Image Processing and Visual Communications

Image Transforms

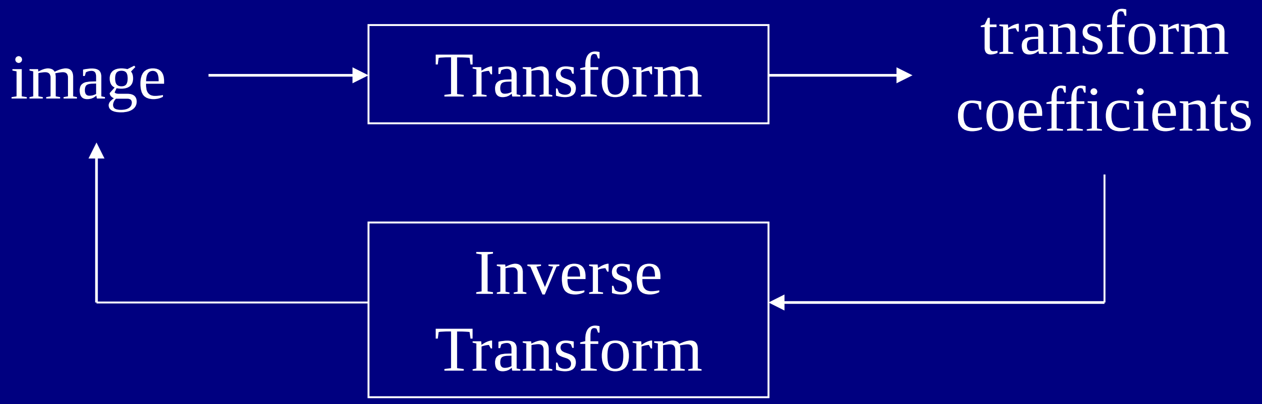
Zhou Wang

Dept. of Electrical and Computer Engineering
University of Waterloo

Outline

- **Why Image Transforms?**
- **Linear Transform and Properties**
  - Linearity
  - Orthogonality
  - Energy Preserving
  - Energy Compaction
  - Sparsity
  - Independence
  - Other Properties
- **Quasi-Linear Transforms**
  - Divisive Normalization
  - Adaptive Basis Selection
  - Adaptive Basis Generation

Why Image Transforms?



- **What Does a Transform Do?**
  - Represents the same image signal in a different but “better” way
  - Decomposes the signal into “easier” to understand components

- **What Do We Gain?**
  - Better mathematical properties
  - Better fit to the nature of the source (images)
  - Better fit to the nature of the receiver (biological visual systems)

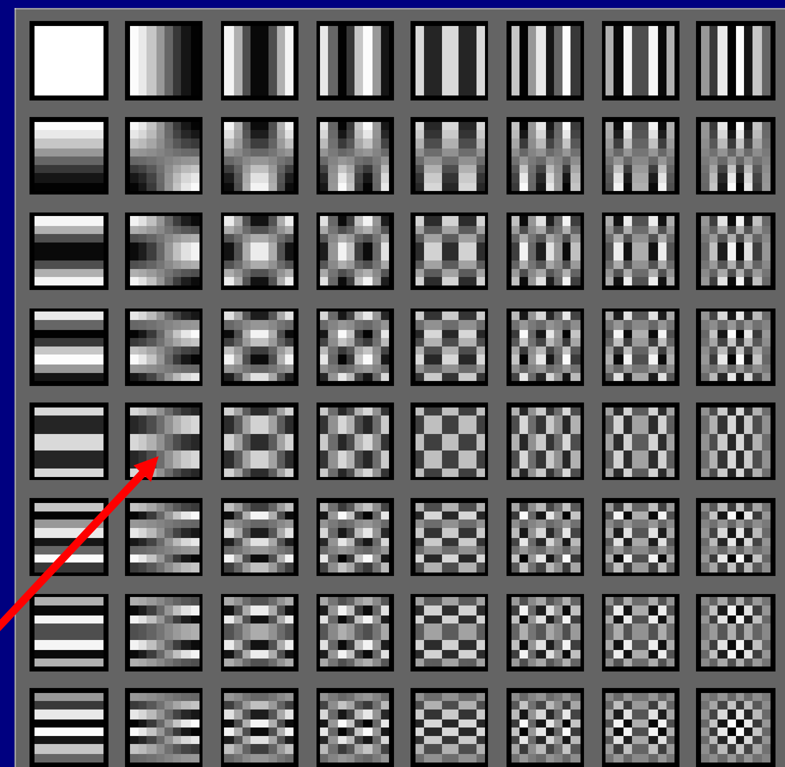
Linear Transform

$$x \longrightarrow \boxed{\mathbf{L}} \longrightarrow \boldsymbol{c} = \mathbf{L}\,\boldsymbol{x}$$

$$\boldsymbol{x} = \mathbf{B}\,\boldsymbol{c} = [\,\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \ \boldsymbol{b}_3 \ldots \boldsymbol{b}_N]\,\boldsymbol{c}$$

$$= c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \ldots + c_N\boldsymbol{b}_N$$

- **Concepts**
  - Linearity
  - Basis
  - Coefficient
  - Complete, incomplete, and overcomplete bases
- **Superposition Principle**
  - Linear combination of bases



