation

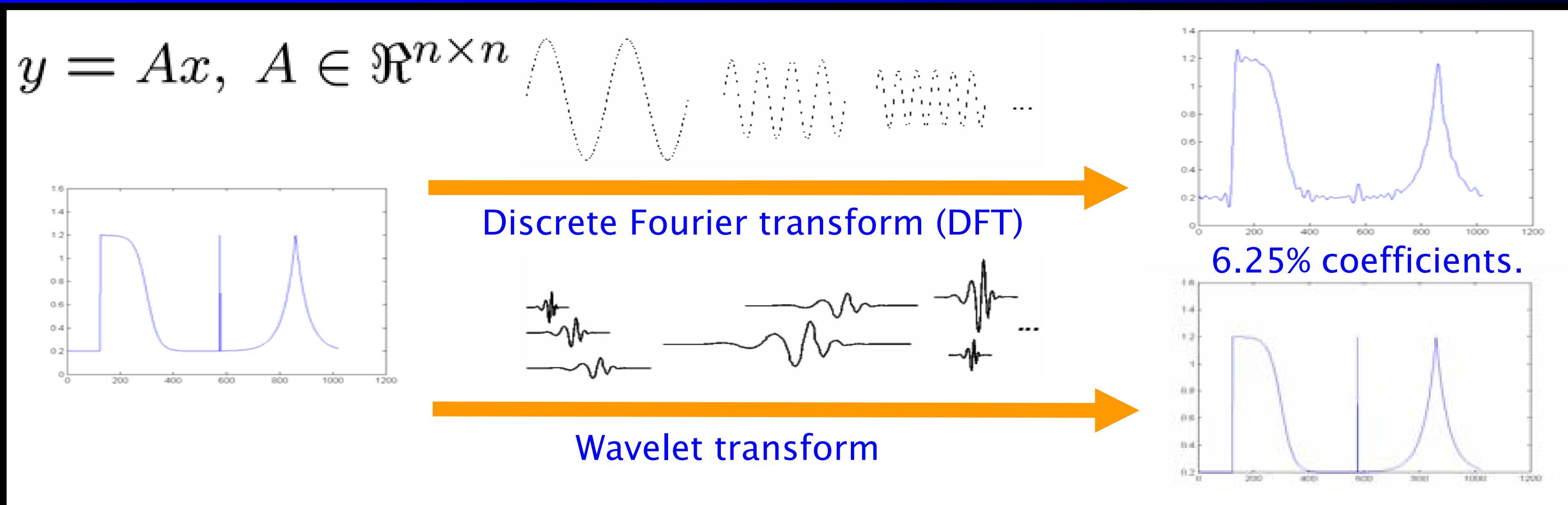
a more compact representation



# Introduction

Fixed Orthogonal Bases (representation, approximation, compression)

- Discrete Fourier transform (DFT) or discrete cosine transform (DCT) (Ahmed '74): JPEG.
- Wavelets (multi-resolution) (Daubechies'88, Mallat'92): JPEG-2000.
- Curvelets and contourlets (Candes & Donoho'99, Do & Veterlli'00)




Unorthogonal Bases (for redundant representations)

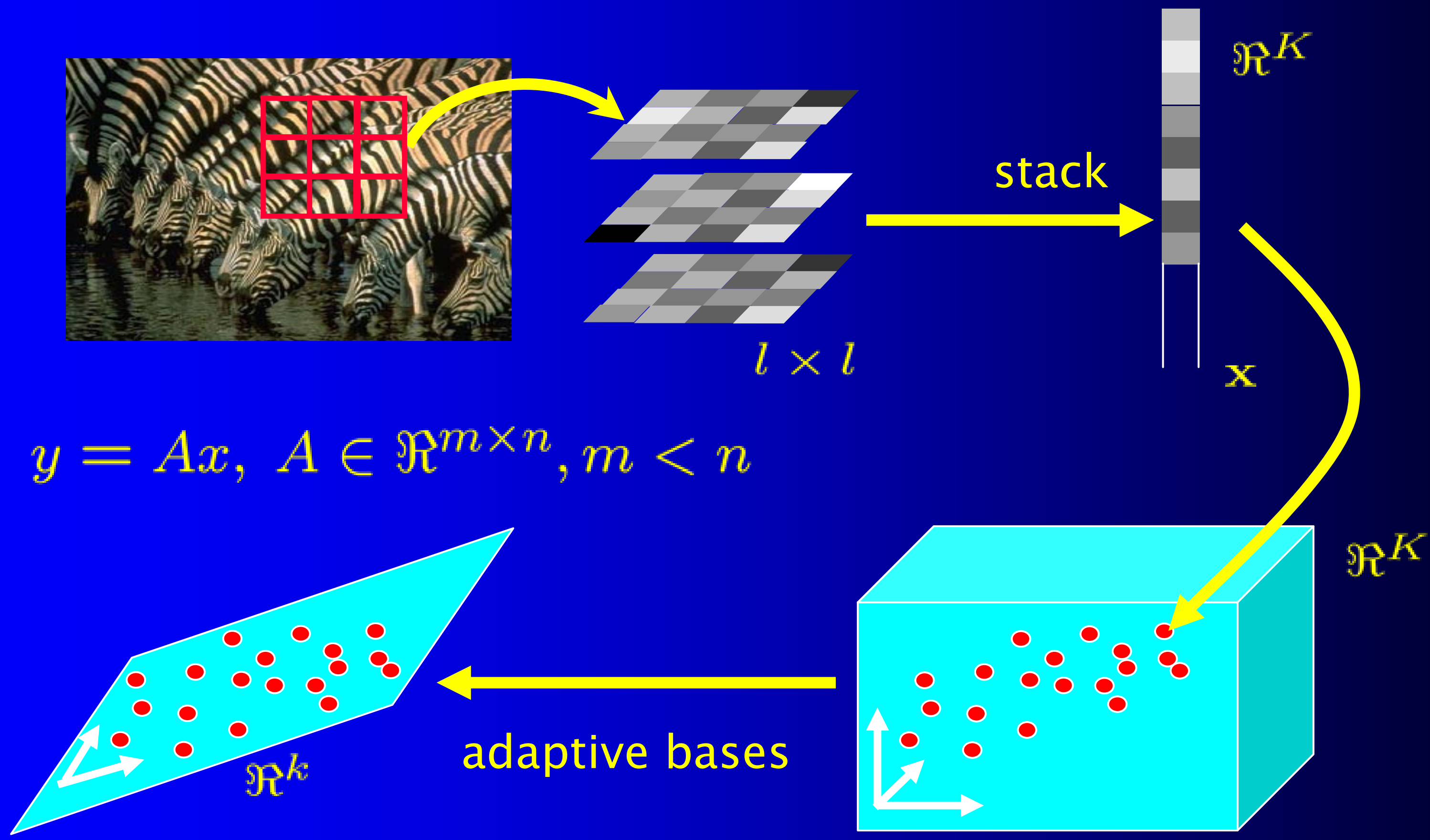
- Extended lapped transforms, frames, sparse representations ($L^p$ geometry)...

$$\min |x|, \text{ s.t } y = Ax, \; A \in \Re^{n \times m}, m > n$$

# Introduction

Adaptive Bases (optimal if imagery data are uni-modal)

- Karhunen-Loeve transform (KLT), also known as PCA (Pearson'1901, Hotelling'33, Jolliffe'86)

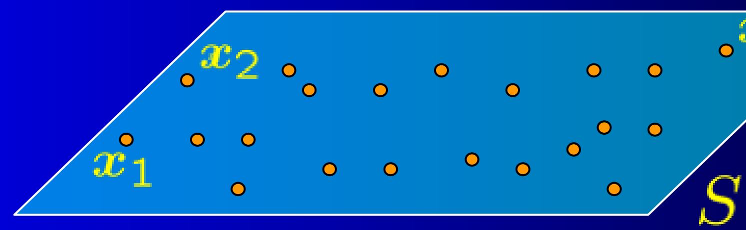


# Introduction – Principal Component Analysis (PCA)

## Dimensionality Reduction

*Find a low-dimensional representation (model) for high-dimensional data.*

Principal Component Analysis (Pearson'1901, Hotelling'1933, Eckart & Young'1936) or Karhunen-Loeve transform (KLT).




$$U \Sigma V^T = [x_1, x_2, \ldots x_N] \in \Re^{K \times N}$$

Basis for $S$

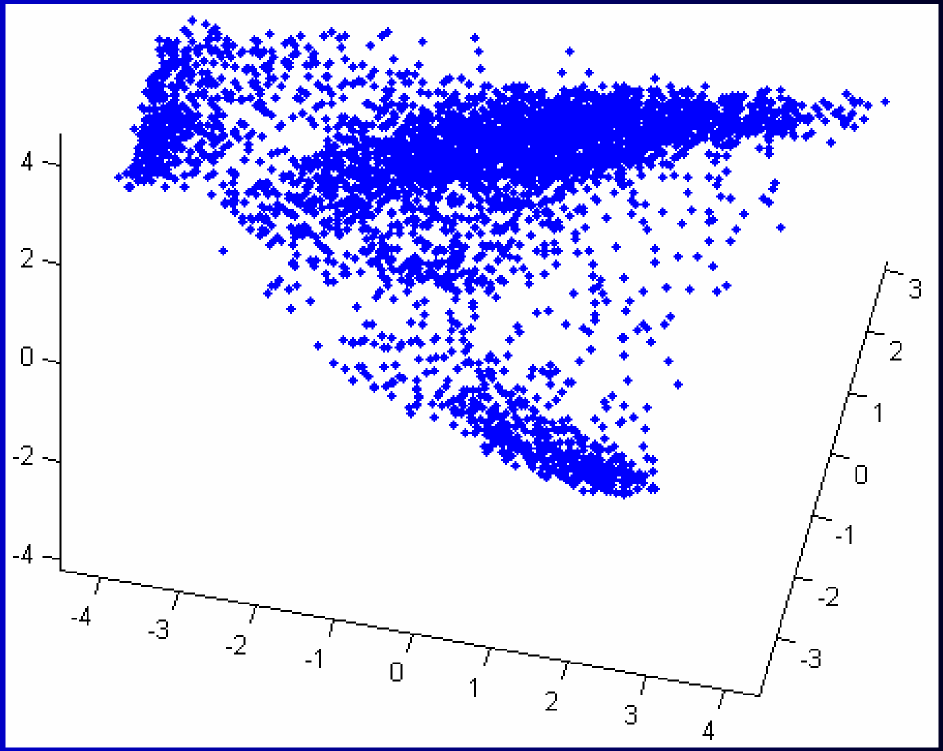
SVD

## Variations of PCA

- Nonlinear Kernel PCA (Scholkopf-Smola-Muller'98)
- Probabilistic PCA (Tipping-Bishop'99, Collins et.al'01)
- Higher-Order SVD (HOSVD) (Tucker'66, Davis'02)
- Independent Component Analysis (Hyvarinen-Karhunen-Oja'01)

# Hybrid Linear Models – Multi-Modal Characteristics

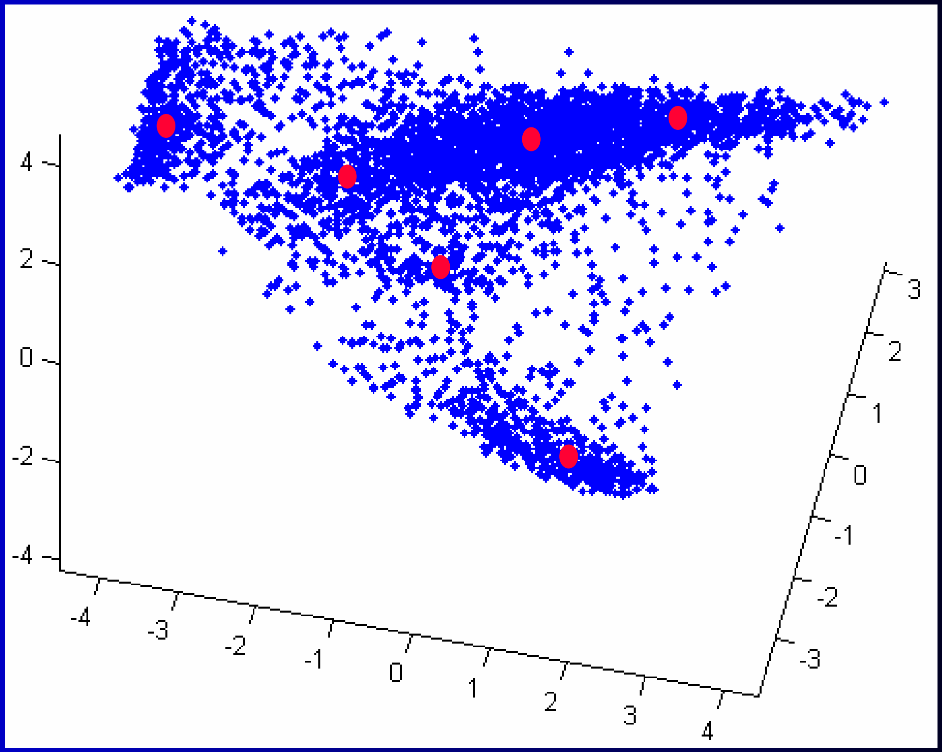
Distribution of the first three principal components of the Baboon image: A clear multi-modal distribution



