E9 205 Machine Learning for Signal Processing

Introduction to Machine Learning of Sensory Signals

19-08-2019

Instructor - Sriram Ganapathy (sriramg@iisc.ac.in)

Teaching Assistant - Prachi Singh (prachisingh@iisc.ac.in)

http://leap.ee.iisc.ac.in/sriram/teaching/MLSP_19/

Feature Extraction

Scope for this course

I. Feature Extraction in Text.

II. Feature Extraction in Speech and Audio signals.

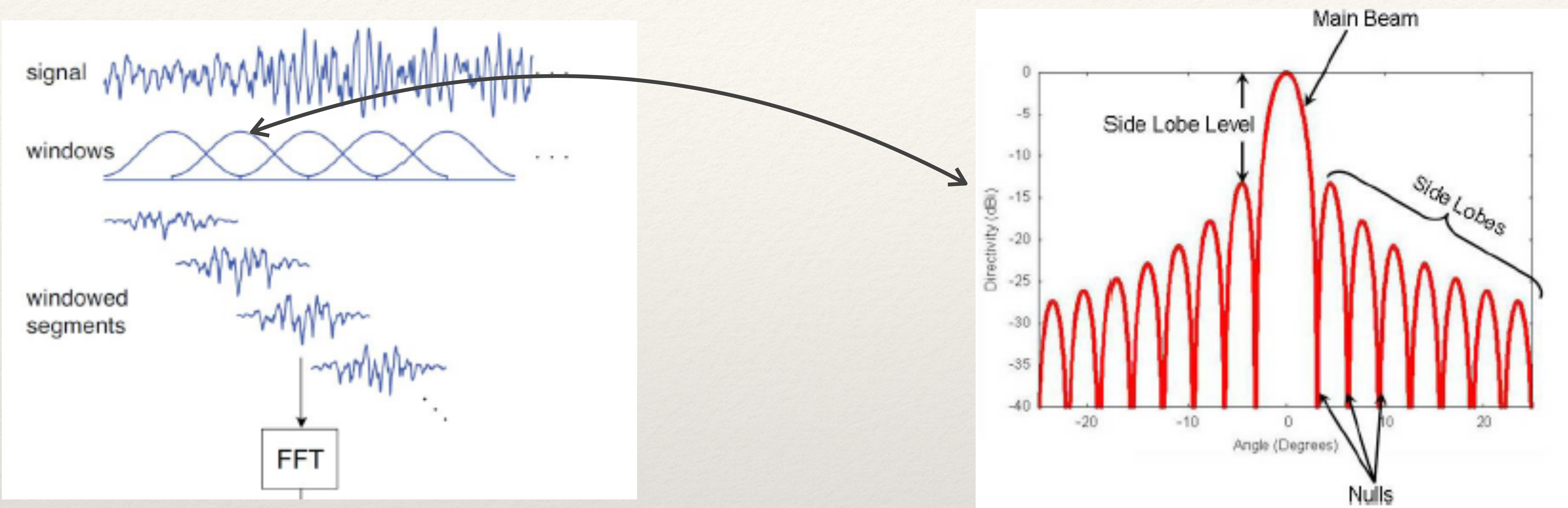
III. Processing of Images.