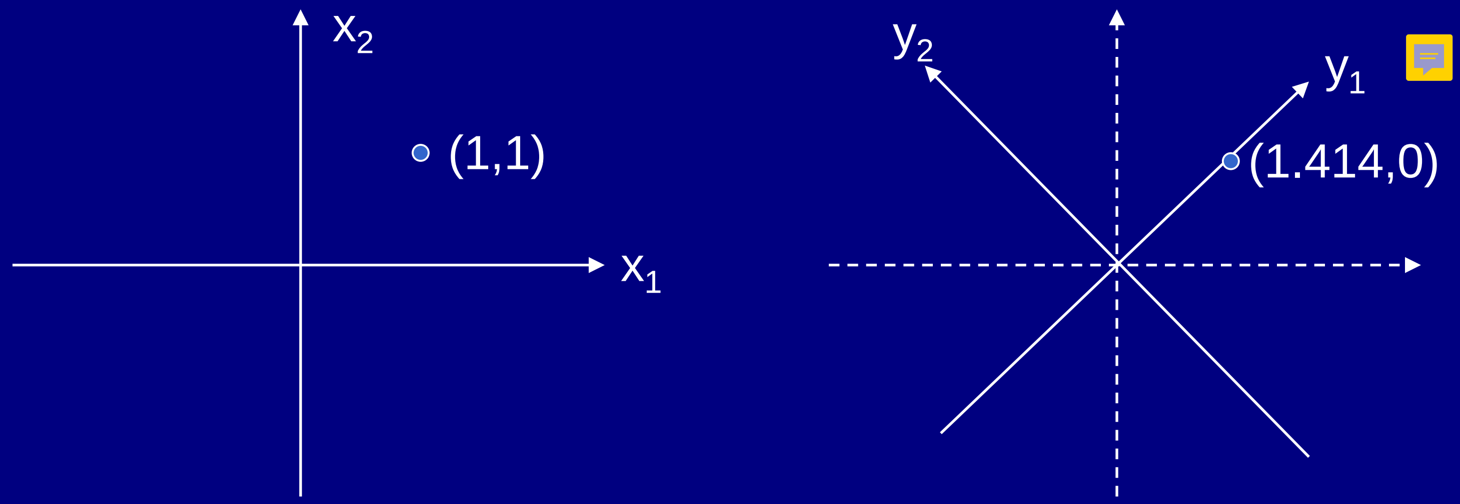
DCT basis

Example: any 8x8 blocks can be represented as a linear combination of DCT bases

Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Linearity

- **Linear Transform → change of the coordinate system**

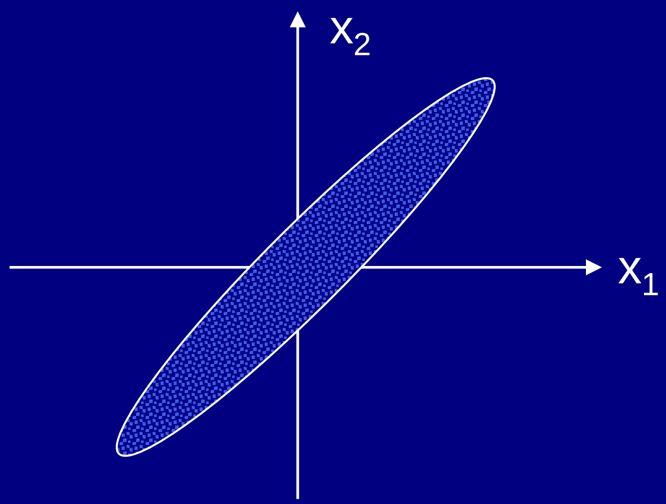


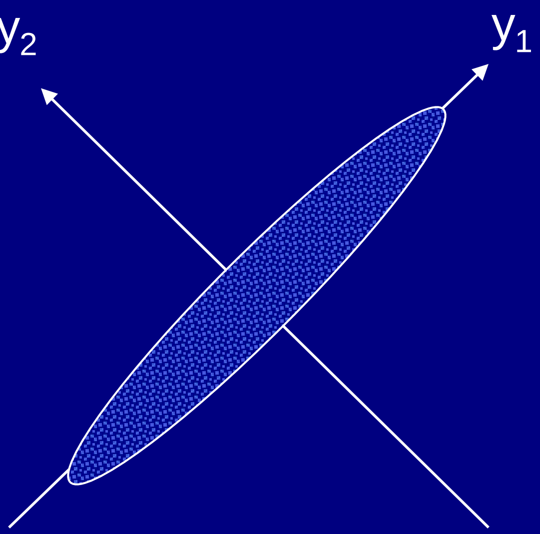
- **Why Do We Love Linear Transforms/Systems?**
  - Nice math properties, large number of math tools
  - First (and useful) approximation of complex systems. Often used as a benchmark for advanced systems

Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Orthogonality

- **Orthogonal Transform: $A^{-1}=A^T$**





- **Unitary Transform: $A^{-1}=A^{*T}$**
- **Both are Nothing But Rotations of Coordinate Systems**
  - No compression/stretching
  - Preserve vector length → energy preserving

Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Energy Preserving

- **A Transform y = Ax is Energy Preserving if for Any x**

$$||\mathbf{y}||^2 = ||\mathbf{x}||^2$$

- **Orthogonal Transforms:**

$$||\mathbf{y}||^2 = \mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{x} = ||\mathbf{x}||^2$$

- **Unitary Transforms:**

$$||\mathbf{y}||^2 = \mathbf{y}^{*T}\mathbf{y} = \mathbf{x}^{*T}\mathbf{A}^{*T}\mathbf{A}\mathbf{x} = \mathbf{x}^{*T}\mathbf{x} = ||\mathbf{x}||^2$$

- **Overcomplete Representations:**
  - Not orthogonal or unitary (**A** is not square), but can also be energy preserving (tight frame)