# Hybrid Linear Models – Multi-Modal Characteristics

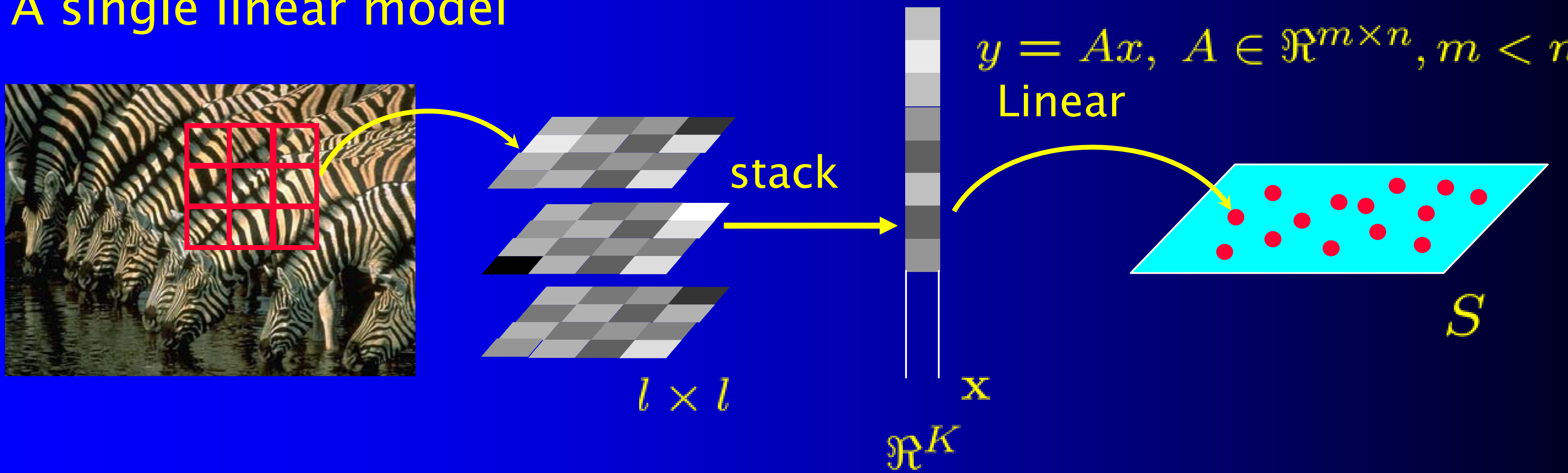
Vector Quantization (VQ)

- multiple 0-dimensional affine subspaces (i.e. cluster means)
- existing clustering algorithms are iterative (EM, K-means)

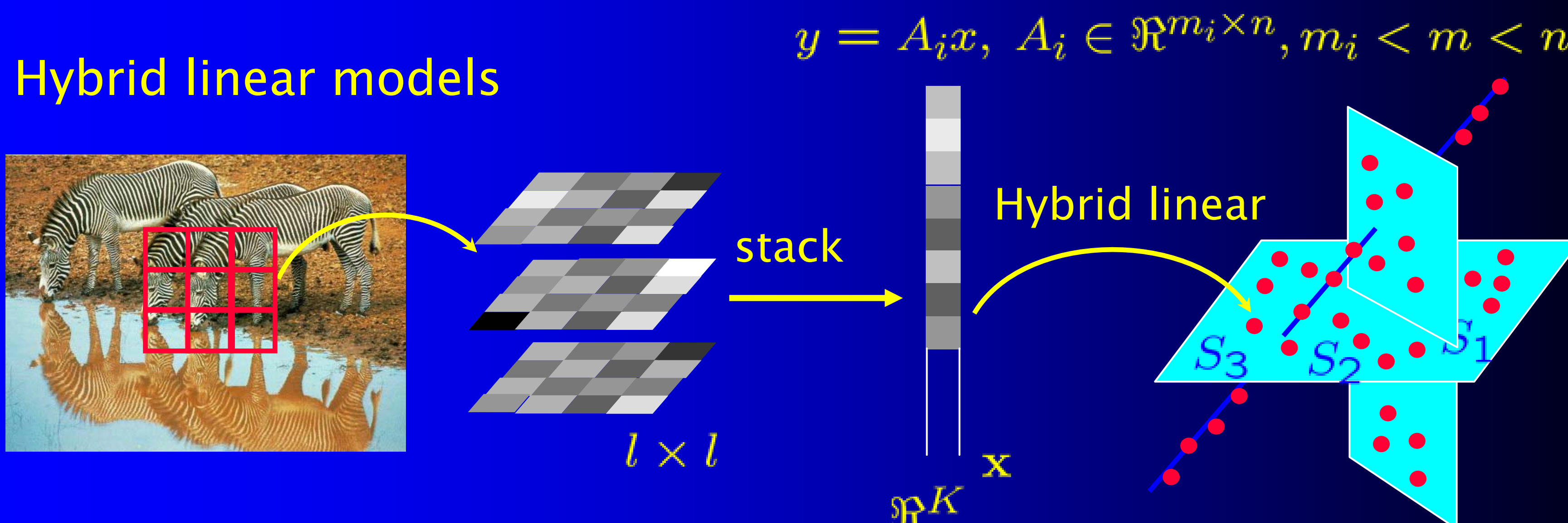


# Hybrid Linear Models – Versus Linear Models

## A single linear model




## Hybrid linear models



# Hybrid Linear Models – Characteristics of Natural Images

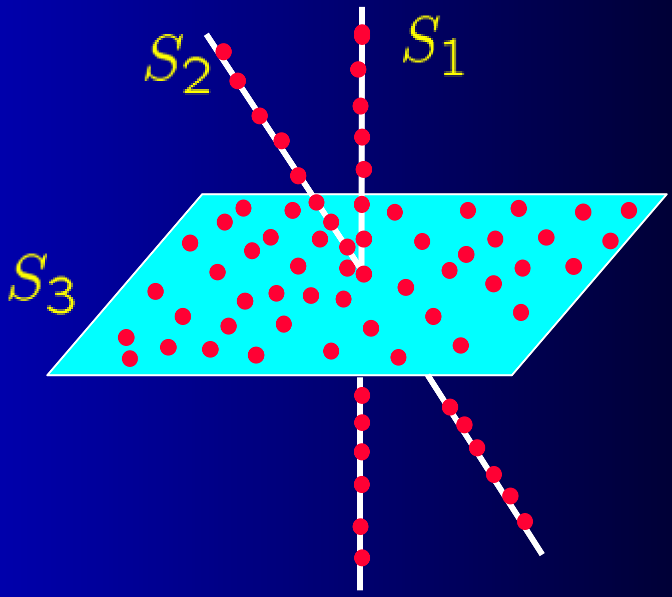
<table>
<tr><th rowspan="2"></th><th colspan="2">Multivariate</th><th rowspan="2">Hybrid<br>(multi-modal)</th><th rowspan="2">Hierarchical<br>(multi-scale)</th><th rowspan="2">High-dimensi<br>(vector-value</th></tr>
<tr><th>1D</th><th>2D</th></tr>
<tr><td>Fourier<br>(DCT)</td><td>X</td><td>X</td><td></td><td></td><td></td></tr>
<tr><td>Wavelets</td><td>X</td><td></td><td rowspan="2"></td><td rowspan="2">X</td><td rowspan="2"></td></tr>
<tr><td>Curvelets</td><td></td><td>X</td></tr>
<tr><td>Random fields</td><td></td><td>X</td><td>X</td><td>X</td><td></td></tr>
<tr><td>PCA/KLT</td><td>X</td><td>X</td><td></td><td></td><td>X</td></tr>
<tr><td>VQ</td><td>X</td><td>X</td><td>X</td><td></td><td>X</td></tr>
<tr><td>Hybrid linear</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td></tr>
</table>

We need a new & simple paradigm to effectively account for all these characteristics simultaneously.

# Hybrid Linear Models – Subspace Estimation and Segmentation

**Hybrid Linear Models (or Subspace Arrangements)**

- the number of subspaces is unknown
- the dimensions of the subspaces are unknown
- the basis of the subspaces are unknown
- the segmentation of the data points is unknown

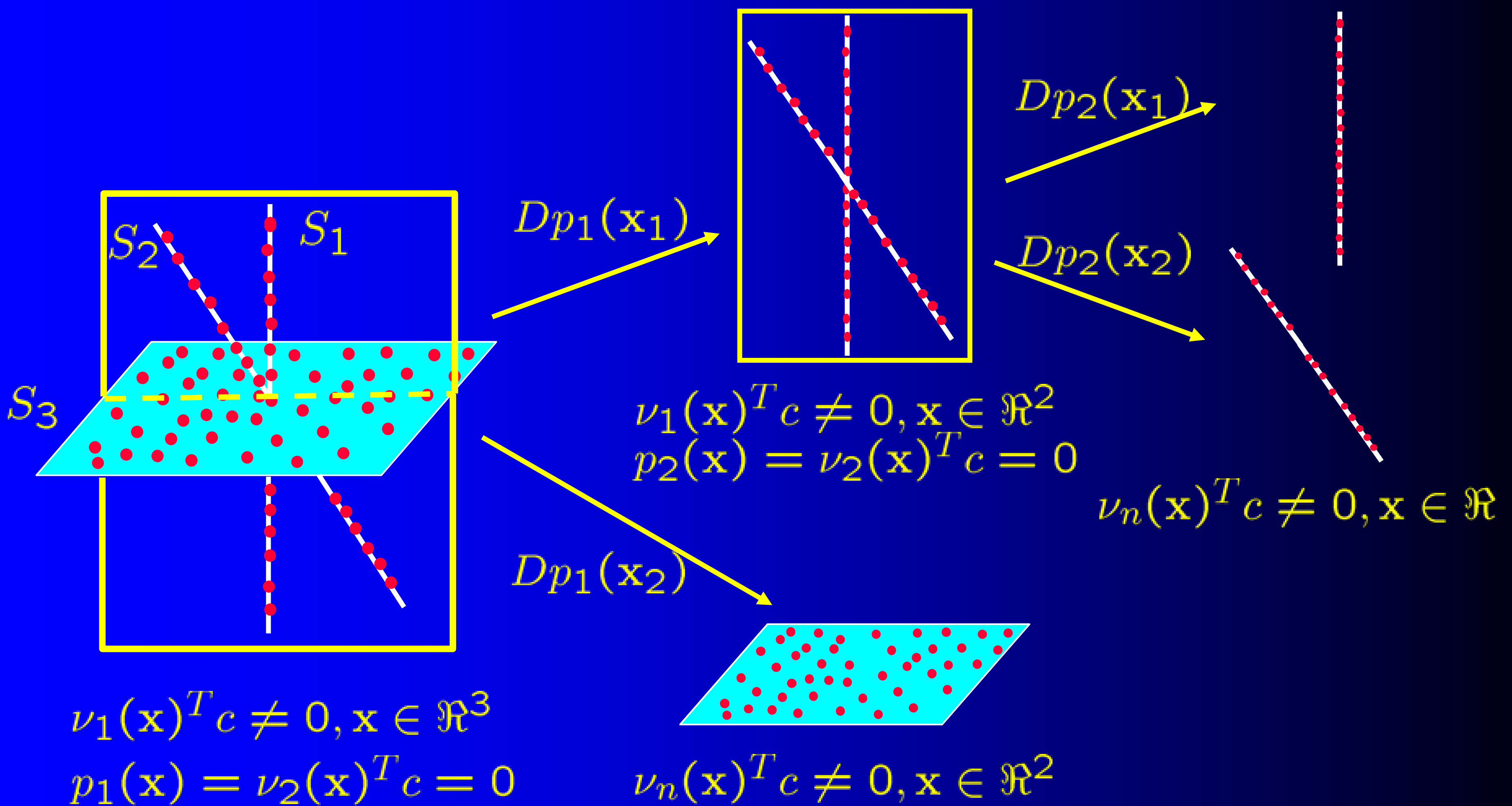



"Chicken-and-Egg" Coupling

- Given segmentation, estimate subspaces
- Given subspaces, segment the data

# Hybrid Linear Models – Recursive GPCA (an Example)

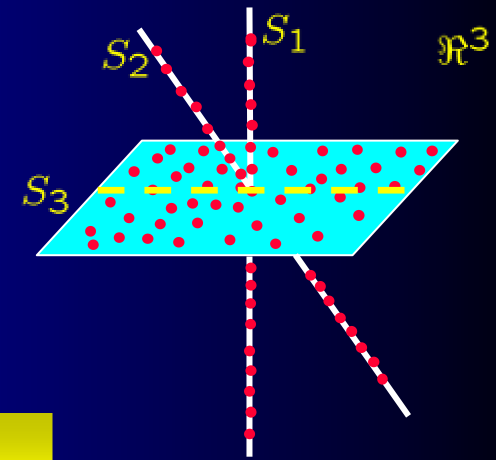
$$p(\mathbf{x}) = c_1 x_1^n + c_2 x_1^{n-1} x_2 + \cdots + c_m x_3^n = \nu_n(\mathbf{x})^T c$$



# Hybrid Linear Models – Effective Dimension

## Model Selection (for Noisy Data)

- Model complexity;
- Data fidelity;




