Sparse Representation of Images with Hybrid Linear Model

Allen Yang Yang

March 15, 2004

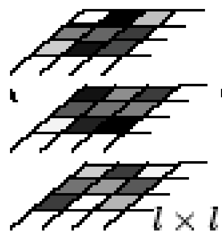
Paper version submitted to ICIP 2004 by

Kun Huang, Allen Y. Yang and Yi Ma

- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications
- Summary

- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications
- Summary

1. Introduction

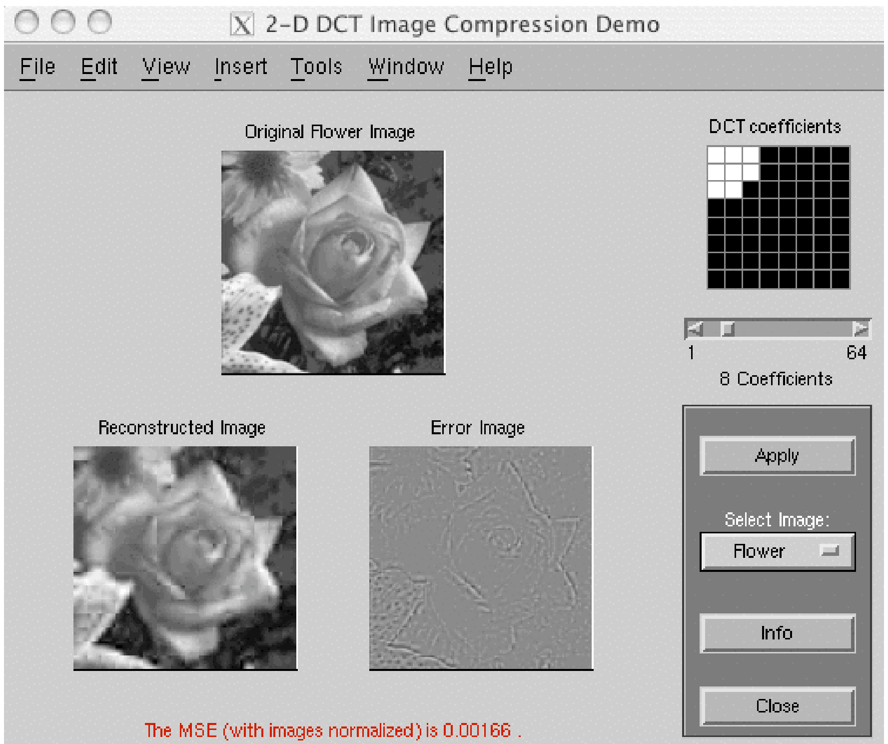


More effective representation?

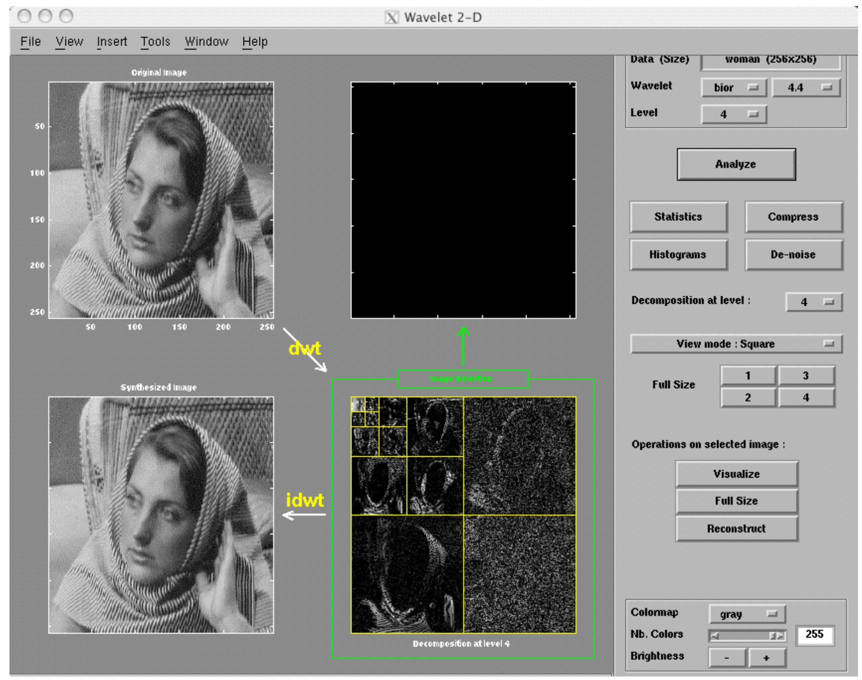
Pixel-based representation

DCT & Wavelet toolbox

dctdemo



wavemenu



Some well-known methods

- Fixed Basis: discrete Fourier transform (DFT) or discrete cosine transform (DCT) [Ahmed74].
- Fixed Basis (multi-resolution): wavelet transforms [Daucechies88, Mallat92].
- Adaptive Basis: Karhunen-Loeve transform (KLT), also known as PCA [Jain89]

Some well-known methods

- KLT (PCA) is known to be the *optimal* transform if signals are assumed to be drawn from one stationary 2nd-order stochastic model [Effros95].
- Some industry standards:

  1. JPEG uses DCT transform.

  2. JPEG-2000 uses multi-resolution bior4.4 wavelet.

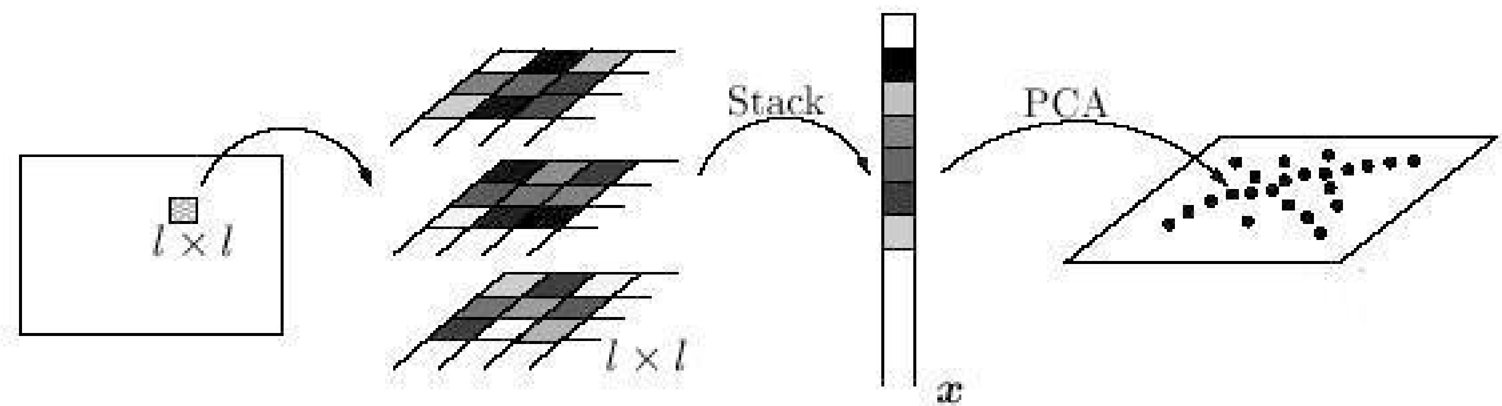
Single linear model assumption

- All the above methods achieve sparse representation by transforming the image with respect to:

  1. single linear models, or

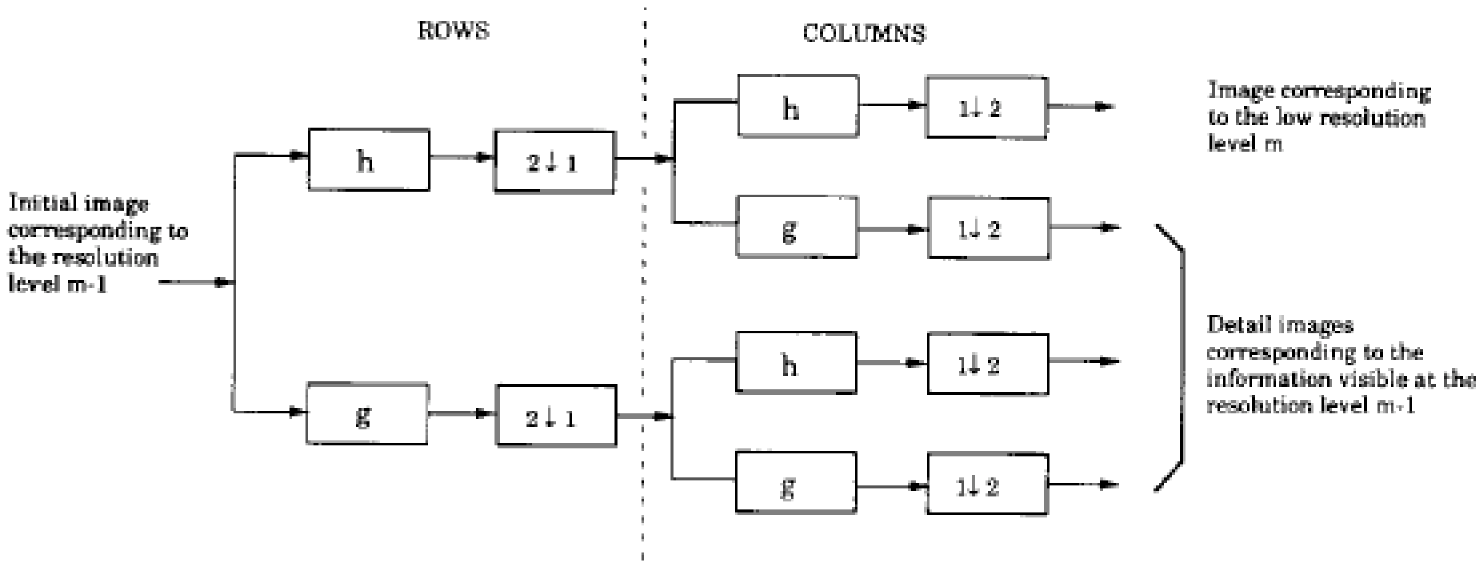
  2. fixed filter bank models.

Single linear model assumption



An illustration of KL transform

Fixed filter banks

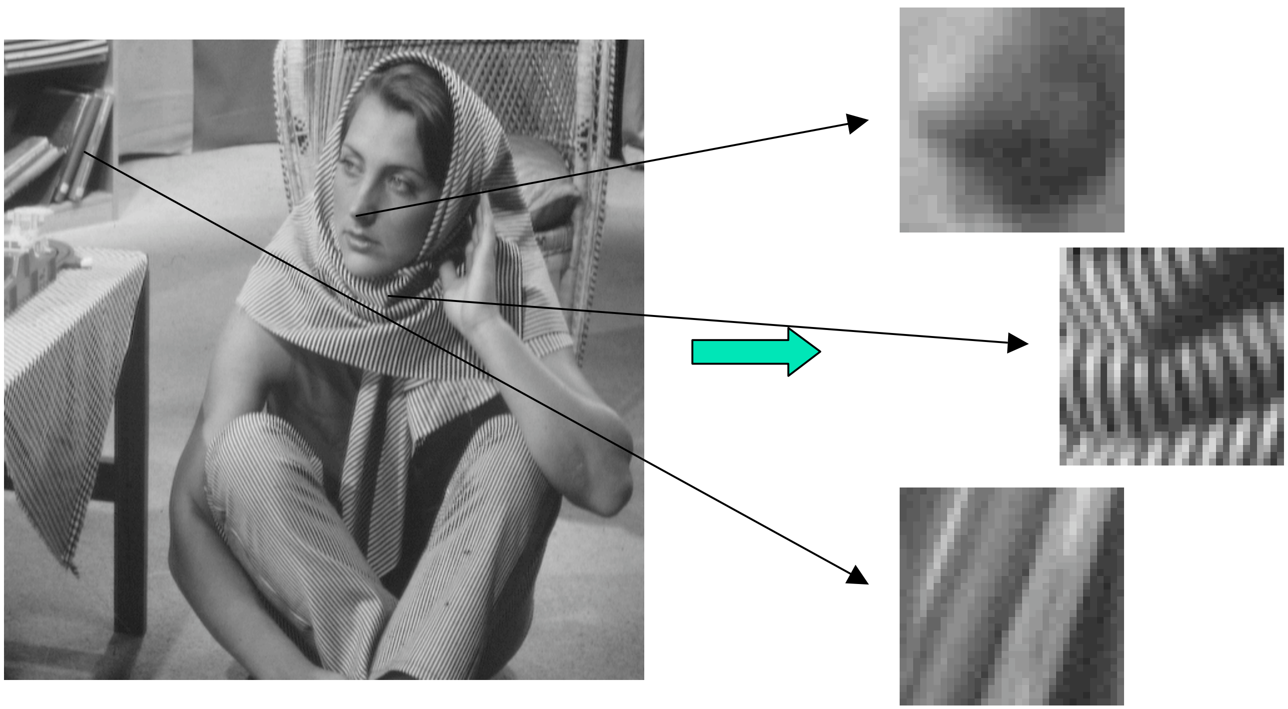


An illustration of wavelet filter bank.

Marc Antonini, et al., *Image coding using wavelet transform*. IEEE trans. on image processing, April 1992, pp 205-220

Hybrid linear model

However, distribution of the signals from an image is NOT a single statistical model in general!

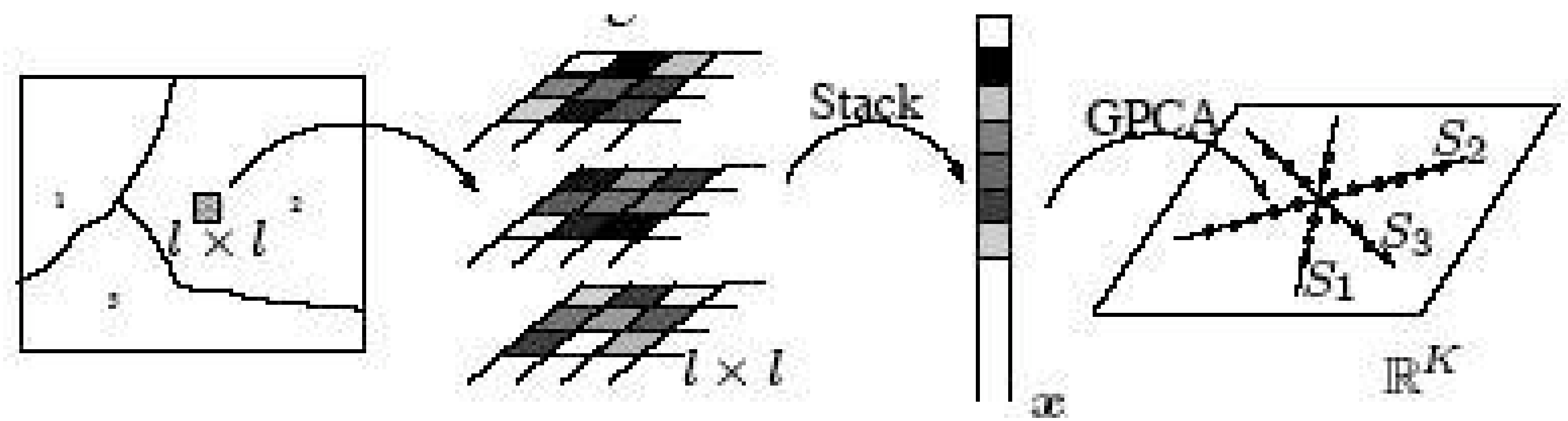


Hybrid linear model

Goals:

1. Segment an image into multiple regions such that within each region the signals obey one linear model.

2. Represent the image based on the multiple linear models we obtain from the segments.

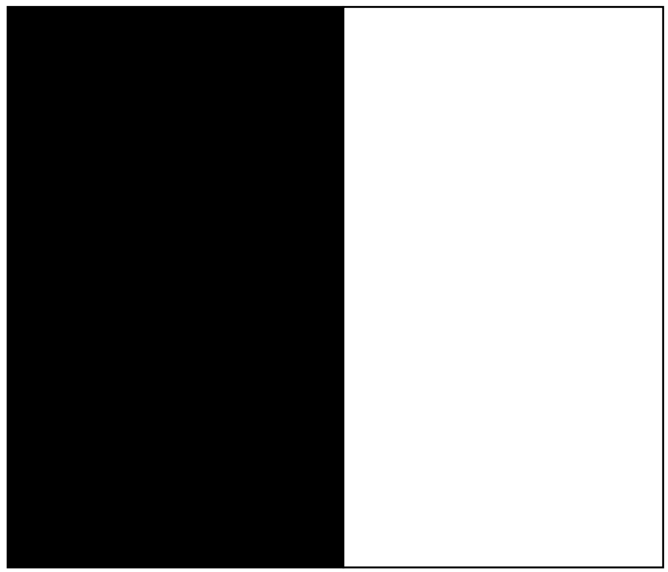


Hybrid linear model

The major difficulties

- Multiple models: without first knowing the basis of each linear model, how to segment the image into different regions?
- Estimation: without knowing the segmentation, how to estimate an adaptive basis for each unknown region.

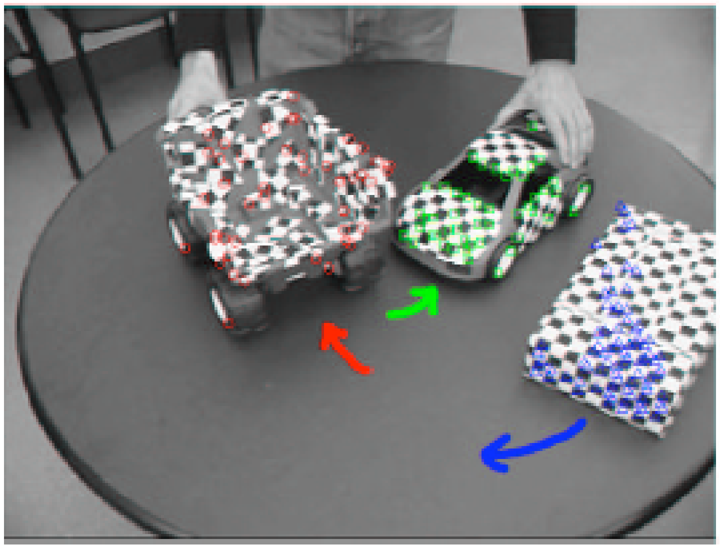
One example

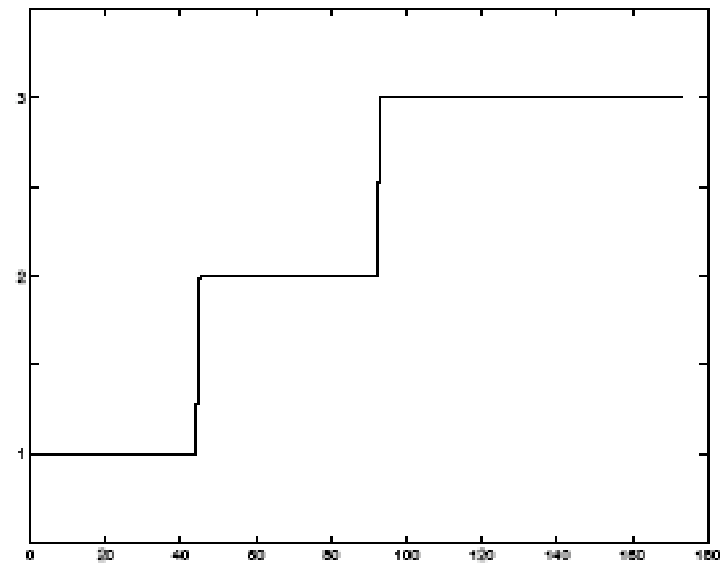


Segmentation based on linear models is different from traditional texture segmentation.

- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications
- Summary

The origin of GPCA (PFA)





Segmentation and estimation of multi-body Fundamental matrix

R. Vidal, S. Soatto, Y. Ma and S Sastry, ECCV workshop on Vision and Modeling of Dynamic Scenes, May 2002

What is in common?

- Data of both problems are subject to a mixture of multiple linear models.
- The segmentation of the data samples and the estimation of the linear models are strongly coupled.
- Both problems were treated as chicken-and-egg problems, and no non-iterative closed-form solution was known. Statistical iterative methods were commonly used (EM, K-mean).

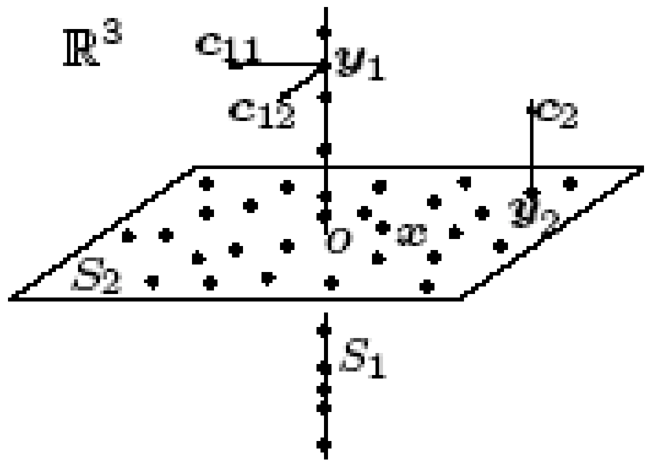
Problem statement

Condition: Given a set of data points sampled from an unknown number of different subspaces in a high dimensional space:

1. Estimate the number of subspaces and their dimensions;
2. Generate a basis for each subspace;
3. Group the given data points into the subspaces.

A case study

Suppose we have sample data in $\mathbb{R}^3$,
drawn from two linear subspaces:
a line $S_1 = \{\mathbf{x} : x_1 = x_2 = 0\}$ and a plane $S_2 = \{\mathbf{x} : x_3 = 0\}$.



De Morgan's rule

The entire data set can be represented as

$$S_1 \bigcup S_2 = \{\mathbf{x} : (x_1 = x_2 = 0) \quad \text{or} \quad (x_3 = 0)\}$$

Using De Morgan's rule, it is equivalent to

$$\{\mathbf{x} : (x_1 \cdot x_3 = 0) \quad \text{and} \quad (x_2 \cdot x_3 = 0)\}$$

Question: how to obtain this subspace description from the data points?

Veronese embedding

We first assume $n = 2$ and $K = 3$. The *Veronese* map is defined as:

$$\nu_n : \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_1 x_3 \\ x_2^2 \\ x_2 x_3 \\ x_3^2 \end{bmatrix} \in \mathbb{R}^{M_n}$$

Given the number of the subspaces $n = 2$ and the dimension of the ambient space $K = 3$, the total number of the monomials is $M_n = \binom{n + K - 1}{n} = 6$.