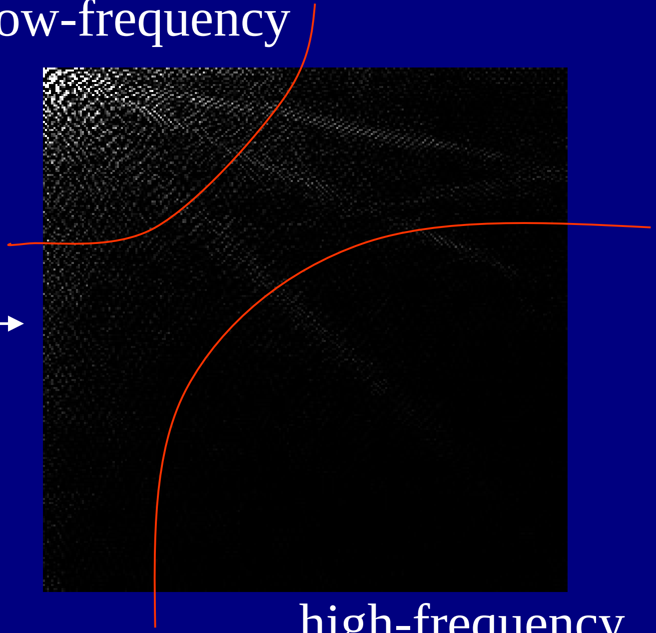
Energy Compaction

- **Energy Measure**

original image

DCT

low-frequency



high-frequency

  – Only 2451 out of 65536 DCT coefficients are **significant** (*Th* = 64)

  – Reasons:

    Natural images: energy concentrates at low frequencies

    DCT: decomposes the image according to frequencies

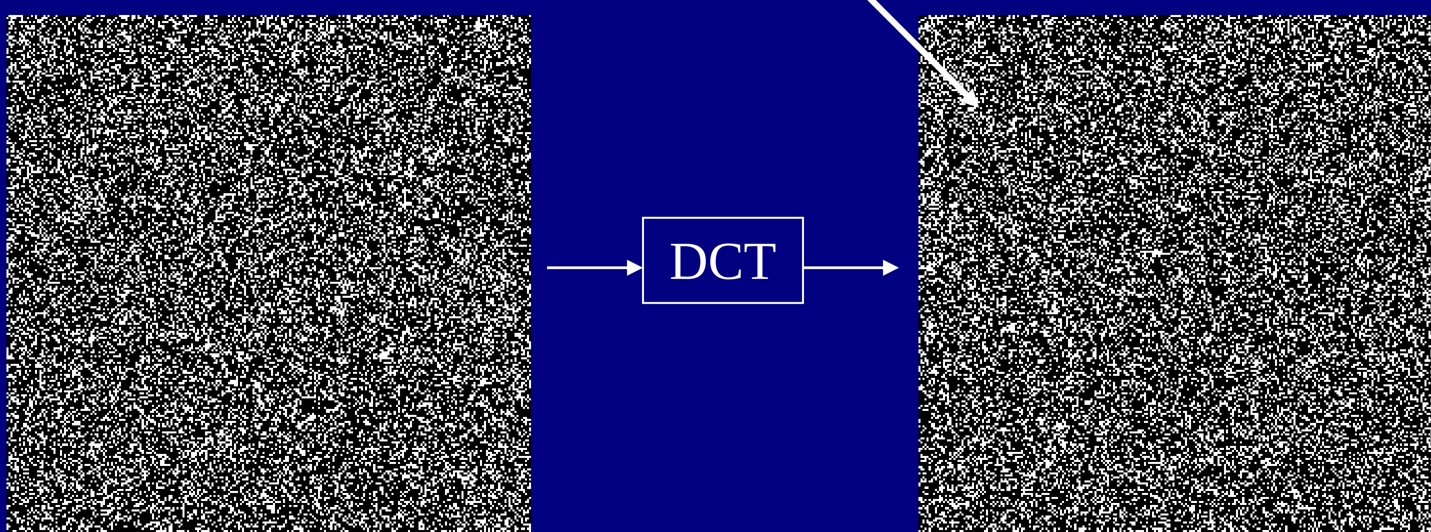
Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Energy Compaction

- **Energy Compaction**
  - A small number of bases “absorb” a large percentage of energy
  - Bases (filters) designed to fit highly probable image structures
  - Depends on the nature of images (not purely a math property)

A counterexample:

NO energy compaction

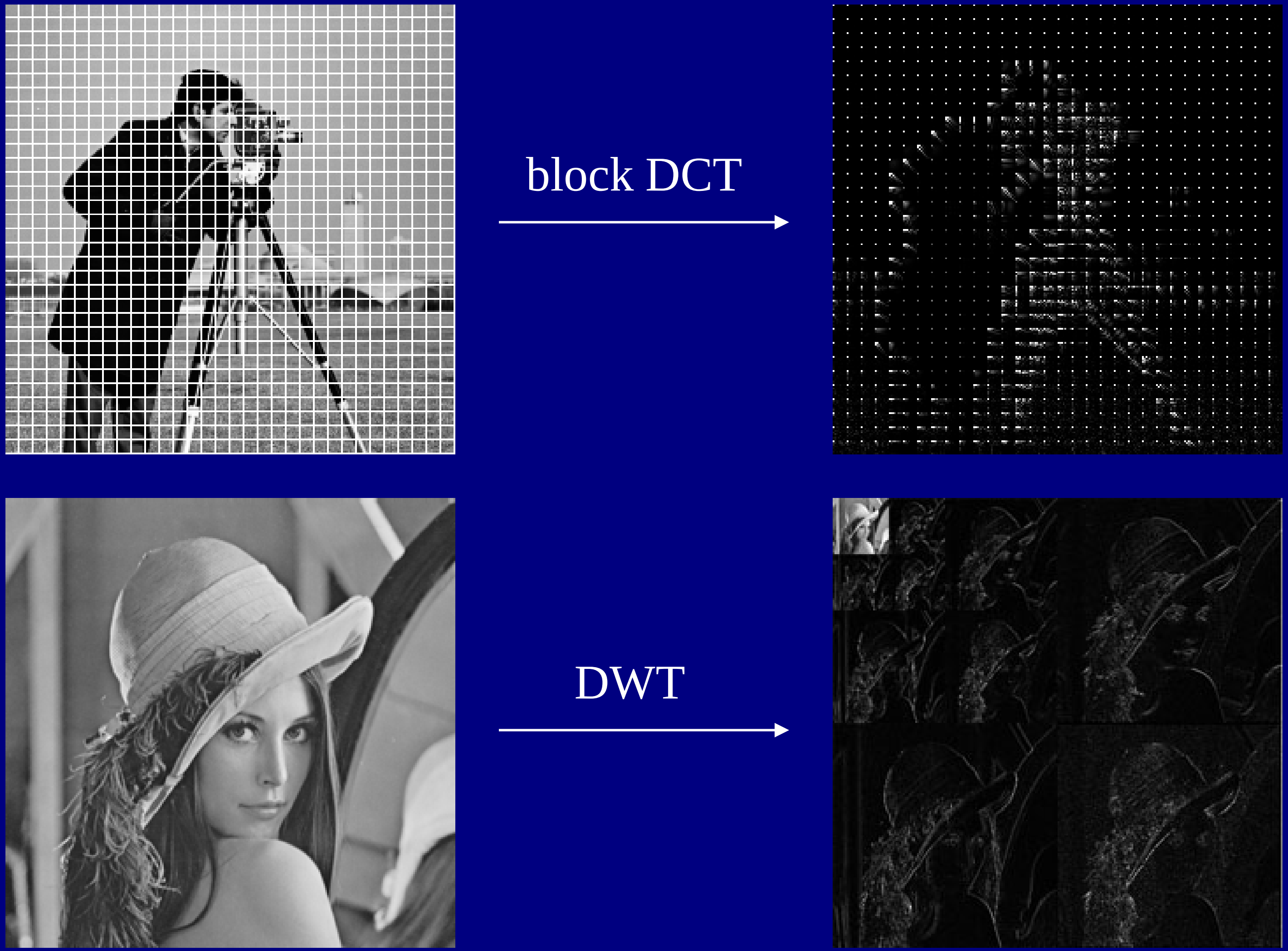


original “image” DCT coefficients

Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Sparsity

- **Sparse Representations**



Ref: *Introduction to Digital Image Processing* lecture notes by Prof. Xin Li

Sparsity

- **Optimal Sparse Representation**
  - Bases trained from a large collection of natural images
  - What do we see?
    wavelets!
    edge detection filters!
    neural responses!