Number of subspaces

$$ED(\mathbf{X}, S) \doteq \frac{1}{N}\sum_{1}^{s} k_i(K - k_i) + \frac{1}{N}\sum_{1}^{s} N_i k_i$$

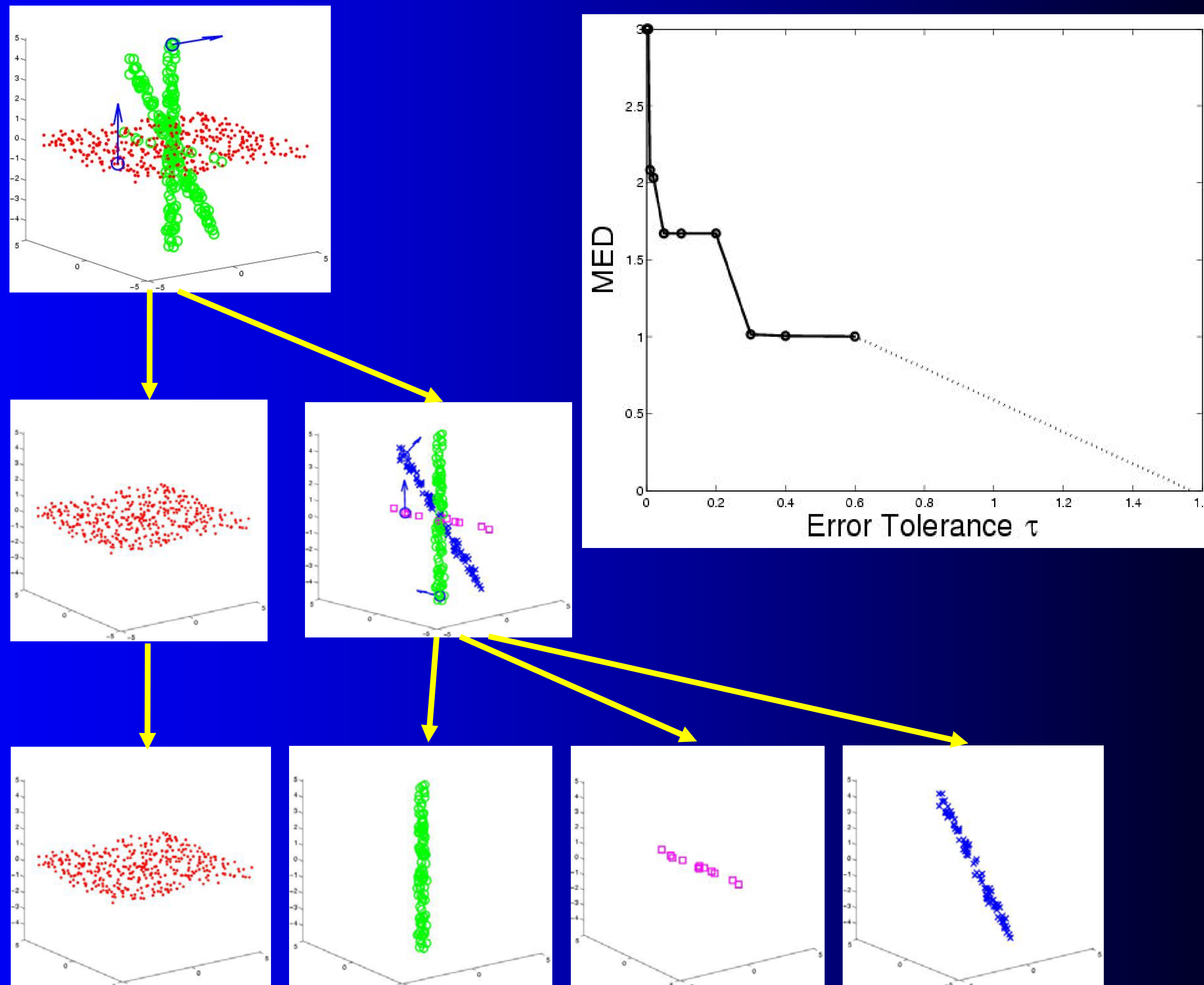
Total number of points

Dimension of each subspace

Number of points in each subspace

Model selection criterion: minimizing effective dimension subject to a given error tolerance (or PSNR)