Veronese embedding

Define the data matrix for N points and $n = 2$ as the following:

$$L_2 \doteq \begin{bmatrix} (x_1^1)^2 & x_1^1 x_2^1 & \cdots & (x_3^1)^2 \\ (x_1^2)^2 & x_1^2 x_2^2 & \cdots & (x_3^2)^2 \\ & \cdots & \cdots & \\ (x_1^N)^2 & x_1^N x_2^N & \cdots & (x_3^N)^2 \end{bmatrix}.$$

This matrix has exact 2-dimensional null space:

$$\mathbf{c}_n = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The basis vectors of the null space of L_n correspond to $p_1 = x_1 x_3$ and $p_2 = x_2 x_3$!

Polynomial differentiation

The mixture of the subspaces in this case is described by the polynomials

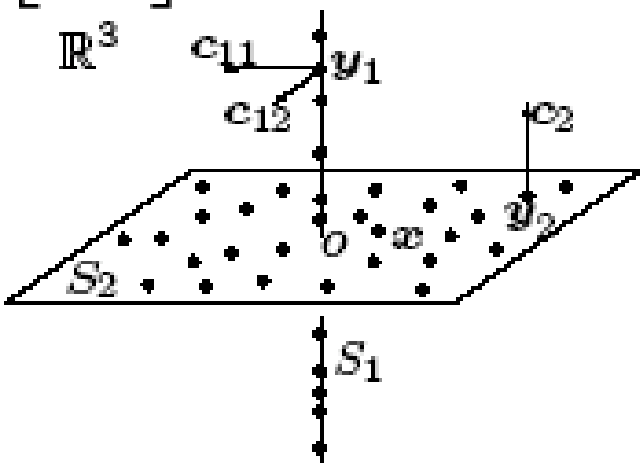
$$P(x) = [p_1(x), p_2(x)] = [x_1x_3, x_2x_3]$$

$$DP(\mathbf{x}) = \begin{bmatrix} \frac{\partial p_1}{\partial x_1} & \frac{\partial p_2}{\partial x_1} \\ \frac{\partial p_1}{\partial x_2} & \frac{\partial p_2}{\partial x_2} \\ \frac{\partial p_1}{\partial x_3} & \frac{\partial p_2}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_3 & 0 \\ 0 & x_3 \\ x_1 & x_2 \end{bmatrix}.$$

Polynomial differentiation

If we know one point in each subspace, for example,
$y_1 = [0,0,1]^T \in S_1$ and $y_2 = [1,1,0]^T \in S_2$,

Then $DP(\mathbf{y}_1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $DP(\mathbf{y}_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$.



The columns of $DP(\mathbf{y}_1)$ span $S_1^\perp$,
and the columns of $DP(\mathbf{y}_2)$ span $S_2^\perp$!!

One last unknown

How do you know in advance that the number of subspaces is 2?

Answer: Search incrementally until Ln drops rank.

Summary of PDA

- Given N points in K-dim.
- Compute the data matrix Ln for $n \geq 2$ until Ln drops rank.
- Compute the null space of Ln, and obtain all the polynomials.
- Find one positive sample on each subspace, evaluate the derivative at this point and find a basis for the subspace.
- Remove all the samples in that subspace from the data set, repeat.

Some practical issues

- How to find one point per subspace?

  There exists a heuristic method, which finds one point that has the shortest distance to each of the subspaces.

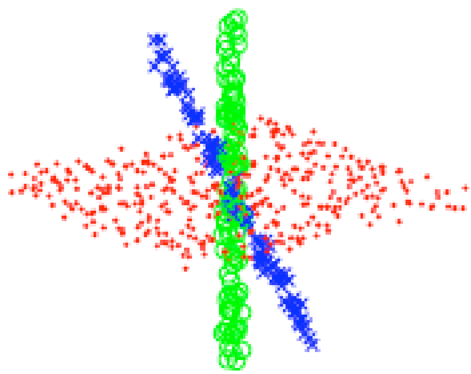
- How to remove a subspaces found from the data set robustly?

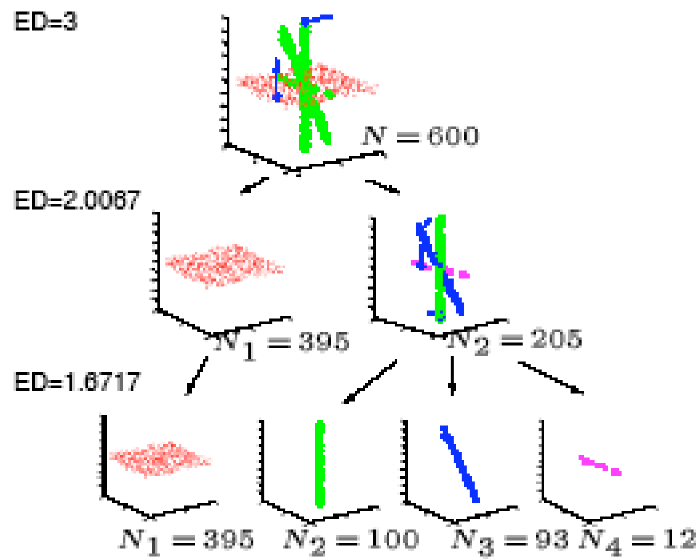
  Instead of removing sample points, heuristically, we can do polynomial division, which is purely algebraic.

A reference

Rene Vidal and Yi Ma, *Clustering subspaces by fitting, differentiating and dividing polynomials*, to be appear in CVPR 2004.

A demo result





- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications
- Summary

Hybrid linear model in images

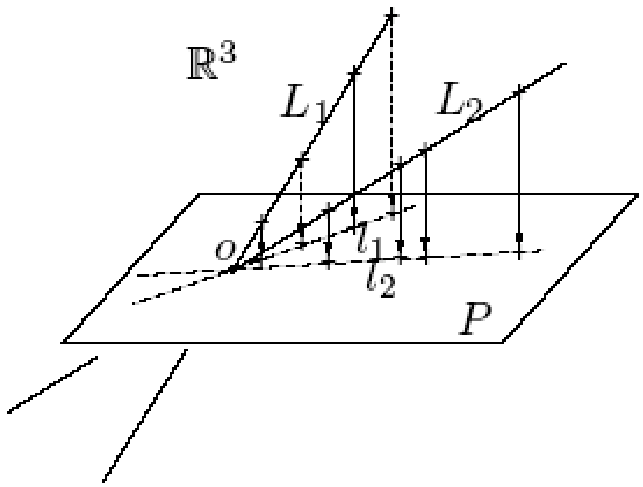
- In principle, one can use the same idea to identify a hybrid linear model directly from an image.

The dimension K for an 8 by 8 by 3 block is 192!

The dimension of the Veronese embedding will be

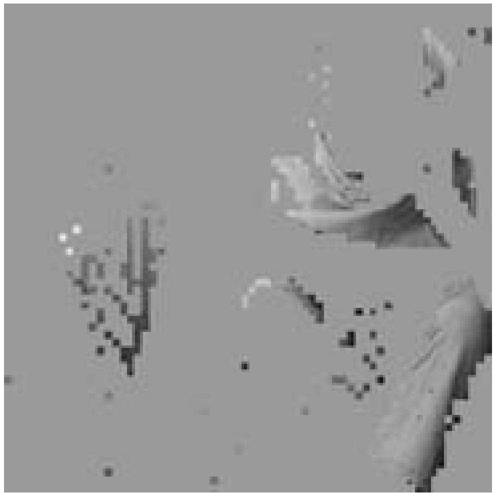
$$M_n = \binom{n+K-1}{K-1}$$

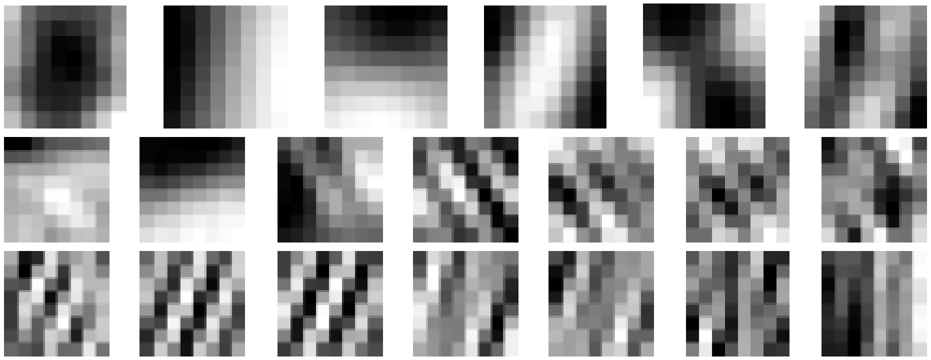
Dimension reduction



If the subspaces are k dimension, the projection $\pi_P : \mathbb{R}^K \rightarrow \mathbb{R}^{k+1}$ in general will preserve the dimension of the subspace and the membership of each sample. There is a one-to-one correspondence between and original subspaces and the projected ones.