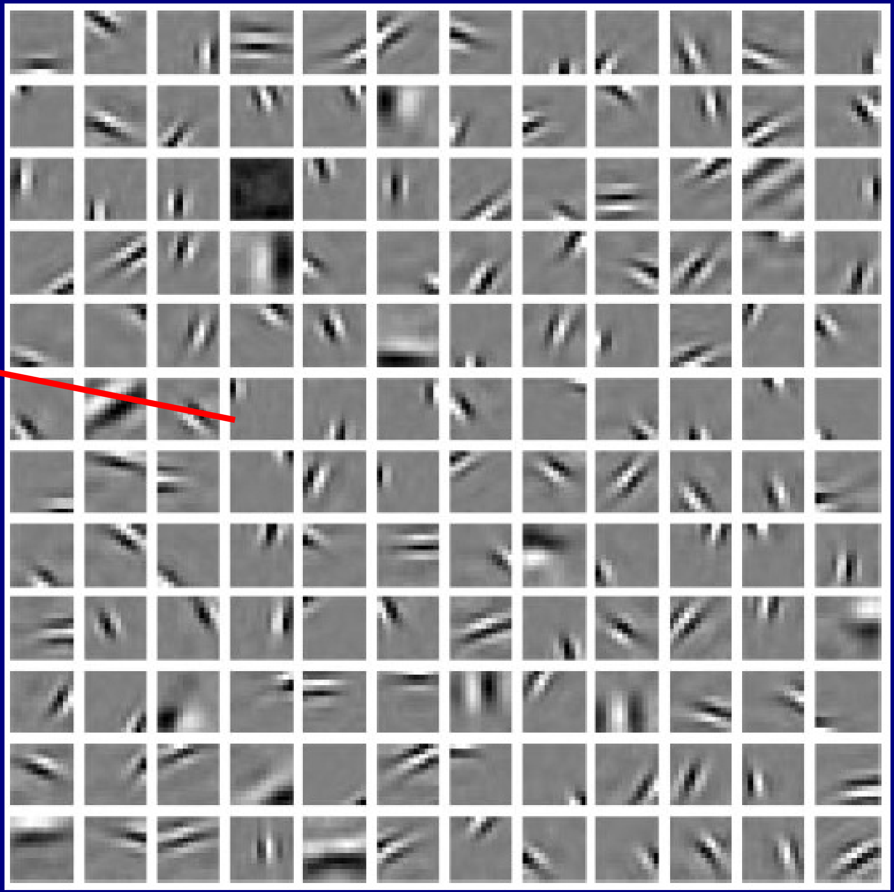


image **linear** bases optimized for sparseness

Olshausen & Field '96

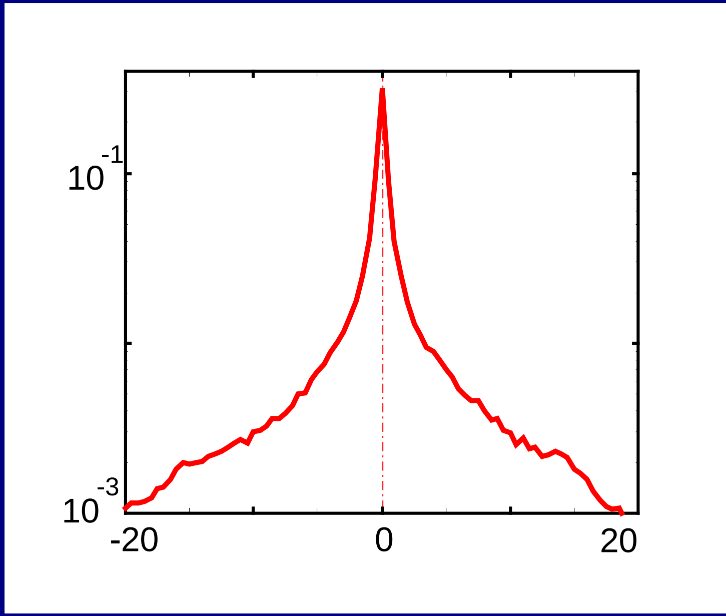
Sparsity

- **Measurement of Sparseness**
  - Count non-zero coefficients (perhaps after thresholding)
  - Entropy of histogram
  - Kurtosis

$$k \equiv \frac{m_4}{m_2^2} \quad \text{or} \quad k \equiv \frac{m_4}{m_2^2} - 3$$

m_n: n-th order central moment

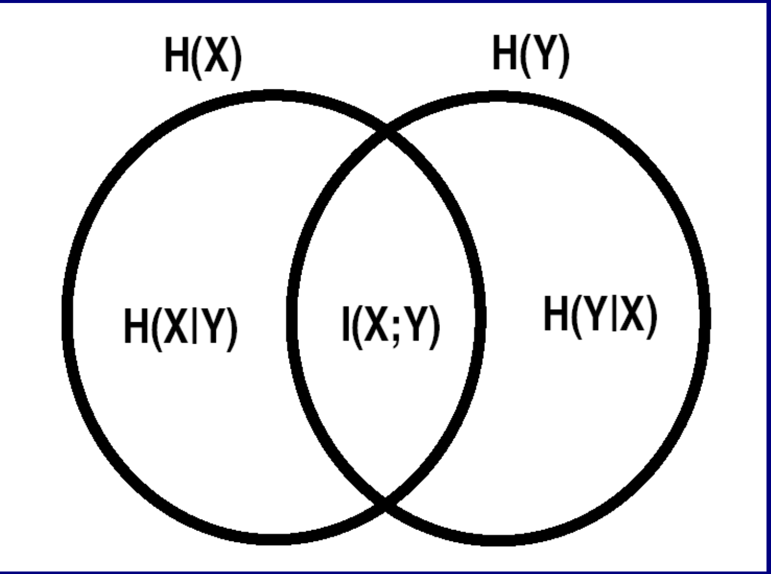
  - How to measure sparseness is still not a completely resolved issue



A typical wavelet coefficient histogram in log scale

Independence

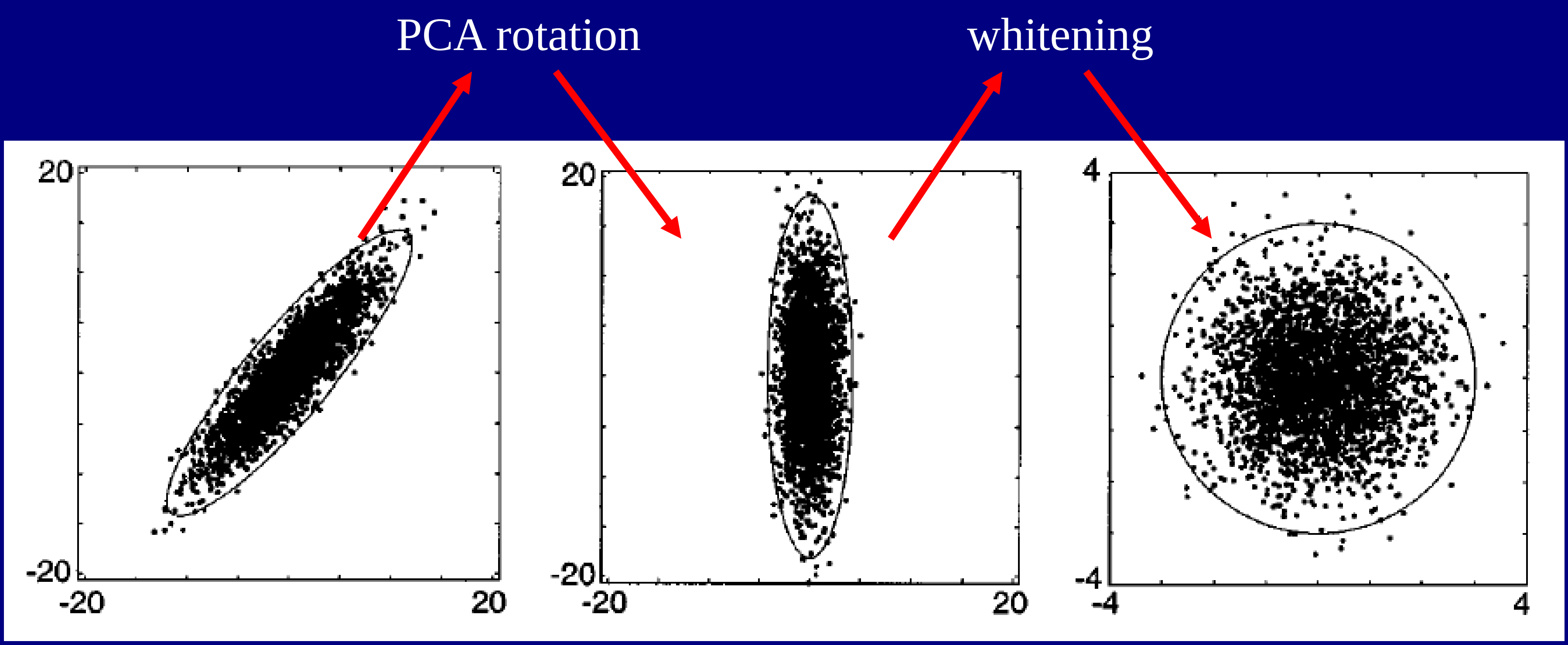
- **From marginal to joint statistics**
  - Why marginal is not enough?
    strong correlations
- **Why independence?**
  - Independent → no mutual information → minimal redundancy
- **Measure of Independence: Mutual Information**