# Hybrid Linear Models – Simulation Results (5% Noise)

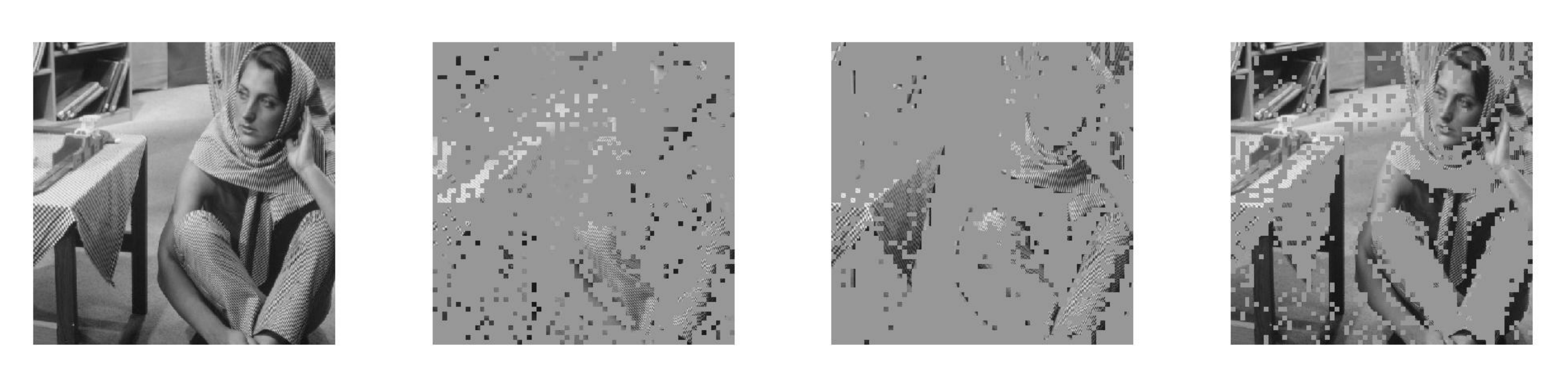


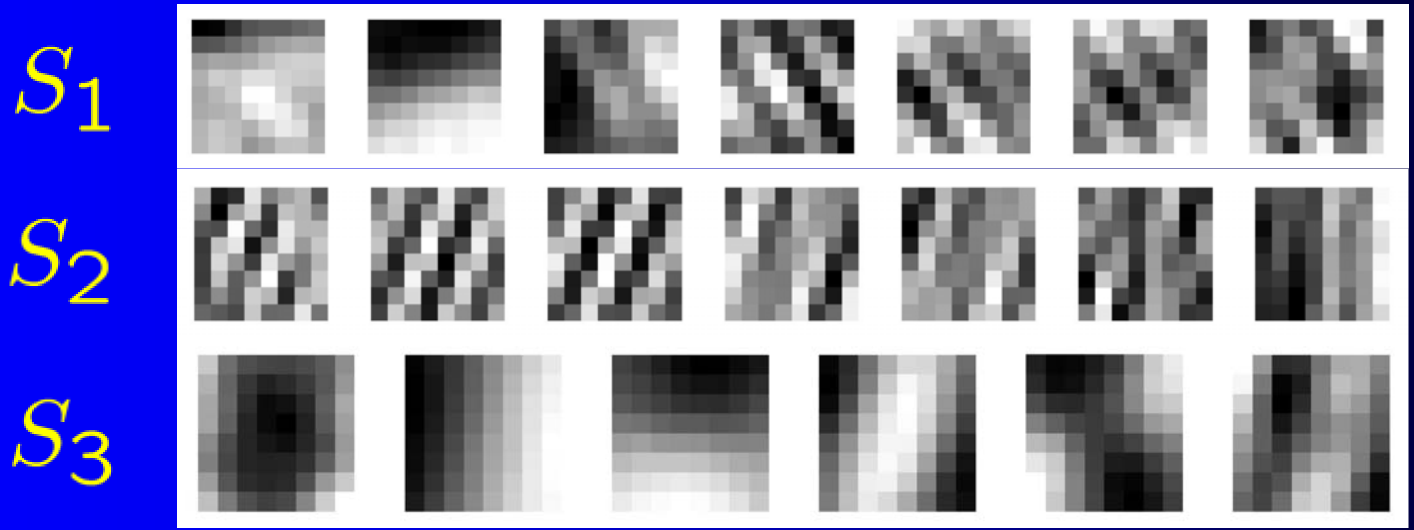

ED=3

ED=2.0067

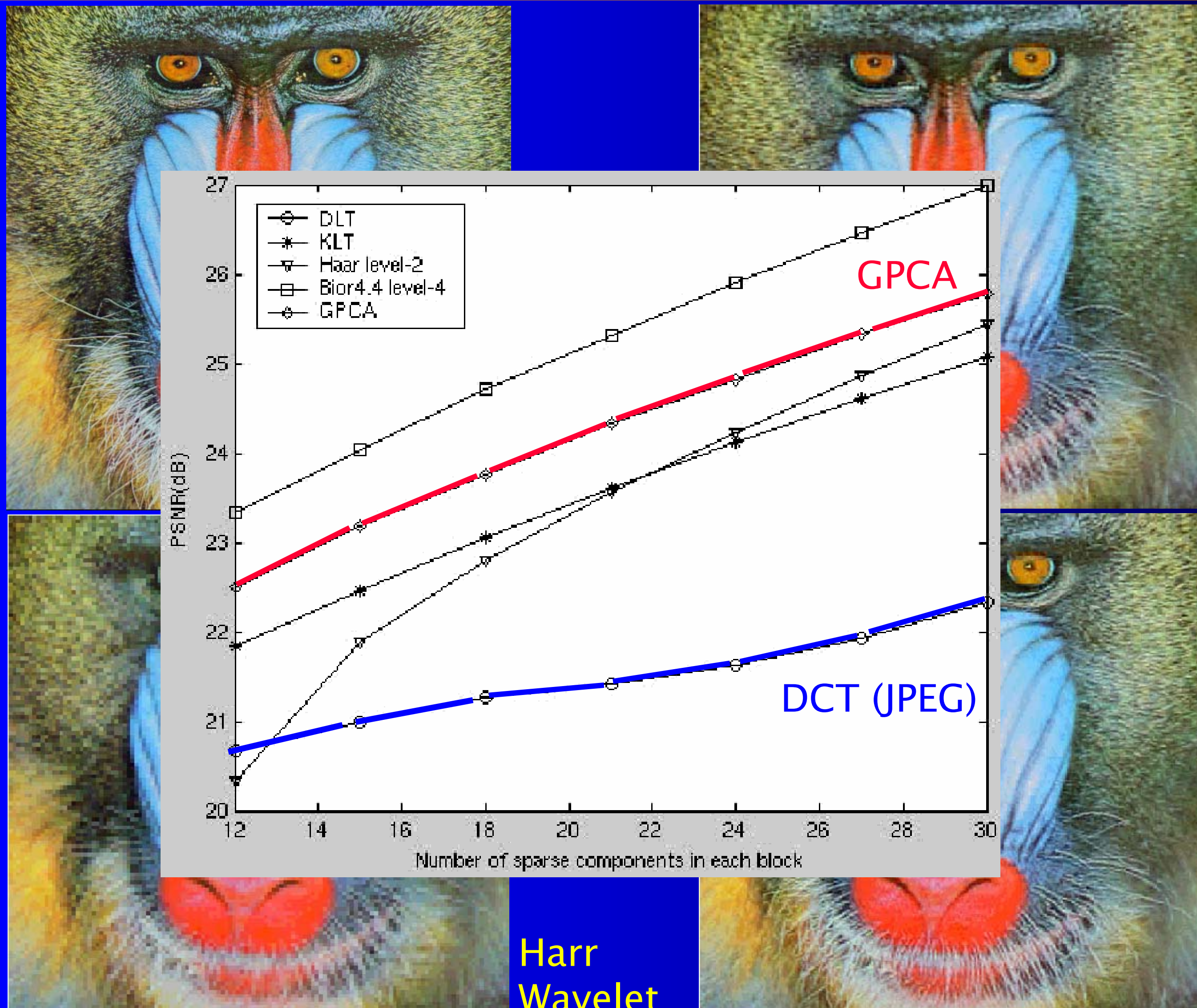
ED=1.6717

# Hybrid Linear Models – Subspaces of the Barbara Image





# Hybrid Linear Models – Lossy Image Representation (Baboon)



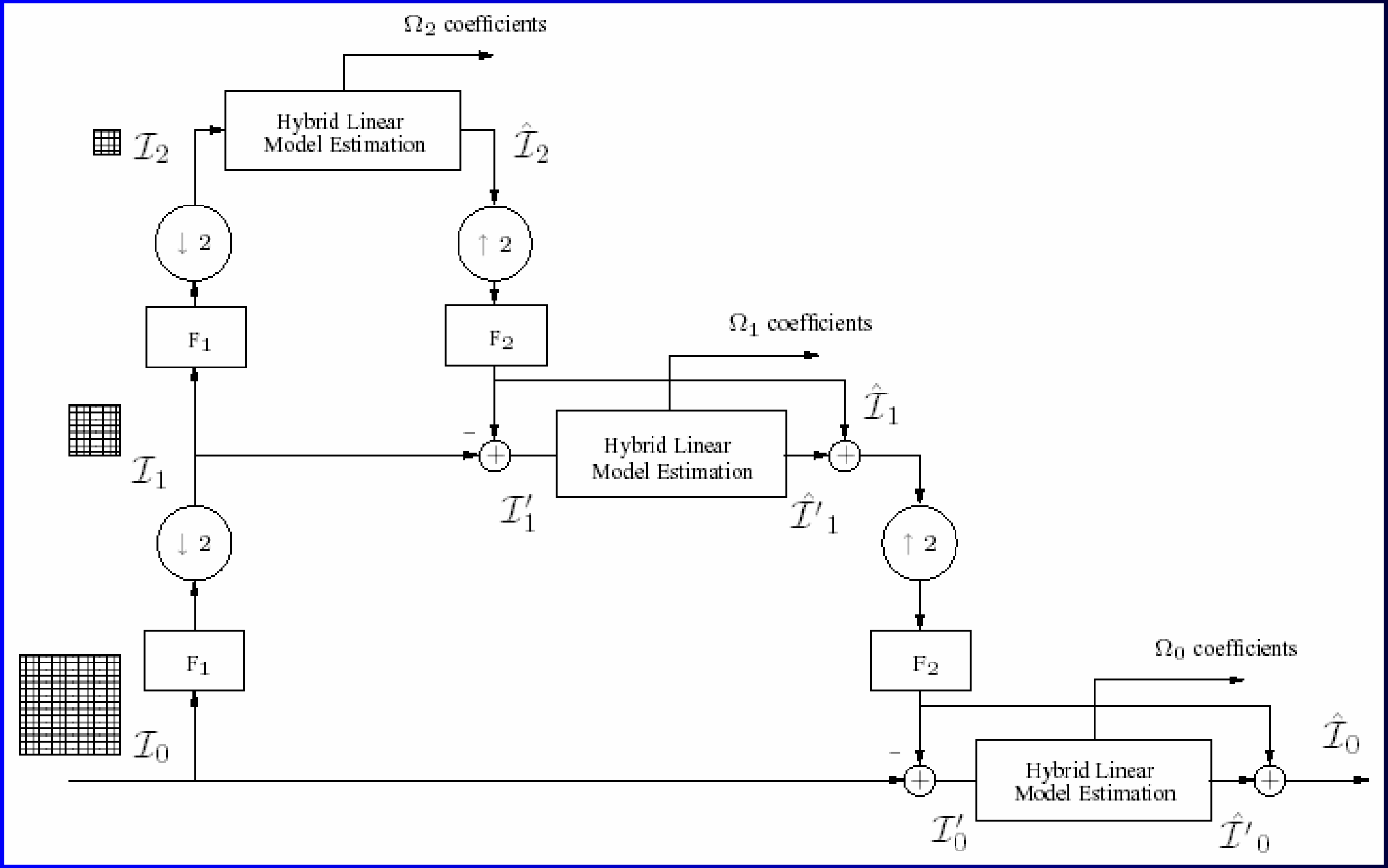

PCA (8x8)

Harr Wavelet

GPCA (8x8)

# Multi-Scale Implementation – Algorithm Diagram

Diagram for a level-3 implementation of hybrid linear models for image representation

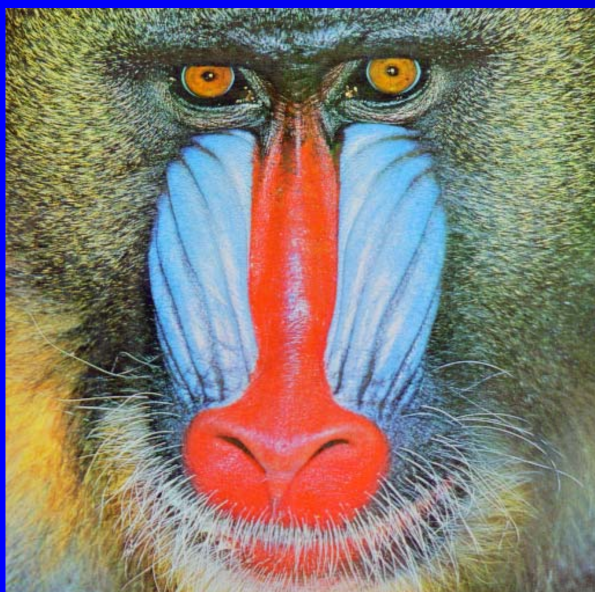



*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

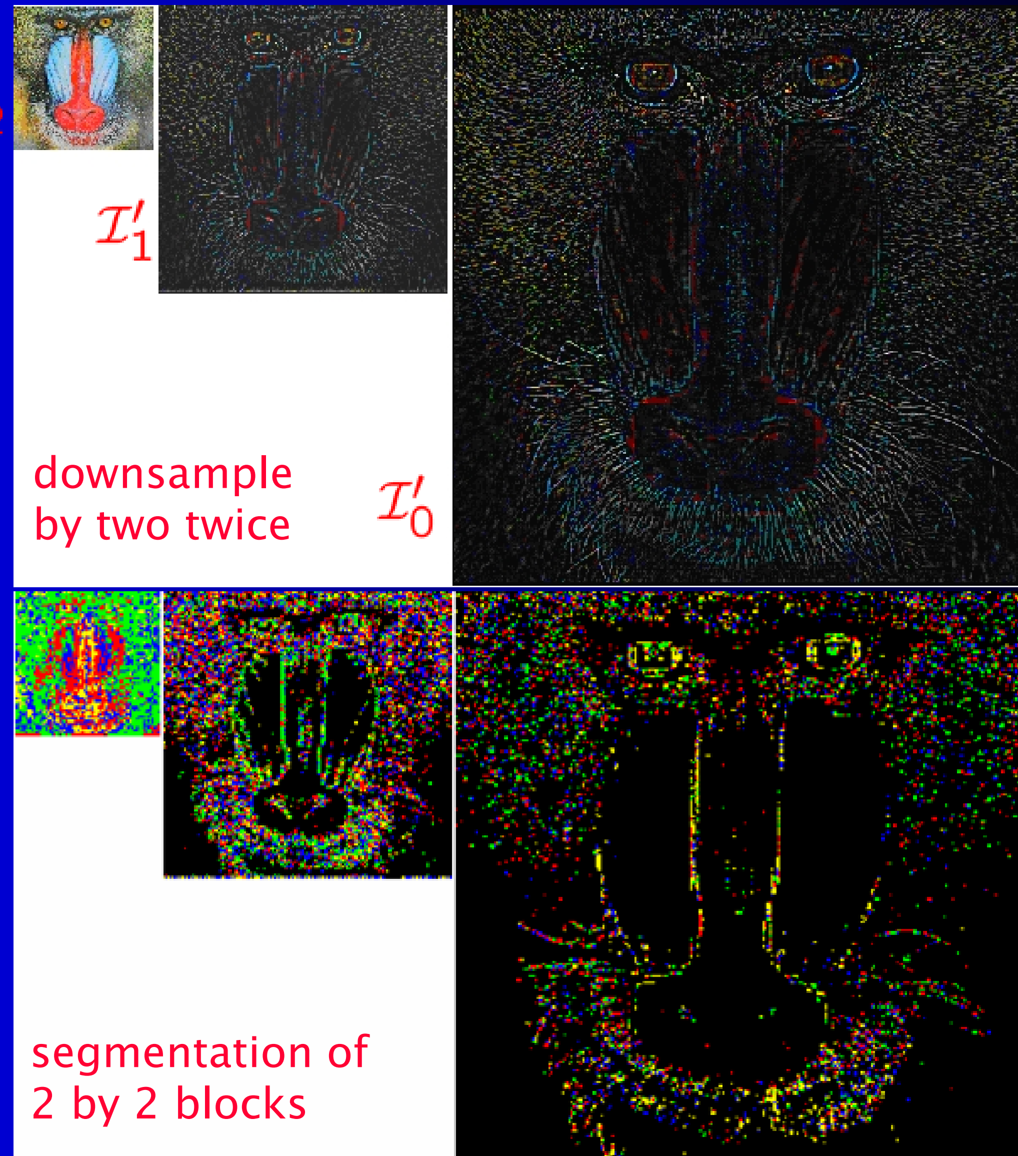
# Multi-Scale Implementation – The Baboon Image

The Baboon image

$\mathcal{I}$



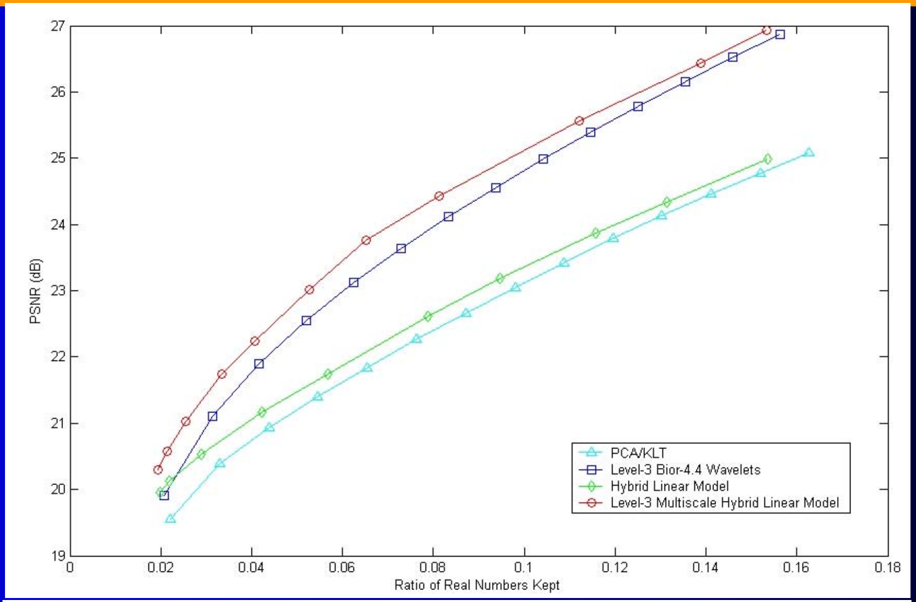
$\mathcal{I}_2$

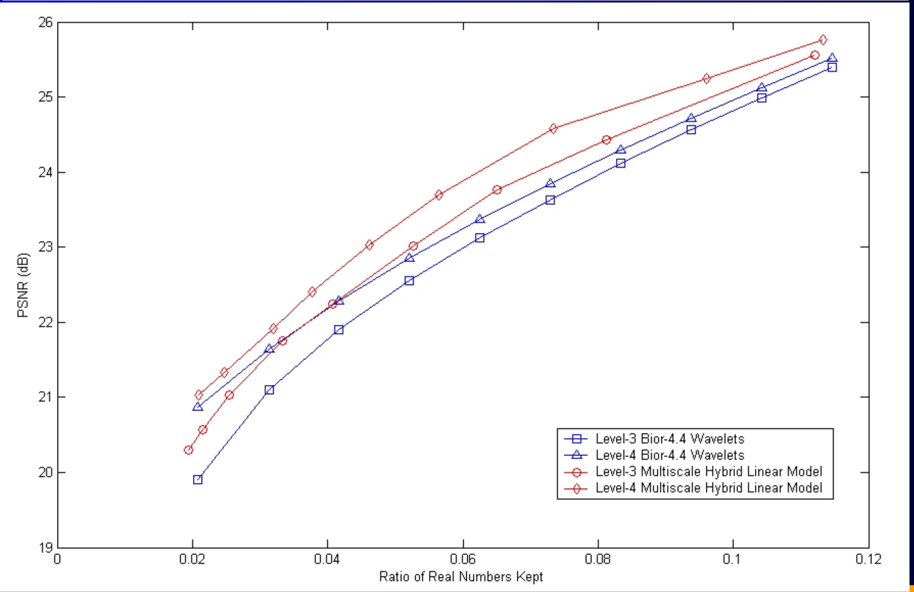



*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

# Multi-Scale Implementation – Comparison with Other Methods

The Baboon image




*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

# Multi-Scale Implementation – Image Approximation

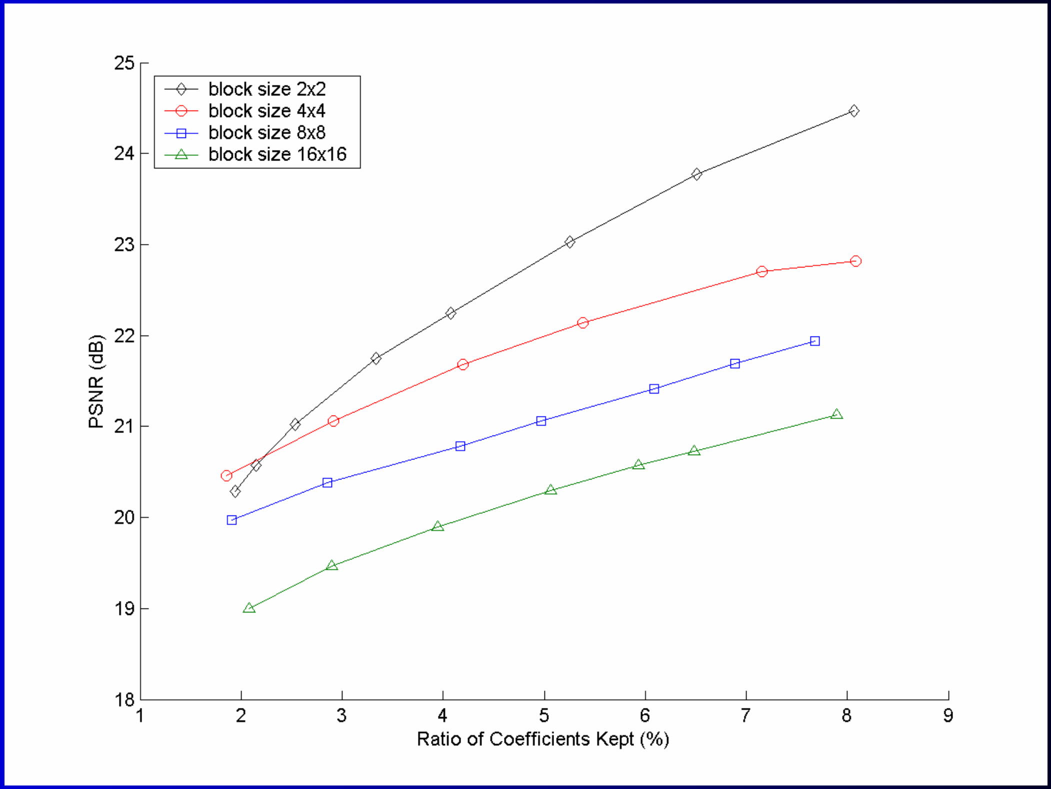
## Comparison with level-3 wavelet (7.5% coefficients)