Established 1911

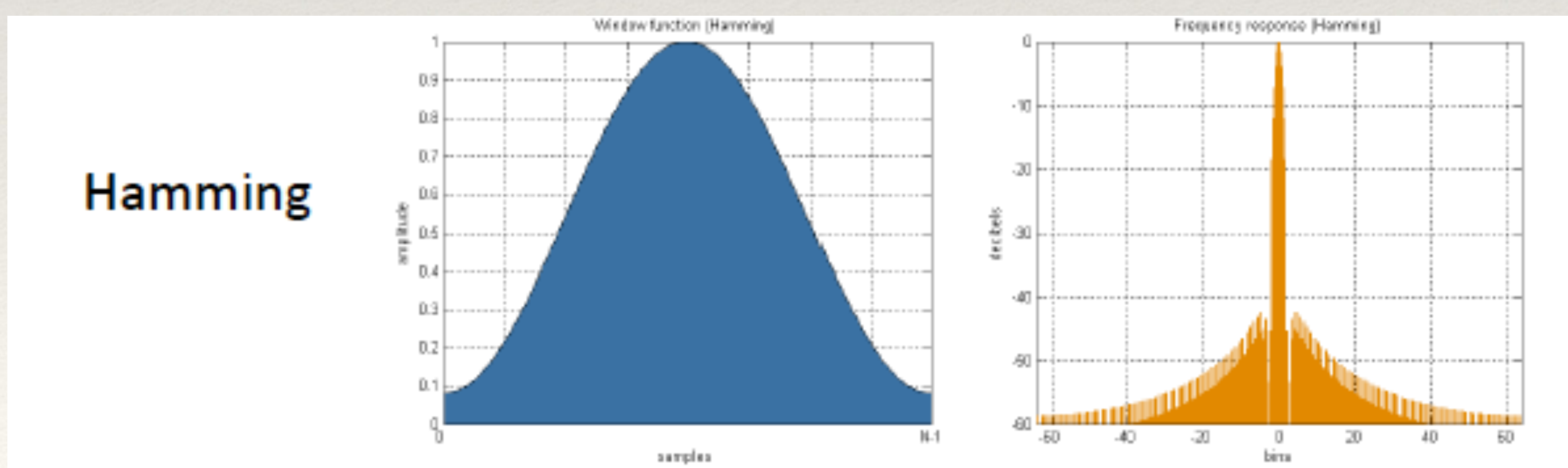
INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Speech and Audio Processing