entropy, conditional entropy and mutual information

Principle Component Analysis

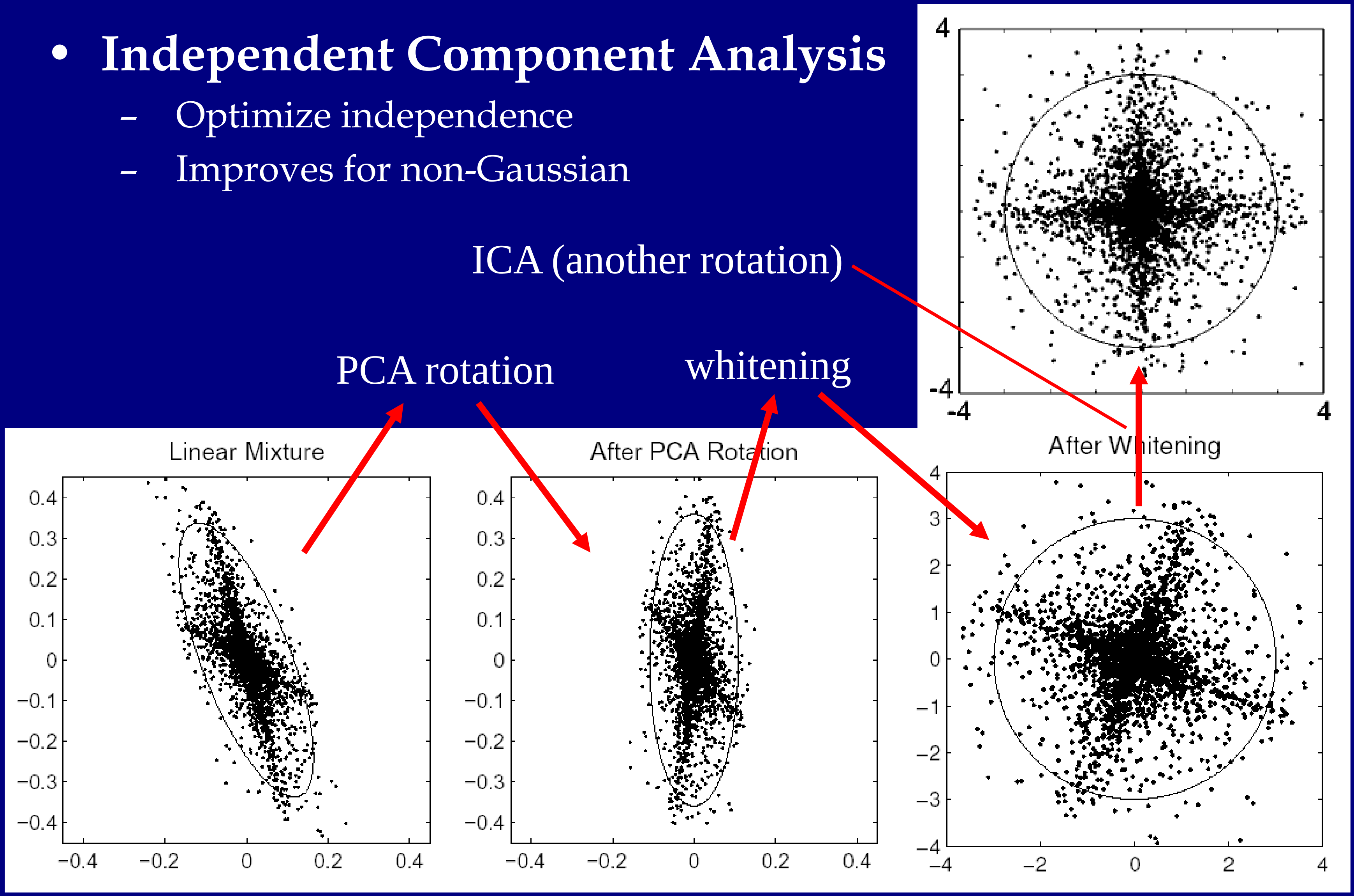
- **Principle Component Analysis**
  - Optimize energy compaction
  - Perfect for Gaussian