Level-3 bior-4.4 wavelets
PSNR=23.94

Level-3 hybrid linear model
PSNR=24.64

# Multi-Scale Implementation – Block Size Effect

The Baboon image



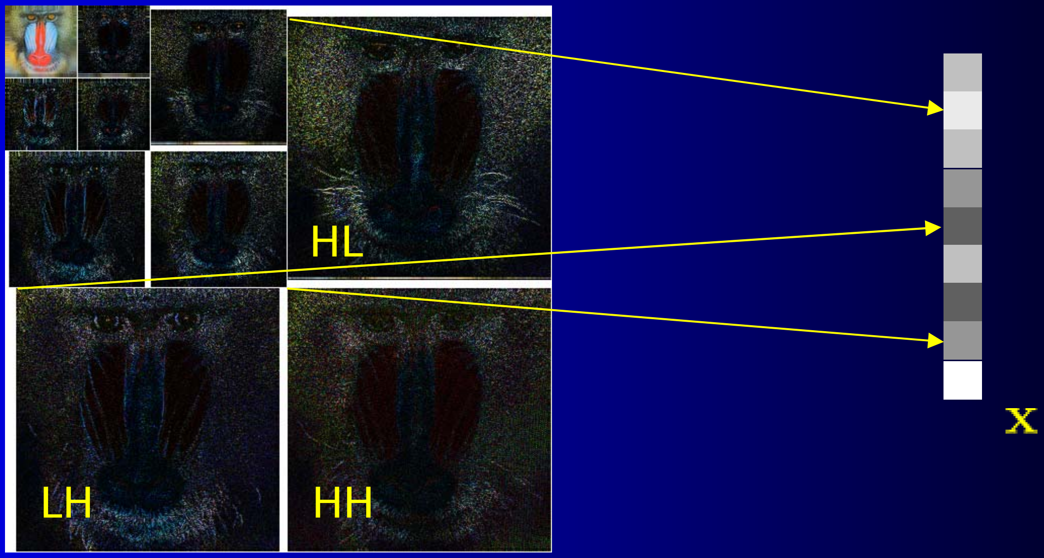

Some problems with the multi-scale hybrid linear model:

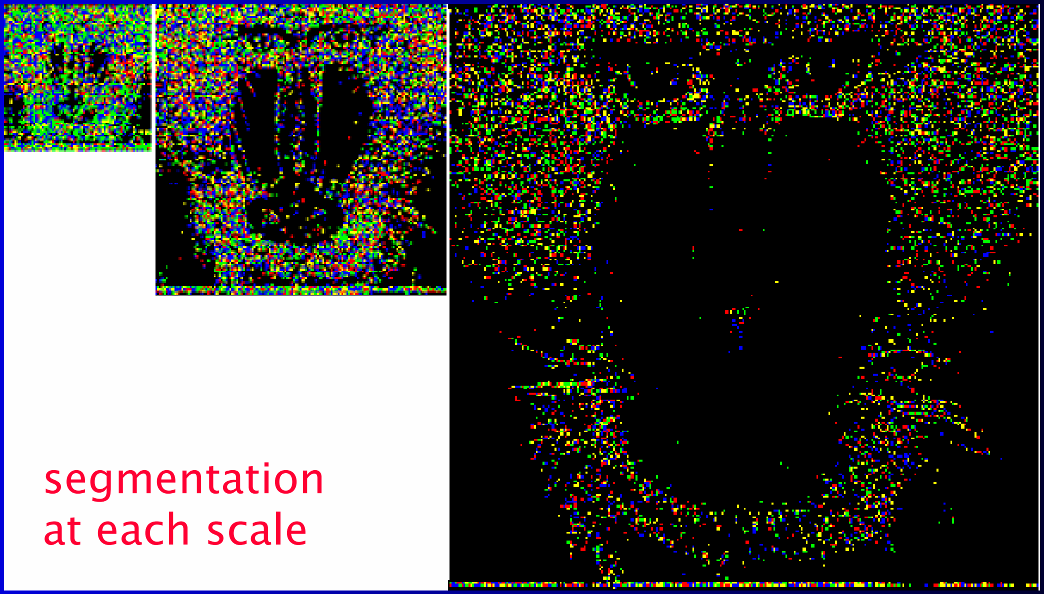
1. has minor block effect;
2. is computationally more costly (than Fourier, wavelets, PCA);
3. does not fully exploit spatial smoothness as wavelets.

*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

# Multi-Scale Implementation – The Wavelet Domain

The Baboon image

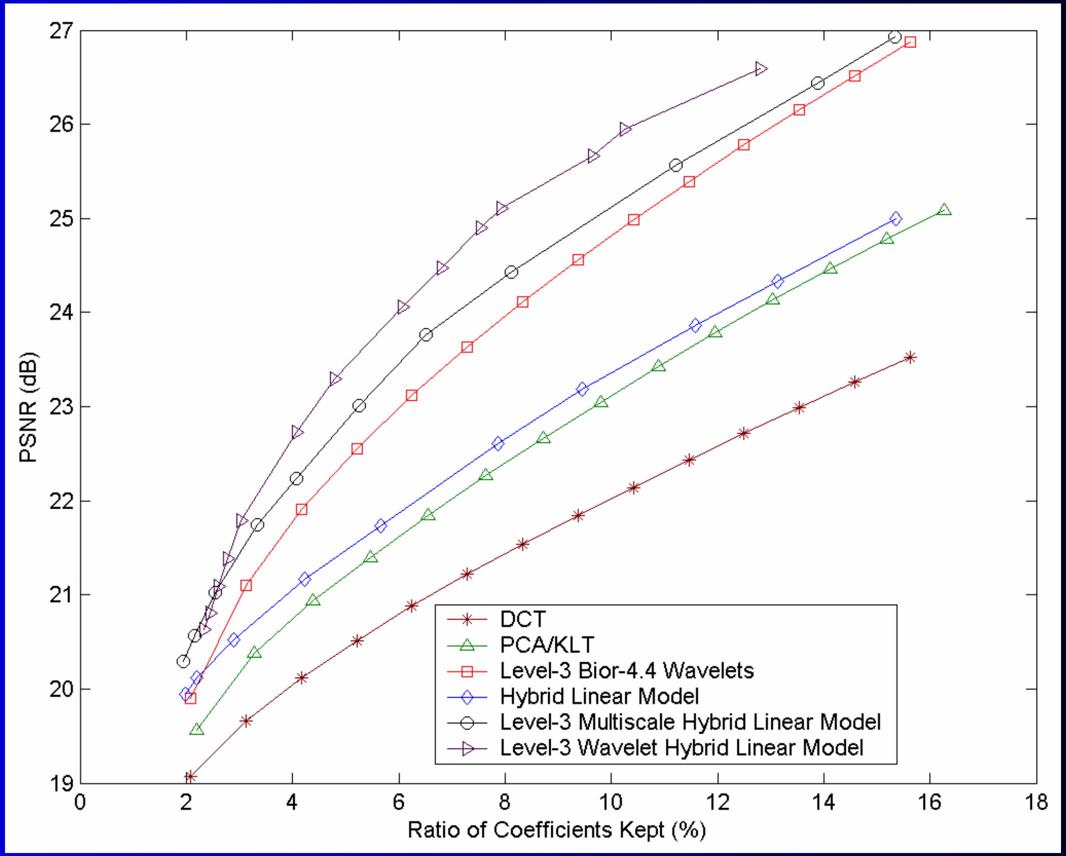



*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

# Multi-Scale Implementation – Wavelets v.s. Hybrid Linear Wavelets

The Baboon image




Advantages of the hybrid linear model in wavelet domain:

1. eliminates block effect;
2. is computationally less costly (than in the spatial domain);
3. achieves higher PSNR.

*Multi-Scale Hybrid Linear Models for Lossy Image Representation*, [Hong-Wright-Ma, TI

## Comparison among several models (7.5% coefficients)

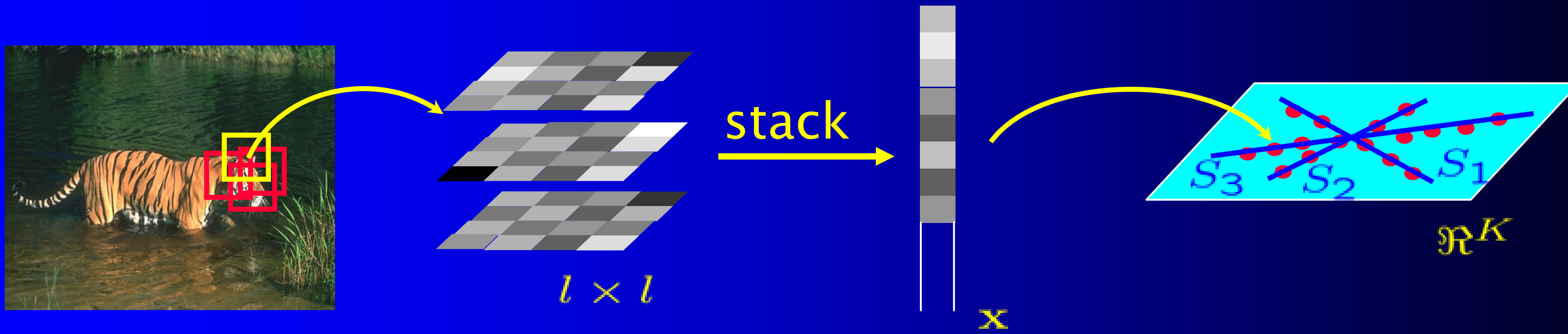
Original Image

Wavelets PSNR=23.94

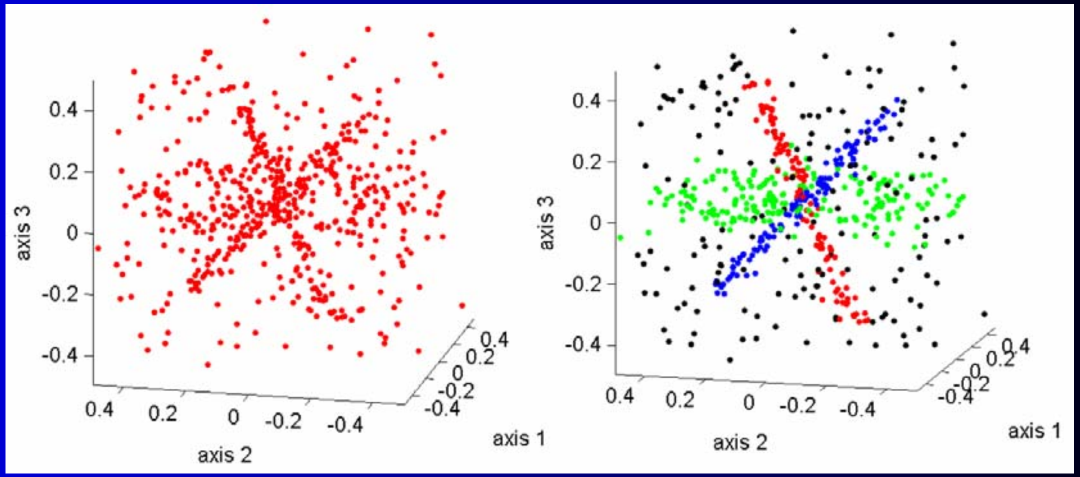
Hybrid model in spatial domain PSNR=24.64

Hybrid model in wavelet domain PSNR=24.88

# Image Segmentation – via Lossy Data Compression




$$\min L^s(X) = L(X_1) + L(X_2) + \cdots + L(X_n) + H(|X_1|, |X_2|, \ldots, |X_n|).$$



# APPLICATIONS – Texture-Based Image Segmentation

Naïve approach:

- Take a 7x7 Gaussian window around every pixel.
- Stack these windows as vectors.
- Clustering the vectors using our algorithm.