Summary of STFT Properties



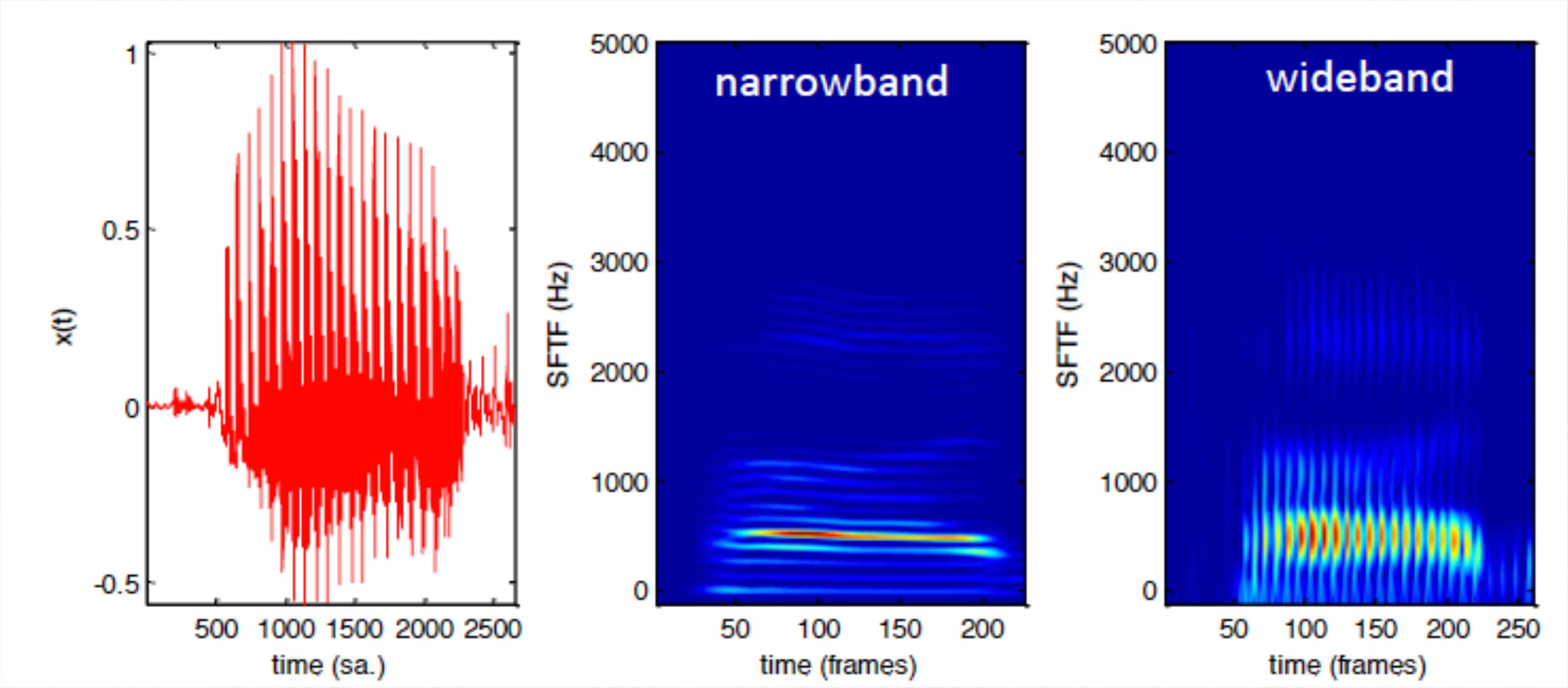
$$X[k, n_0]$$



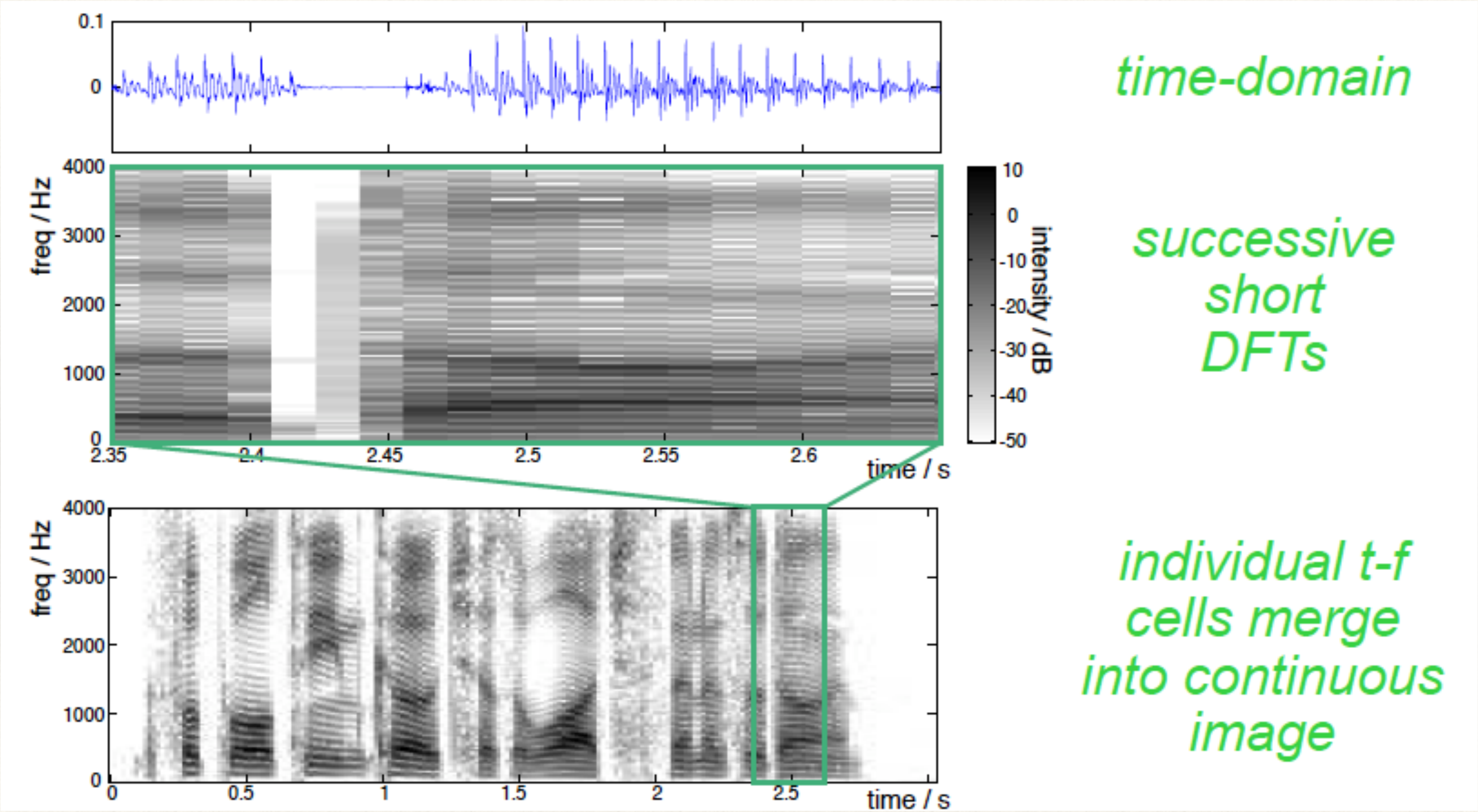
http://en.wikipedia.org/wiki/Window_function