GPCA result on Barbara





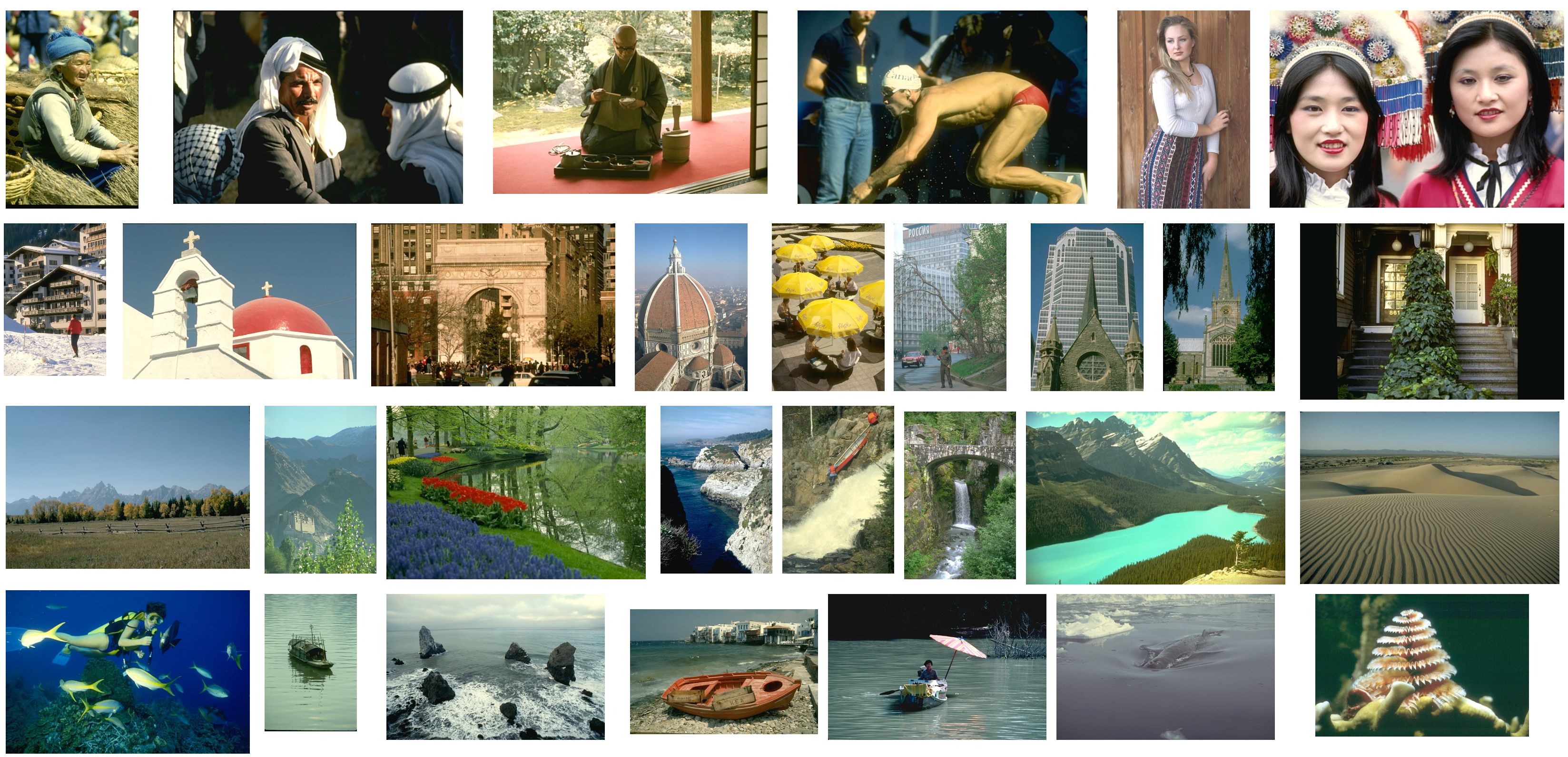
- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications

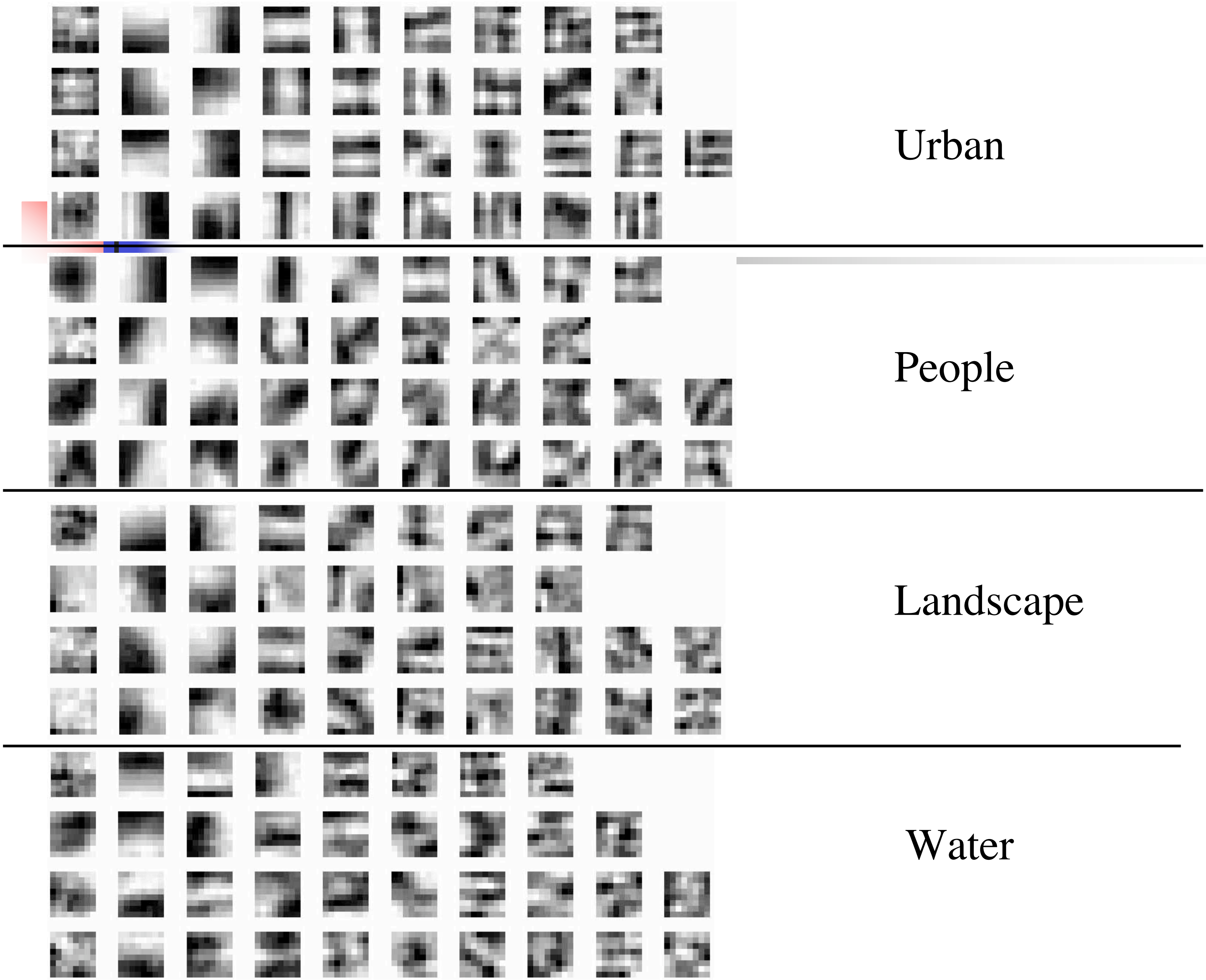
  1. Sparse codes for image ensembles

  2. Lossy image representation
- Summary

1. Sparse coding for image ensembles

- What will the sparse basis be for different categories of images?