A few results:

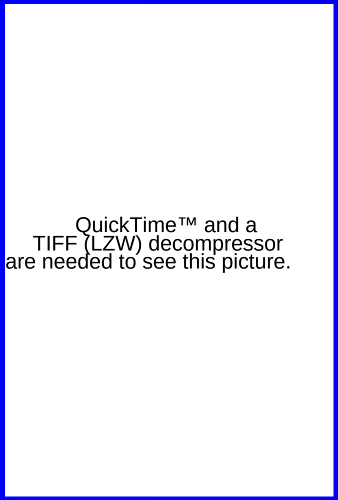
*Segmentation of Multivariate Mixed Data via Lossy Coding and Compression*, [Ma-Derksen-Hong-Wright, PAMI'07]

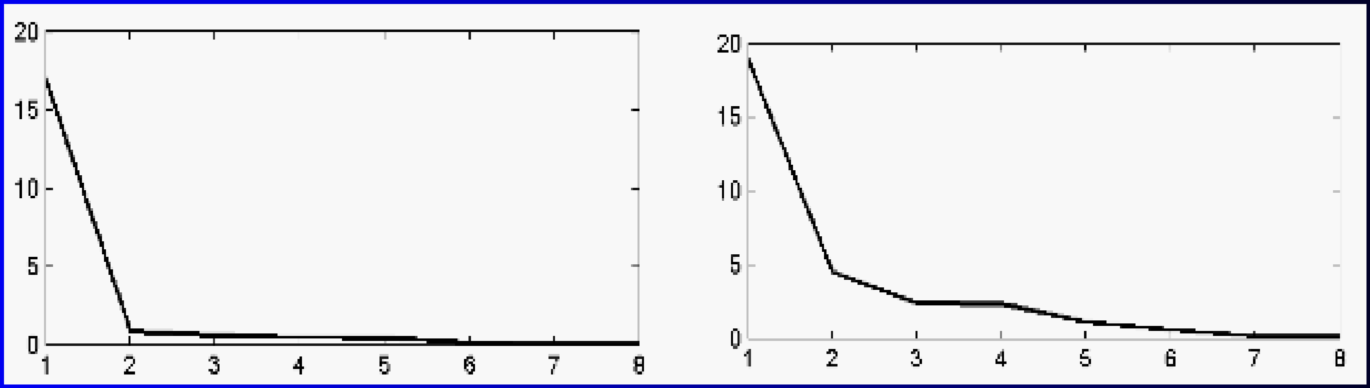
# APPLICATIONS – Distribution of Texture Features

**Question**: why does such a simple algorithm work at all?

**Answer**: Compression (MDL/MCL) is well suited to mid-level texture segmentation.

Using a single representation (e.g. windows, filterbank responses) for texture different complexity ⇒ redundancy and degeneracy, which can be exploited f clustering / compression.




Above: singular values of feature vectors from two different segments of the image at left.

# APPLICATIONS – Compression-based Texture Merging (CTM)

## Problem with the naïve approach:

Strong edges, segment boundaries

QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.

QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.

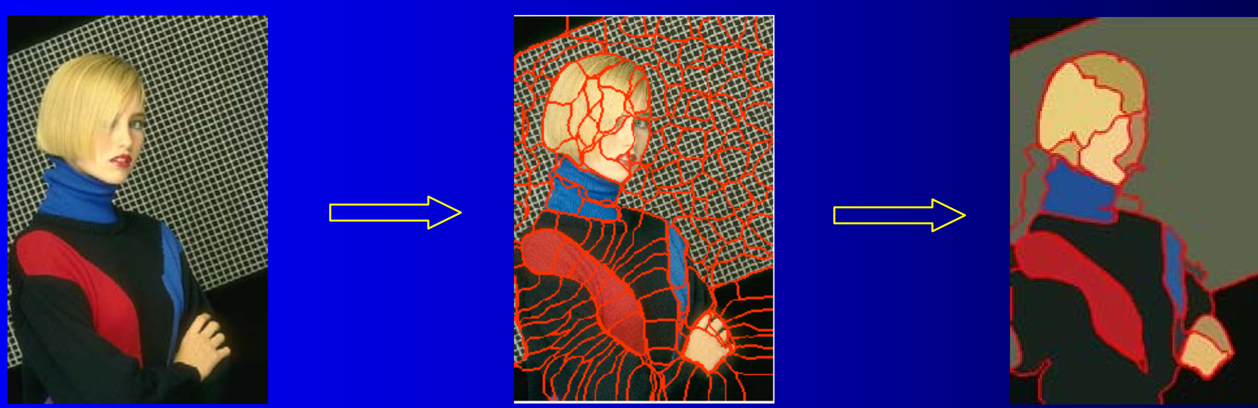
QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.

## Solution:

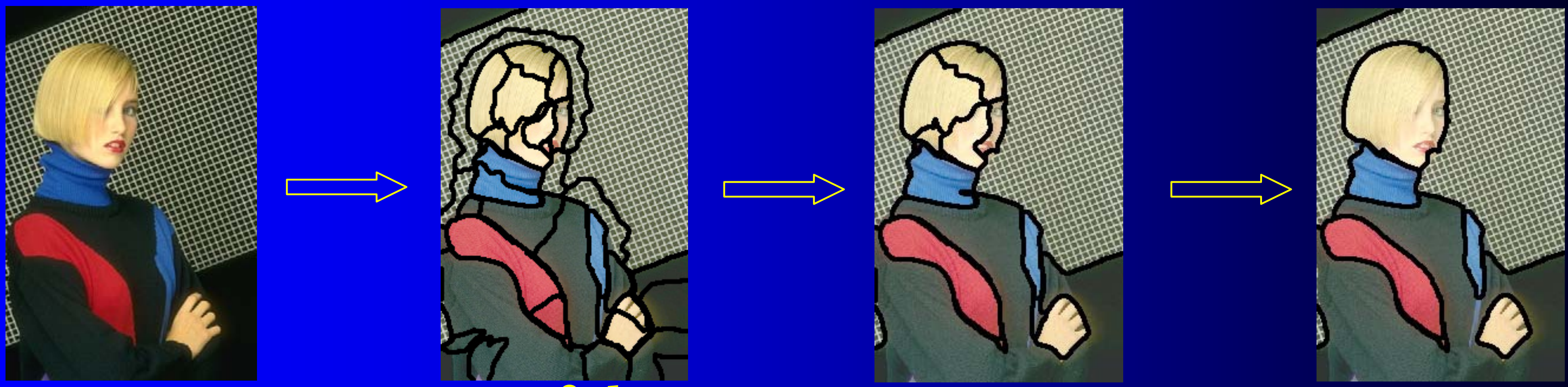
Low-level, edge-preserving **over-segmentation** into small homogeneous regions.

**Simple features**: stacked Gaussian windows (7x7 in our experiments).

Merge adjacent regions to **minimize coding length** (“compress” the features).

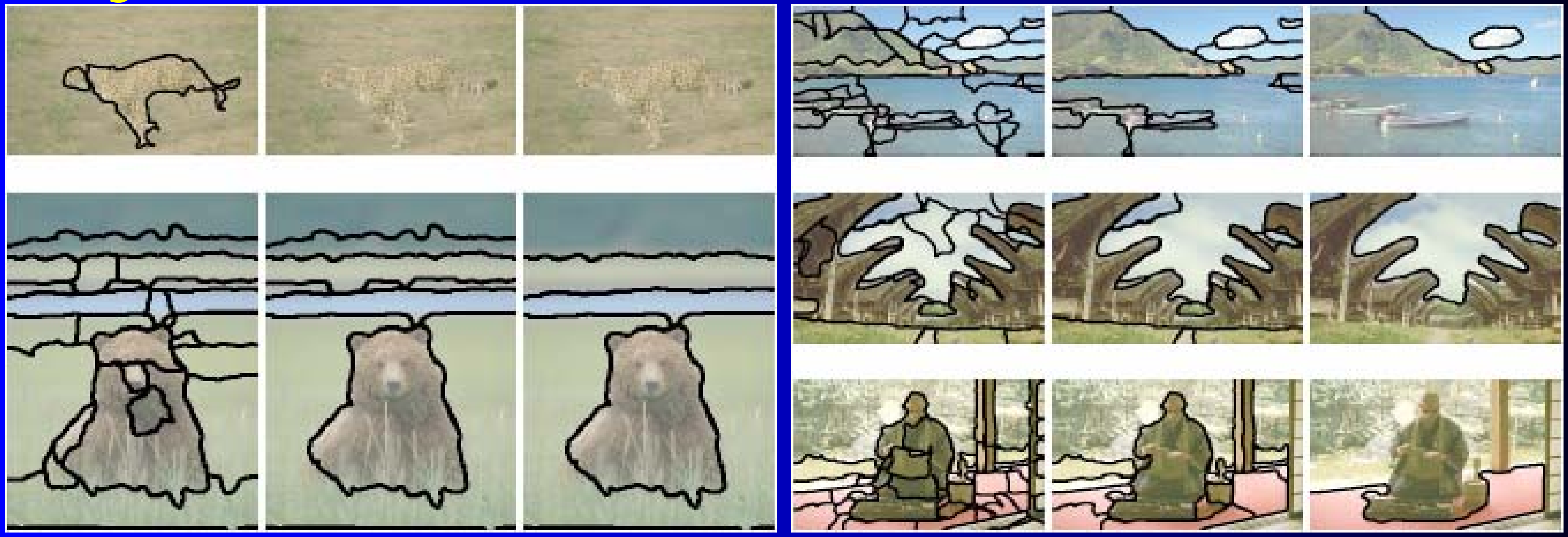


# APPLICATIONS – Hierarchical Image Segmentation via CTM

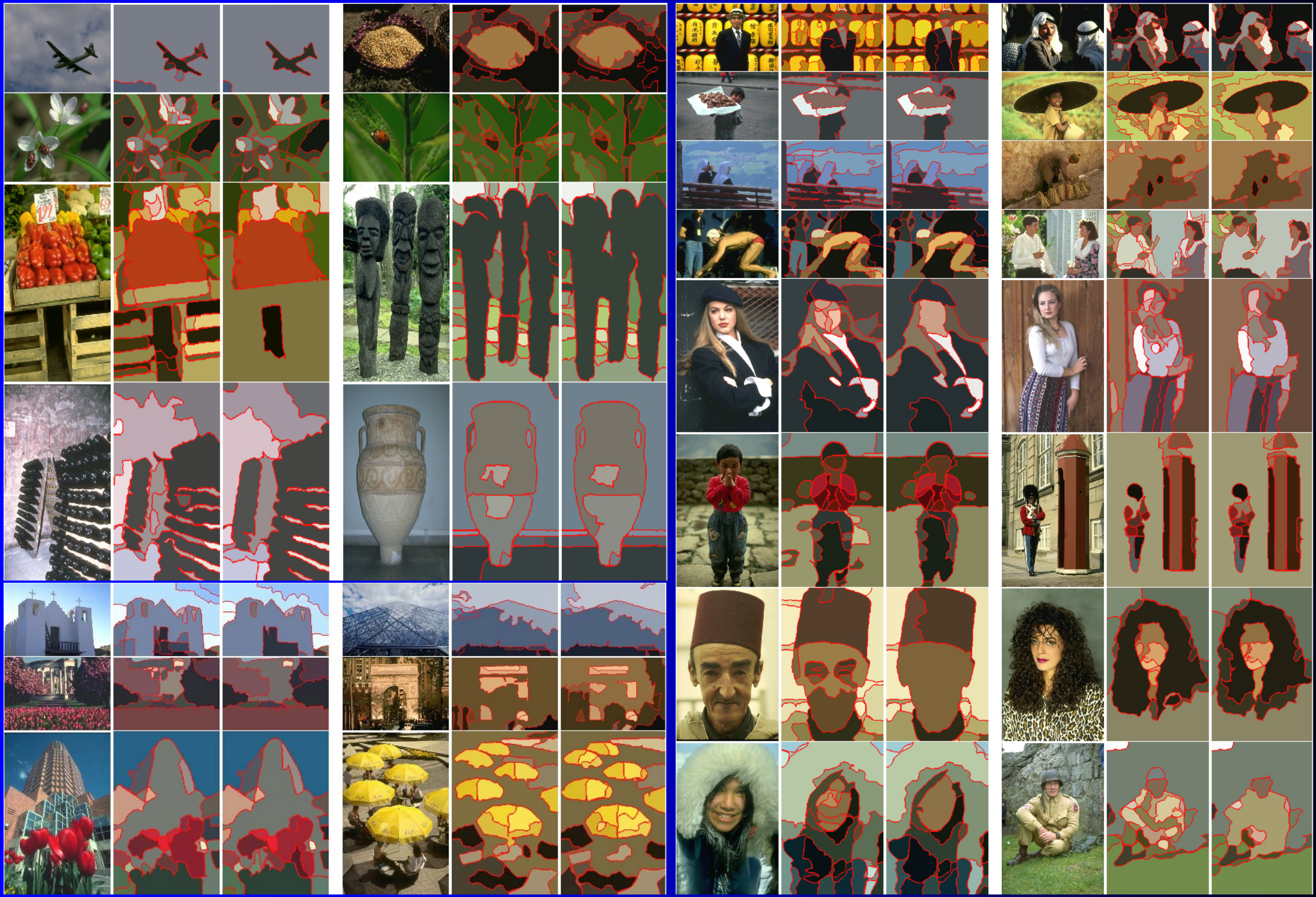


$\varepsilon = 0.1$ $\varepsilon = 0.2$ $\varepsilon = 0.4$

Lossy coding with varying distortion $\varepsilon$ => hierarchy of segmentations



# APPLICATIONS – CTM: Qualitative Results



# APPLICATIONS – CTM: Quantitative Evaluation and Comparison

Berkeley Image Segmentation Database

| | PRI | VoI | GCE | BDE |
|---|---|---|---|---|
| Humans | 0.8754 | 1.1040 | 0.0797 | 4.994 |
| CTM (0.1) | 0.7561 | 2.4640 | **0.1767** | **9.4211** |
| CTM (0.15) | 0.7627 | 2.2035 | 0.1846 | 9.4902 |
| CTM (0.2) | 0.7617 | **2.0236** | 0.1877 | 9.8962 |
| Mean-Shift | 0.7550 | 2.477 | 0.2598 | 9.7001 |
| NCuts | 0.7229 | 2.9329 | 0.2182 | 9.6038 |
| FH | **0.7841** | 2.6647 | 0.1895 | 9.9497 |