Simoncelli & Olshausen ’01

Independent Component Analysis

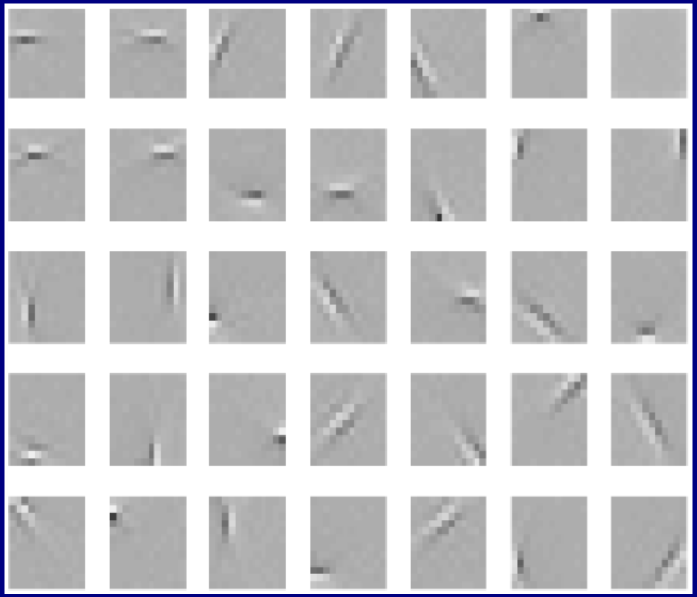
- **Independent Component Analysis**
  - Optimize independence
  - Improves for non-Gaussian



Simoncelli & Olshausen ’01

Independent Component Analysis

- **Independent Component Analysis for Natural Images**



After Bell & Sejnowski '97

- Bases trained from a large collection of natural images
- What do we see?

  wavelets!

  edge detection filters!

  neural responses!

- **Debate about ICA in image processing**
  - Natural images are generally **NOT** linear combinations of independent sources, e.g., occlusion (nonlinear for sure)
  - Note: ICA is still a linear transform

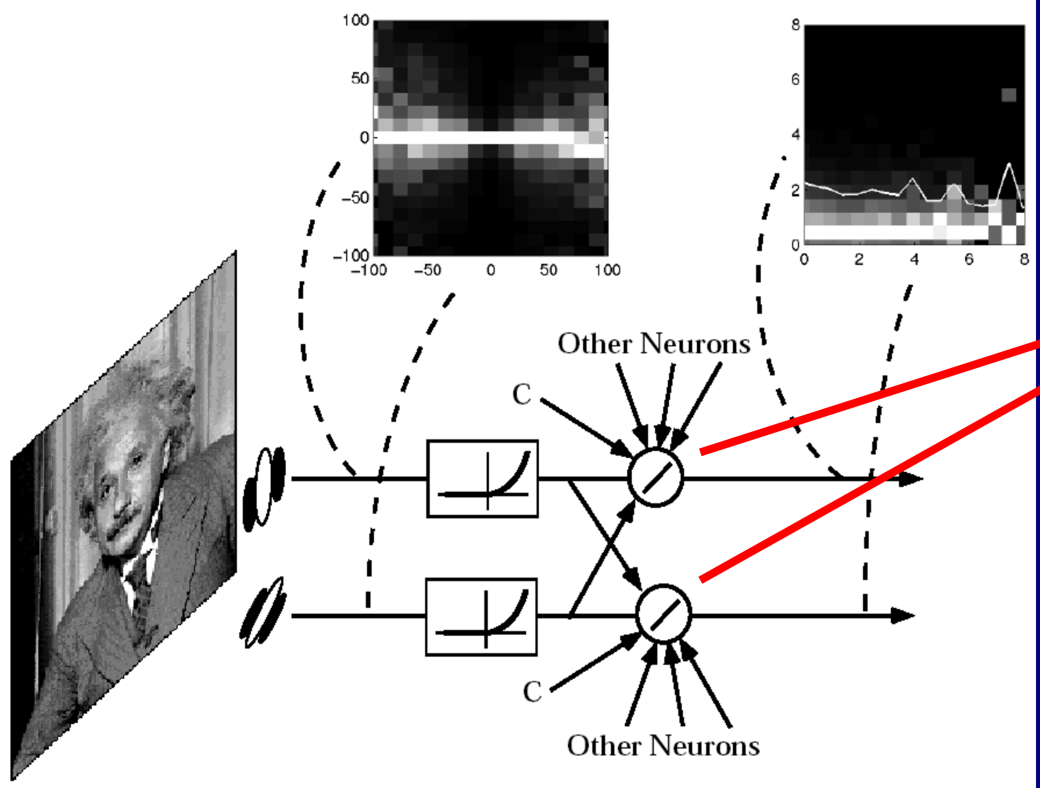
Ref: *Representation and Analysis of Visual Images* lecture notes by Prof. Eero P. Simoncelli

Other Properties

- Linear Phase
- Compact Support
- Scale-Invariance
- Translation-Invariance
- Rotation-Invariance
- Separable and Directional Filters
- Steerability
- Spatial vs. Frequency Implementations
- Integer Implementation
- ……

Divisive Normalization

- **Improving Independence: Why ICA Still not Enough?**
  - Natural images are generally **NOT** linear combinations of independent sources
- **A Promising Method: Divisive Normalization**



Wainwright, Schwartz & Simoncelli '02

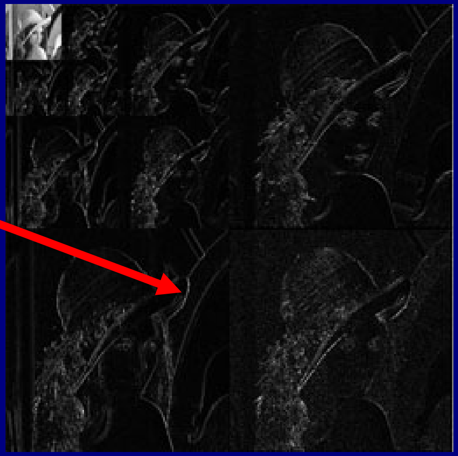
$$R_j \equiv \frac{L_j^2}{\sigma_j^2 + \sum_k w_{jk} L_k^2}$$