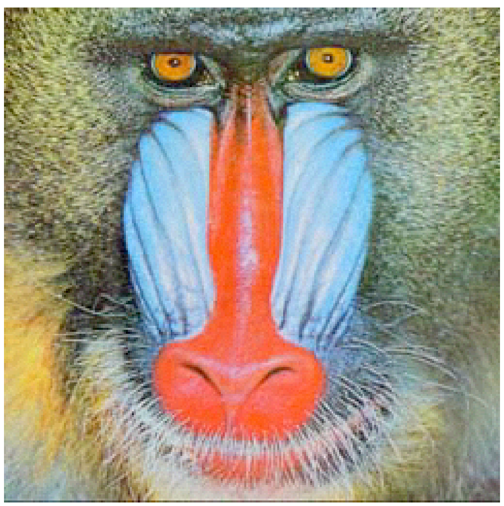


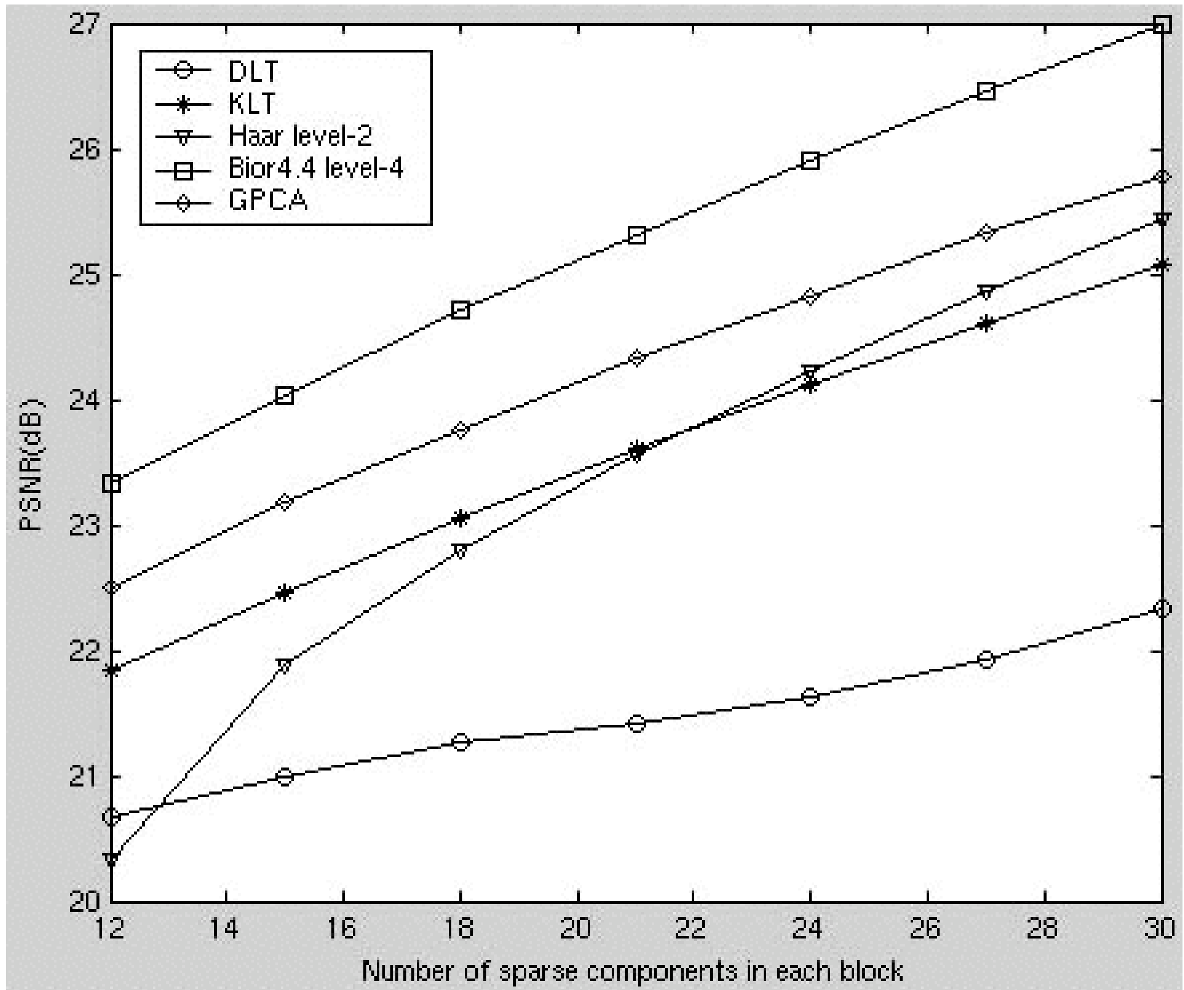

Urban
People
Landscape
Water

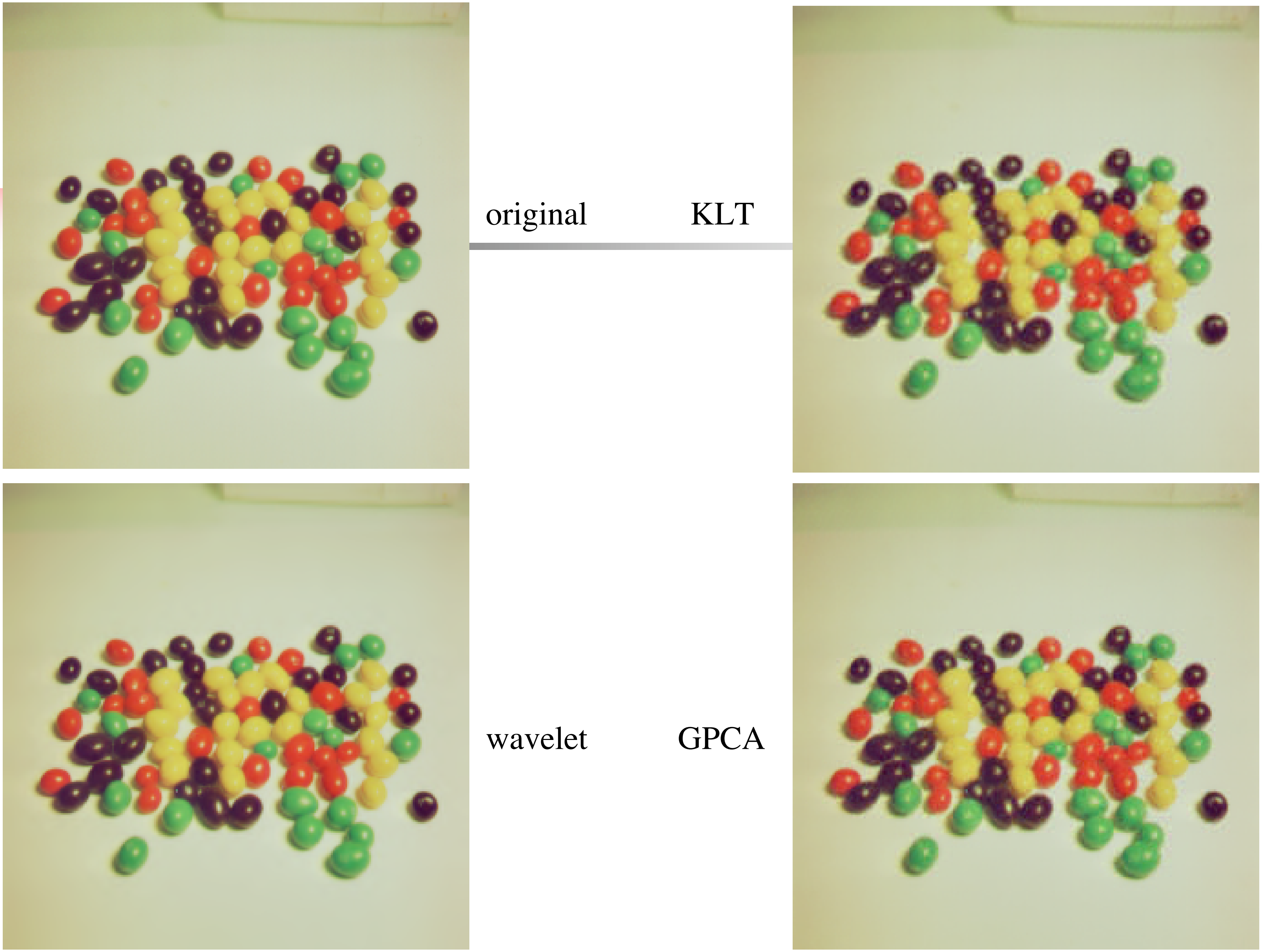
2. Lossy image representation

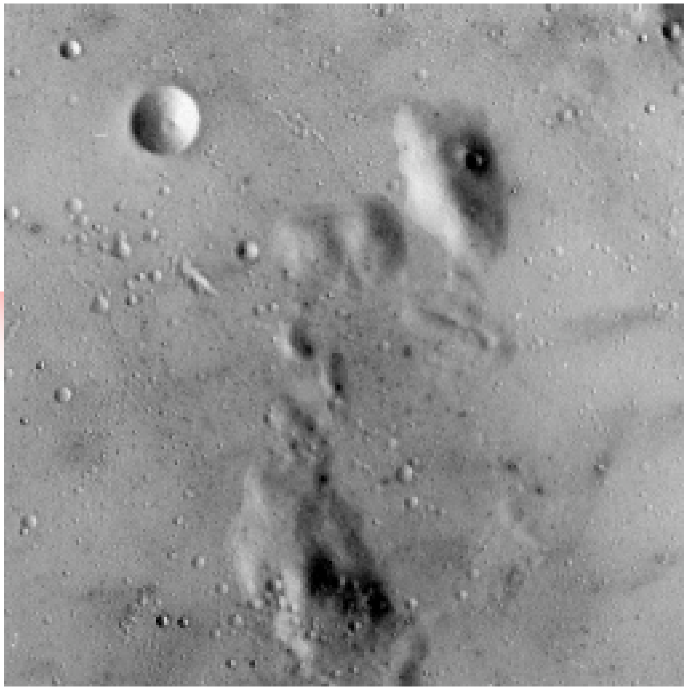
- The hybrid linear model offers an efficient representation of an image.
- After each subspace is identified, we may compress each subspace based on its unique linear model.

Compression result



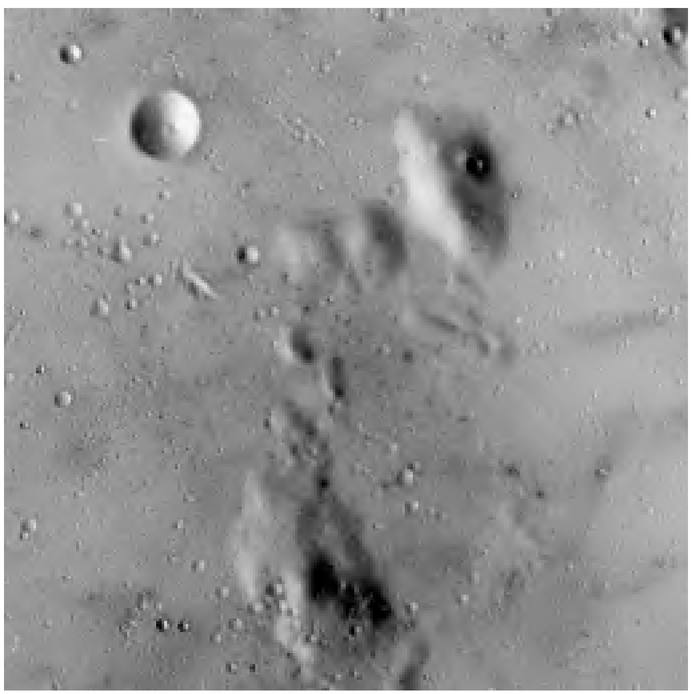



original
KLT
wavelet
GPCA



original

KLT



wavelet

GPCA

original

KLT

wavelet

GPCA

Table 1: Compression quality comparison using the PSNR (dB)*.

Test images	KLT	level-3 Bior4.4	GPCA
Lena grayscale: 512×512	31.5578	38.2735	32.5745
Lena with numbers embedded grayscale: 512×512	20.6257	27.6285	22.7146
Man with a camera grayscale: 256×256	25.0862	33.9286	26.3532
Lake with a sailboat color: 512×512	28.9998	31.1965	29.4768
F-16 jet color: 512×512	34.5049	37.62492	35.6760
Pine with mountains color: 256×256	28.6259	29.9005	29.4511

Conclusion

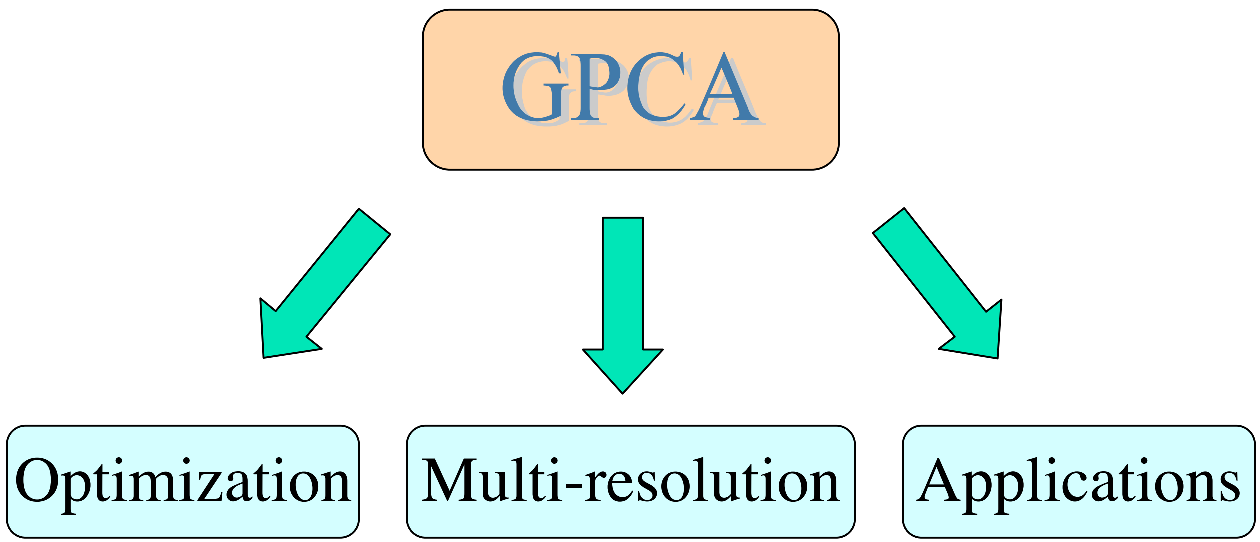
- If most part of an image is covered by the same type of objects with similar statistics, GPCA is only marginally better than KLT.
- When a image contains multiple different regions, such as foreground and background, change of objects, GPCA matches closely to the bior4.4 transform.
- Wavelet performs very good at high frequency signals, GPCA has the advantage of estimating linear models from the lumped pixel vectors.