PRI: Probabilistic Rand Index [Pantofaru 2005]

VoI: Variation of Information [Meila 2005]

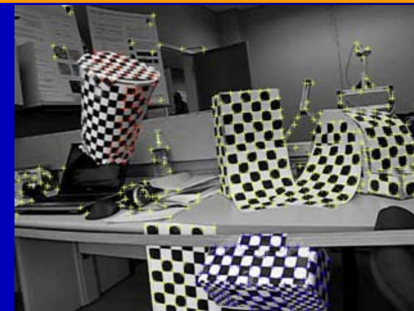
GCE: Global Consistency Error [Martin 2001]

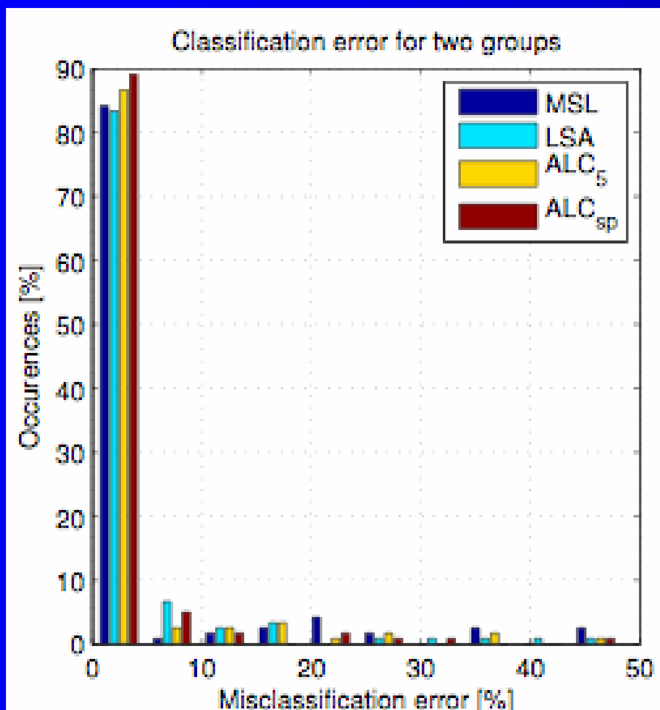
BDE: Boundary Displacement Error [Freixenet 2002]

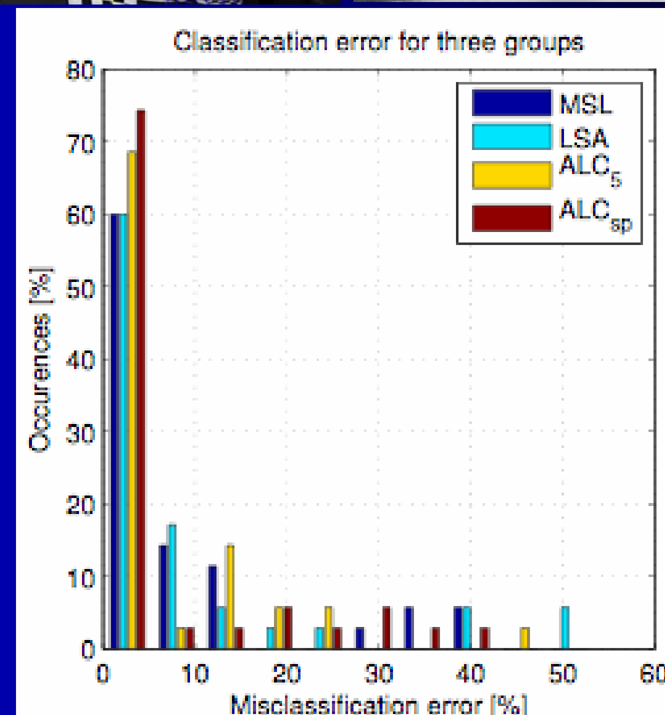
# Other Applications: Multiple Motion Segmentation (on Hopkins155)

QuickTime™ and a Cinepak decompressor are needed to see this picture.

QuickTime™ and a Cinepak decompressor are needed to see this picture.






Two Motions: MSL 4.14%, LSA 3.45%, ALC 2.40%, and work with up to 25% outliers.

Three Motions: MSL 8.32%, LSA 9.73%, ALC 6.26%.

Shankar Rao, Roberton Tron, Rene Vidal, and Yi Ma, to appear in CVPR'08

# Other Applications – Clustering of Microarray Data

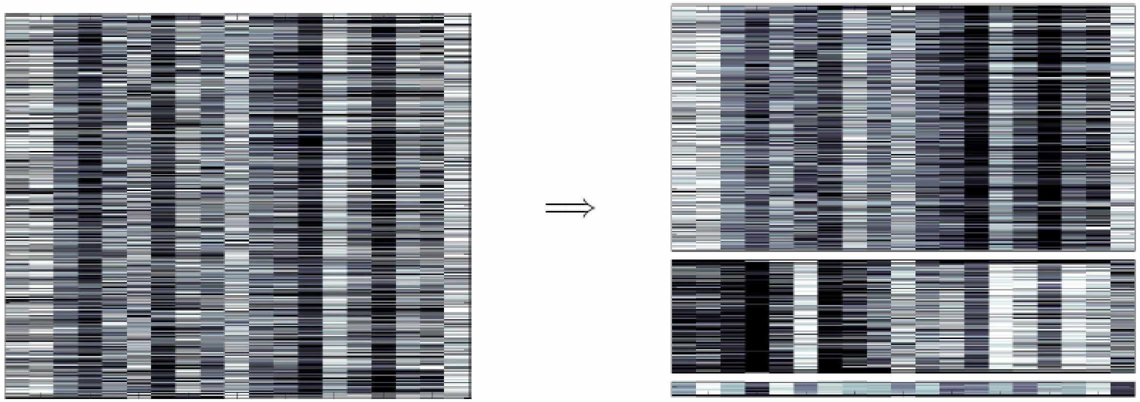
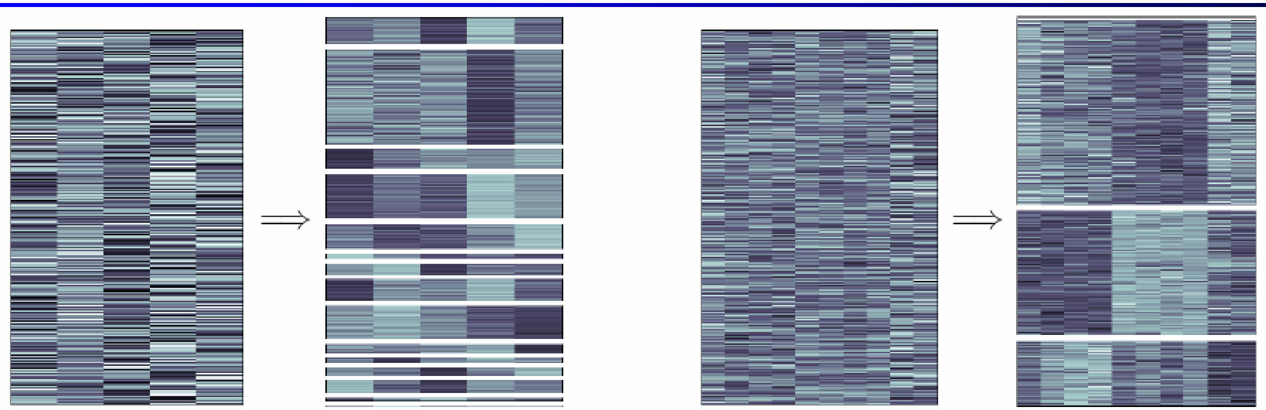


Fig. 14. Segmentation of microarray data. Left: raw data. Each row represents the expression level of a single gene. Right: Three distinct clusters are found, visualized by reordering the rows.

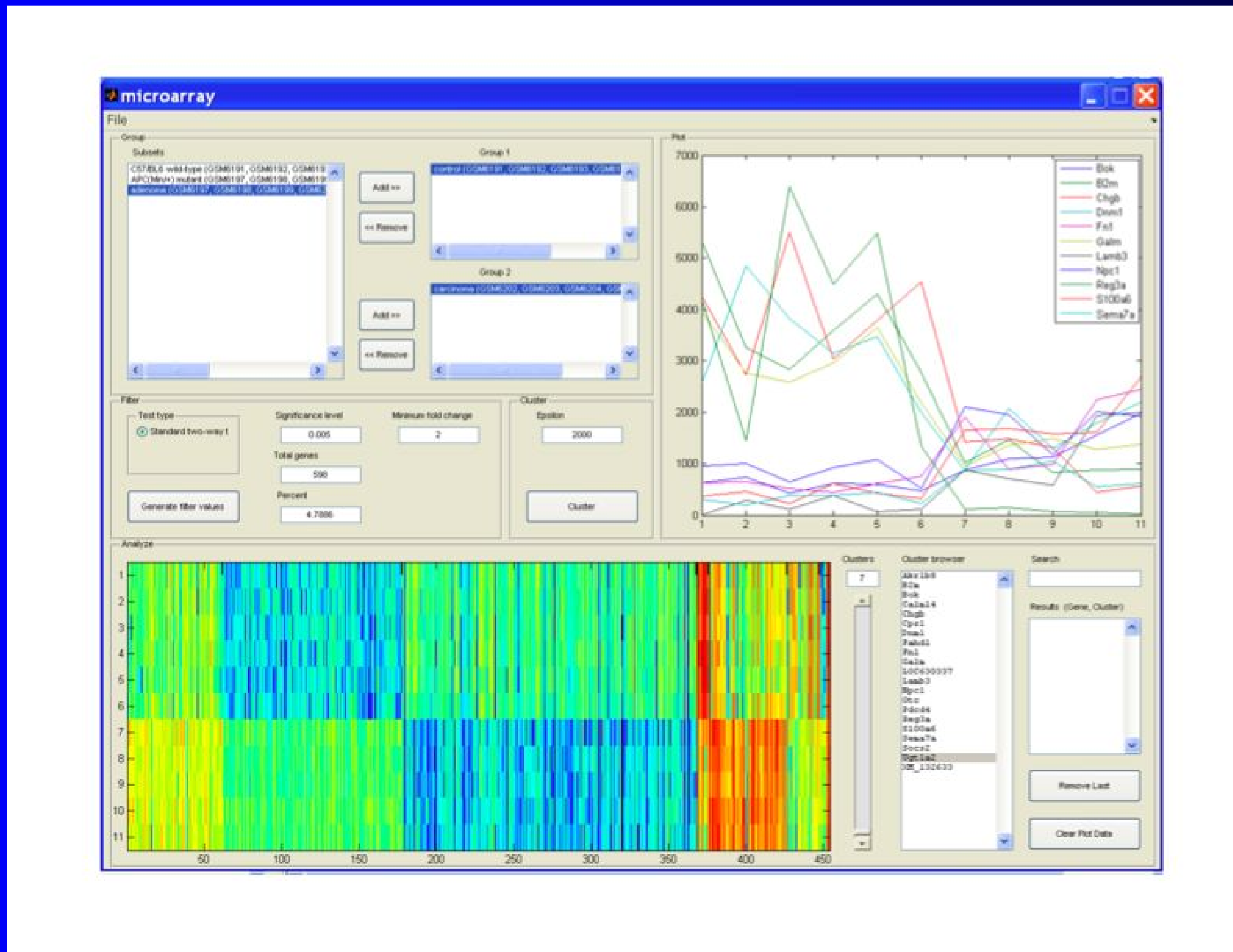


(a) (b) (c) (d)

Fig. 15. Results on two microarray datasets. (a) raw yeast data. (b) segmentation, visualized by reordering rows. The greedy algorithm discovers a number of distinct clusters of varying size. (c) raw leukemia data. (d) segmentation. Three clusters are found.