Narrowband versus Wideband

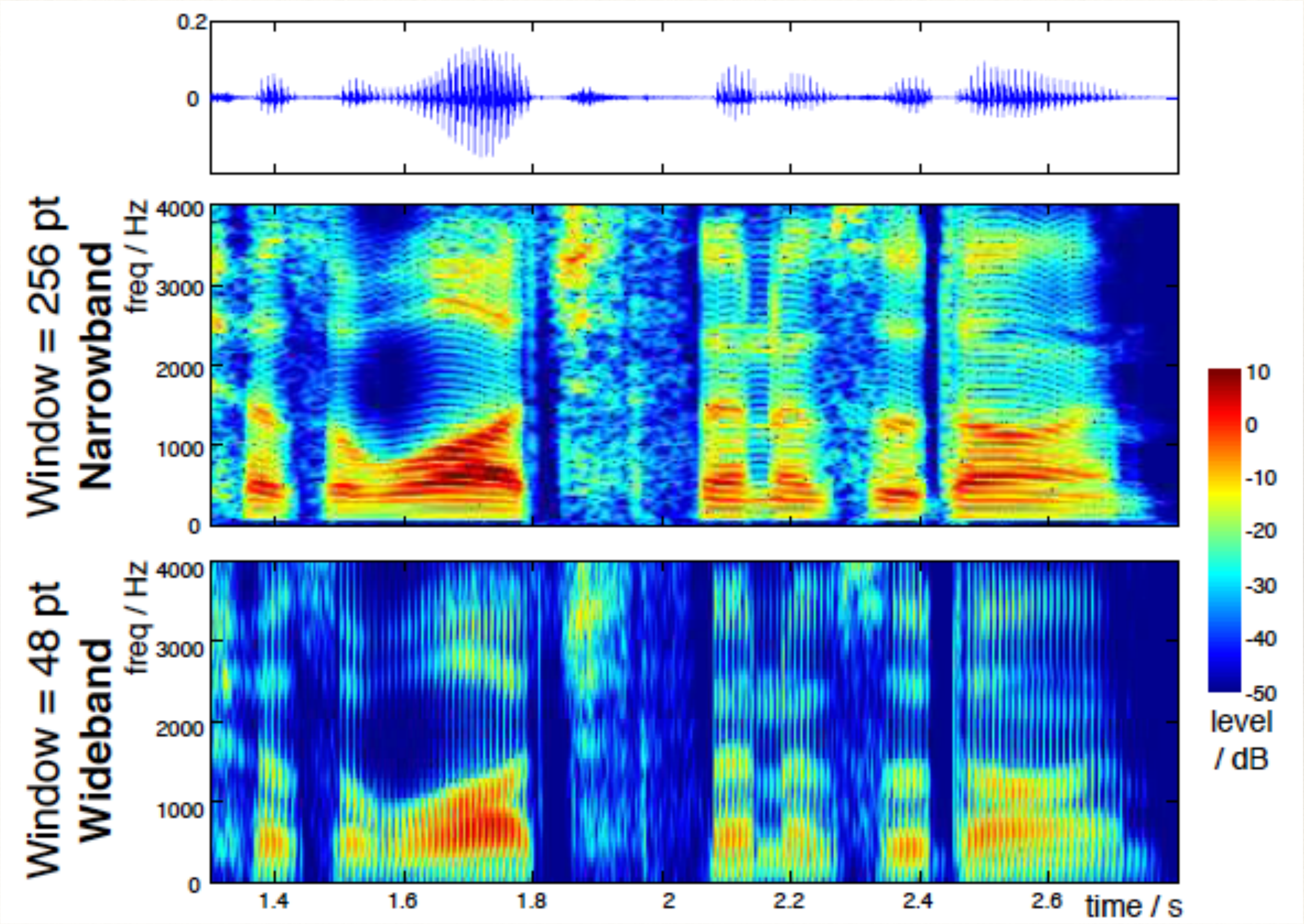
- Short windows - poor frequency resolution - wideband spectrogram
- Long windows - poor time resolution - narrowband spectrogram