- Introduction
- Identifying hybrid linear models via GPCA
- Hybrid linear model for images
- Applications
- Summary

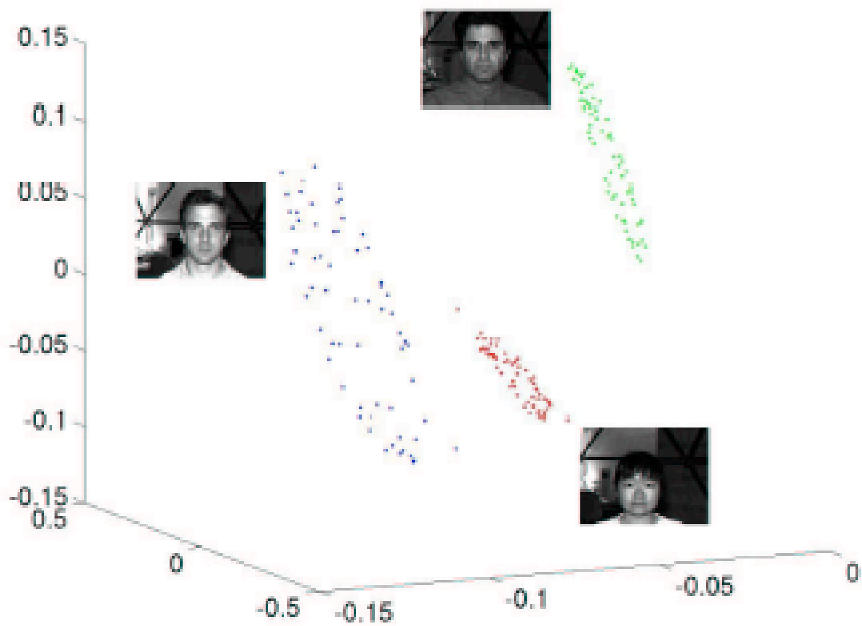
Summary

- Hybrid linear model: effective representation of images.
- The main advantage of GPCA: ability to solve the segmentation and estimation simultaneously.
- The solution is noniterative and global.
- Shown applications in
  - Image sparse coding
  - Lossy image representation and compression

Future research



Face recognition



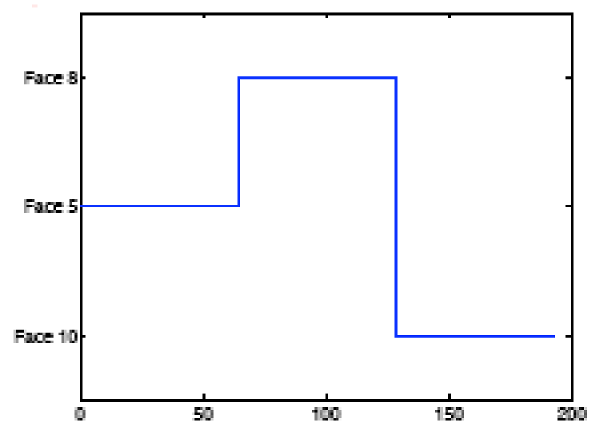


Figure 4: Clustering a subset of the Yale Face Database B consisting of 64 frontal views under varying lighting conditions for subjects 2, 5 and 8. Left: Image data projected onto the three principal components. Right: Clustering of the images given by PDA.

Video segmentation

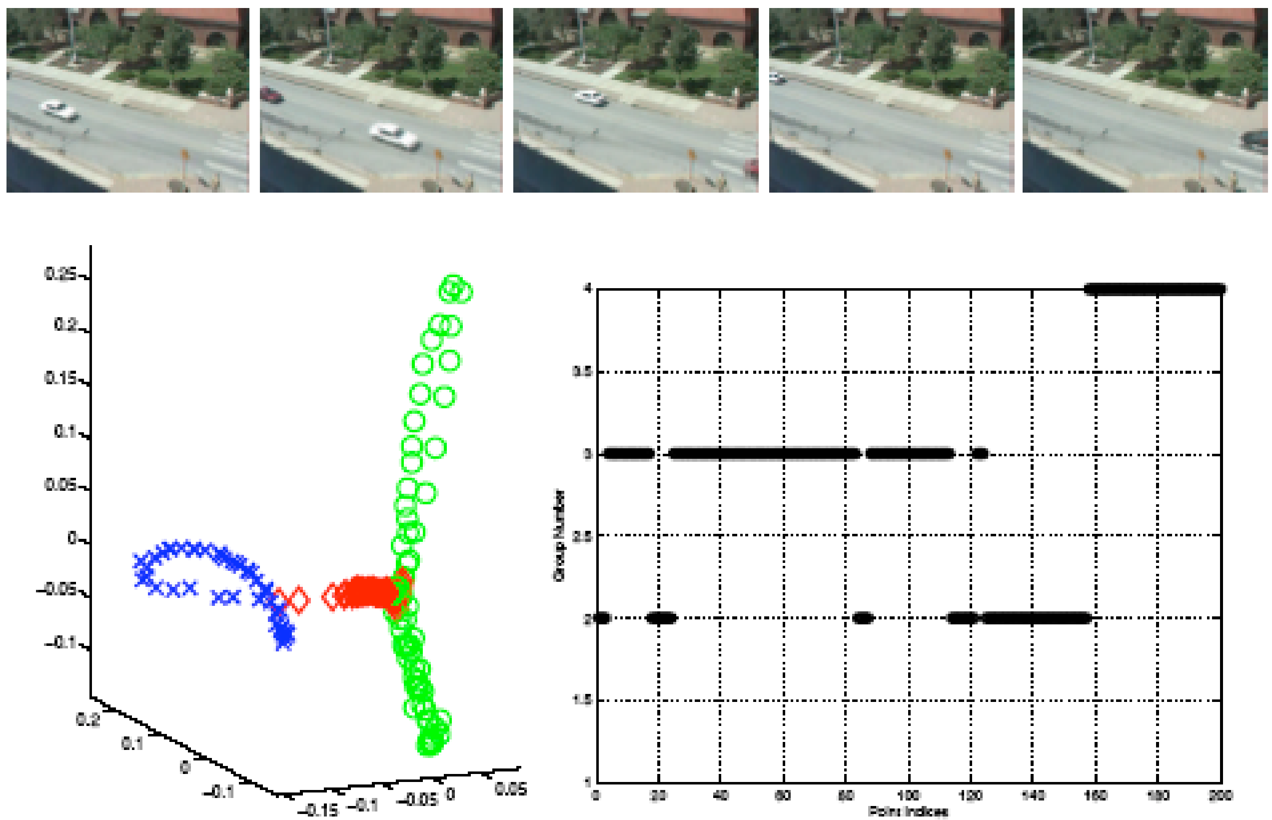


Figure 9: Results for segmenting a video sequence of 200 frames, each color is for a different segment.

Sparse Representation of Images with Hybrid Linear Model

Thank You!

yangyang@uiuc.edu