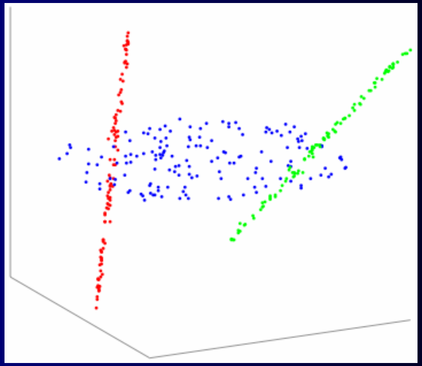
*Segmentation of Multivariate Mixed Data*, [Ma-Derksen-Hong-Wright, PAMI’

# Other Applications – Clustering of Microarray Data




*Segmentation of Multivariate Mixed Data*, [Ma-Derksen-Hong-Wright, PAMI’

# Other Applications – Supervised Classification



**Premises:** Data $\{y\}$ lie on an arrangement of subspaces

$$A = S_1 \cup S_2 \cup \cdots \cup S_n.$$

**Unsupervised Clustering**

– Generalized PCA

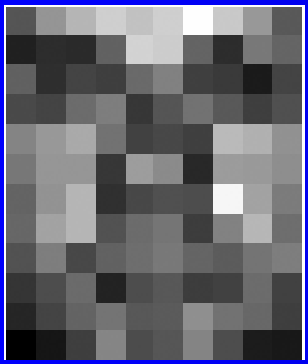
$$y = A_i \vec{\beta}_i, \; i = 1, \ldots, n.$$

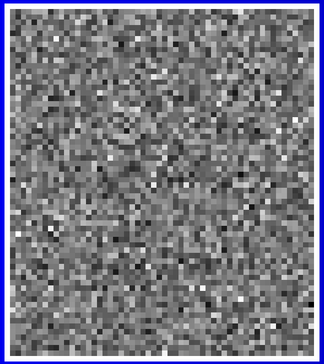
**Supervised Classification**

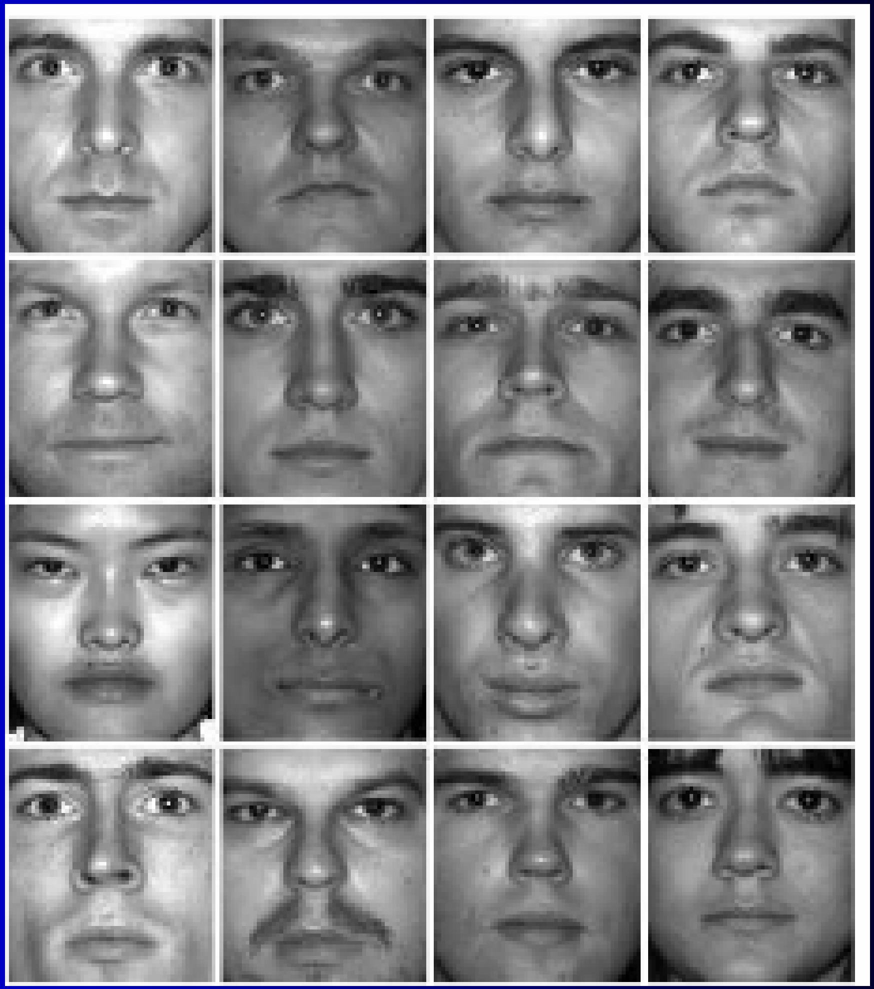
– Sparse Representation

$$y = [A_1, \ldots, A_n] x.$$

# Other Applications – Robust Face Recognition







*Robust Face Recognition via Sparse Representation, to appear in PAMI 2008*

# Other Applications: Robust Motion Segmentation (on Hopkins155)

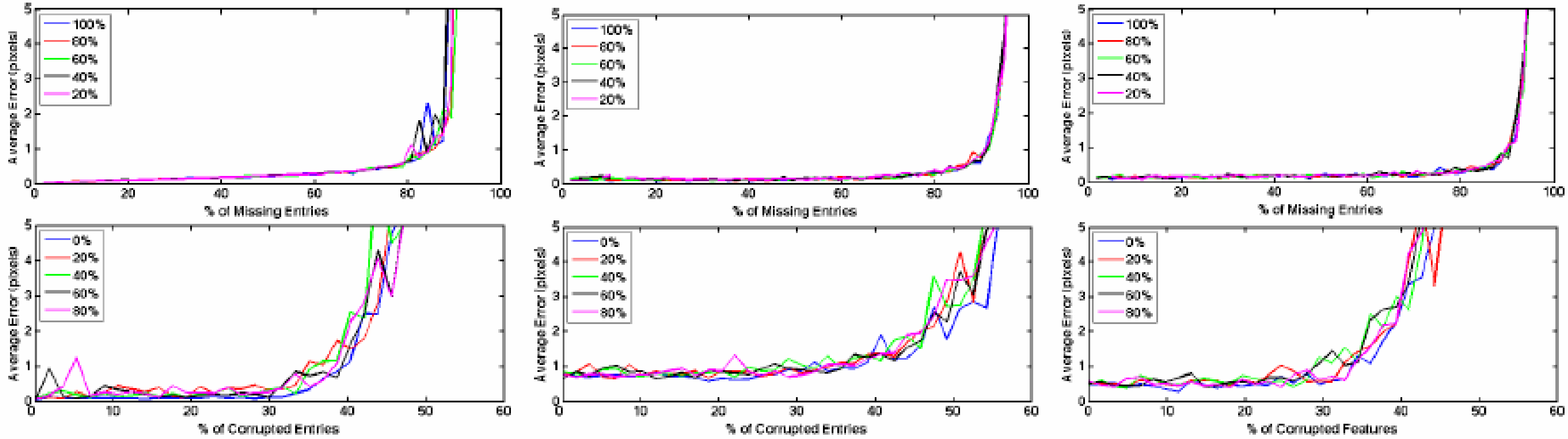
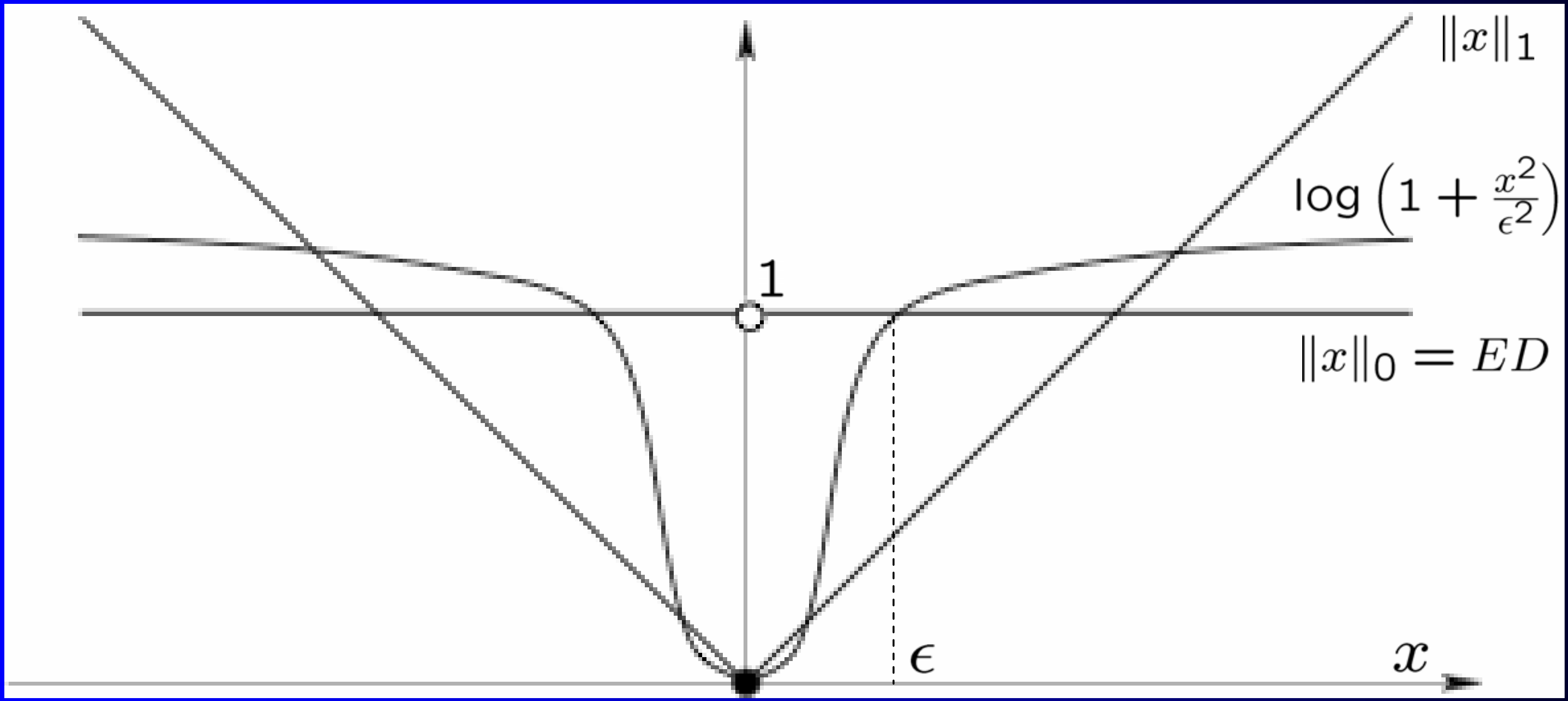



Figure 2. Errors of recovered trajectories for the sequences: "1R2RC" (left), "arm" (center), and "cars10" (right). Top: Results for our $\ell^1$-based trajectory completion. The different colored plots are for experiments with varying percentage of the dataset used for completion. Bottom: Results for our $\ell^1$-based detection and repair of corrupted trajectories. The different colors represent experiments with varying percentage of corrupted trajectories in the dataset.

Dealing with incomplete or mistracked features with dataset 80% corrupted!

Shankar Rao, Roberto Tron, Rene Vidal, and Yi Ma, to appear in CVPR'08

# Three Measures of Sparsity: Bits, L_0 and $L_1$-Norm




Reason: High-dimensional data, like images, do have compact, compressible, sparse structures, in terms of their geometry, statistics, and semantics.

# Conclusions

- Most imagery data are high-dimensional, statistically or geometrically heterogeneous, and have multi-scale structures.
- Imagery data require hybrid models that can adaptively represent different subsets of the data with different (sparse) linear models.
- Mathematically, it is possible to estimate and segment hybrid (linear) models non-iteratively. GPCA offers one such method.
- Hybrid models lead to new paradigms, new principles, and new applications for image representation, compression, and segmentation.

## Future Directions

- Mathematical Theory
  - Subspace arrangements (algebraic properties).
  - Extension of GPCA to more complex algebraic varieties (e.g., hybrid multilinear, high-order tensors).
  - Representation & approximation of vector-valued functions.
- Computation & Algorithm Development
  - Efficiency, noise sensitivity, outlier elimination.
  - Other ways to combine with wavelets and curvelets.
- Applications to Other Data
  - Medical imaging (ultra-sonic, MRI, diffusion tensor…)
  - Satellite hyper-spectral imaging.
  - Audio, video, faces, and digits.
  - Sensor networks (location, temperature, pressure, RFID…)
  - Bioinformatics (gene expression data…)

## Acknowledgement

- People
  - Wei Hong, Allen Yang, John Wright, University of Illinois
  - Rene Vidal of Biomedical Engineering Dept., Johns Hopkins University
  - Kun Huang of Biomedical & Informatics Science Dept., Ohio-State University
- Funding
  - Research Board, University of Illinois at Urbana-Champaign
  - National Science Foundation (NSF CAREER IIS-0347456)
  - Office of Naval Research (ONR YIP N000140510633)
  - National Science Foundation (NSF CRS-EHS0509151)
  - National Science Foundation (NSF CCF-TF0514955)

Generalized Principal Component Analysis:
Modeling and Segmentation of Multivariate Mixed Data

Rene Vidal, Yi Ma, and Shankar Sastry
Springer-Verlag, to appear

# Thank You!