Every coefficient is normalized by a weighted sum of the square of its neighborhood

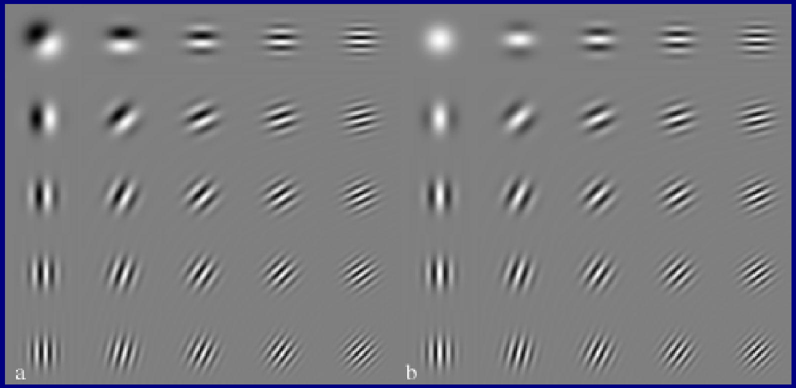
This is also a good model for neurons in visual cortex

Adaptive Basis Selection

- **Idea: Bases Selected Based on Image being Analyzed**
  - Simple implementation: choose large coefficients



- **Dictionary-based Methods**
  - Create (large) basis dictionary
  - Matching Pursuit [Mallat & Zhang '93]



Gabor dictionary

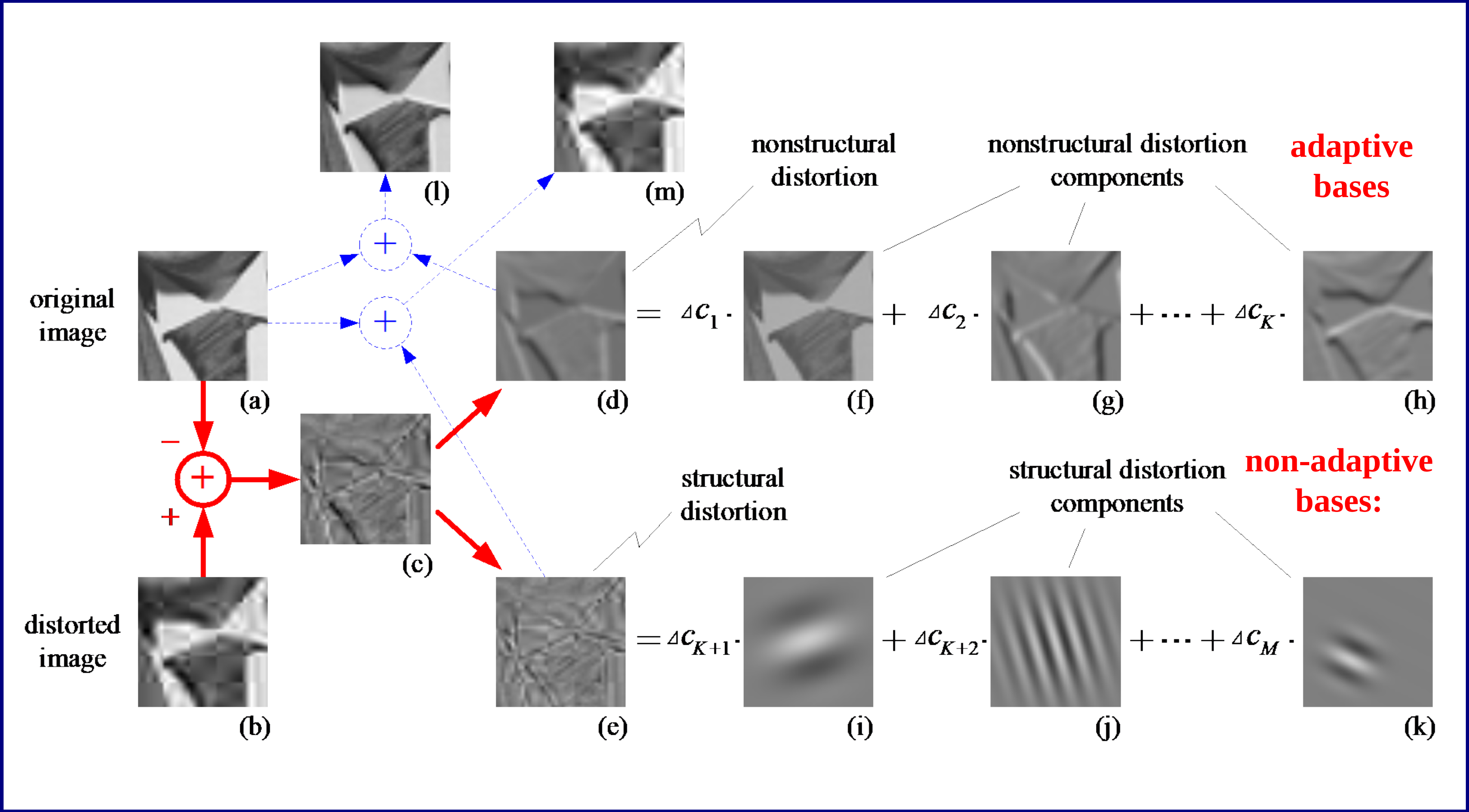
reconstructed images with 2500/5000 bases

Lee, IEEE Trans. PAMI '96

Bergeaud & Mallat, ICIP'95

Adaptive Basis Generation

- **Idea: Bases Computed from Image being Analyzed**



Wang & Simoncelli, ICIP'05