Spectrogram of Real Sounds



Narrowband versus Wideband



Dan Ellis, "STFT Tutorial"

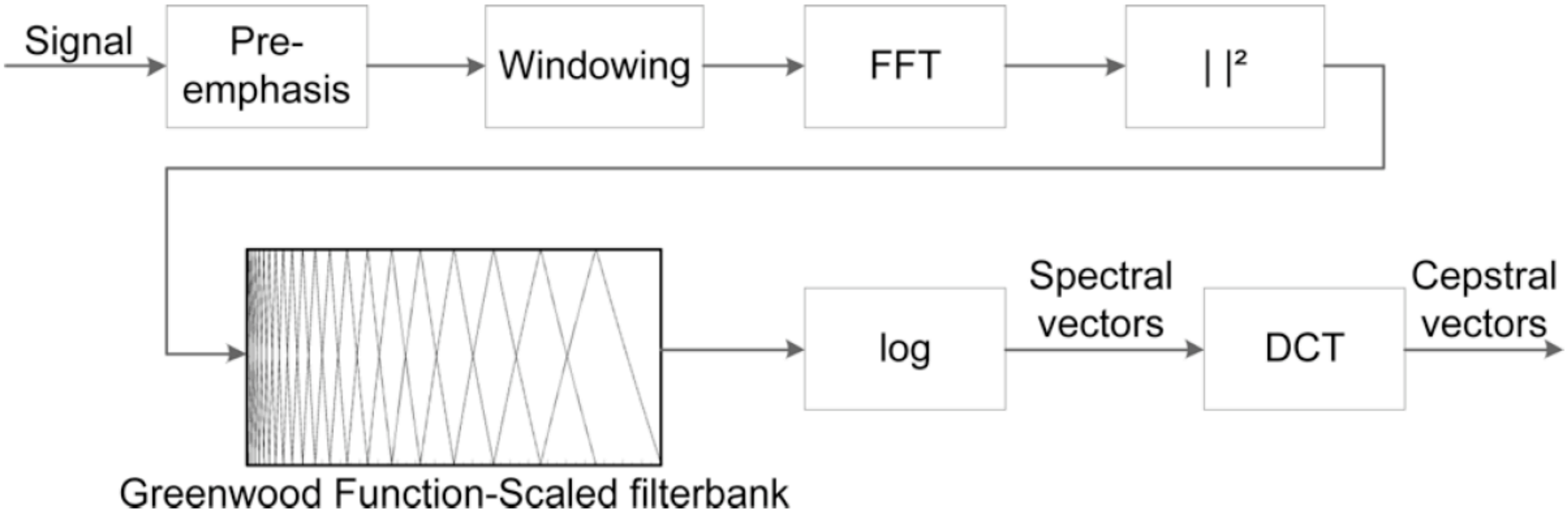
Mel Frequency Cepstral Coefficients

MFCC

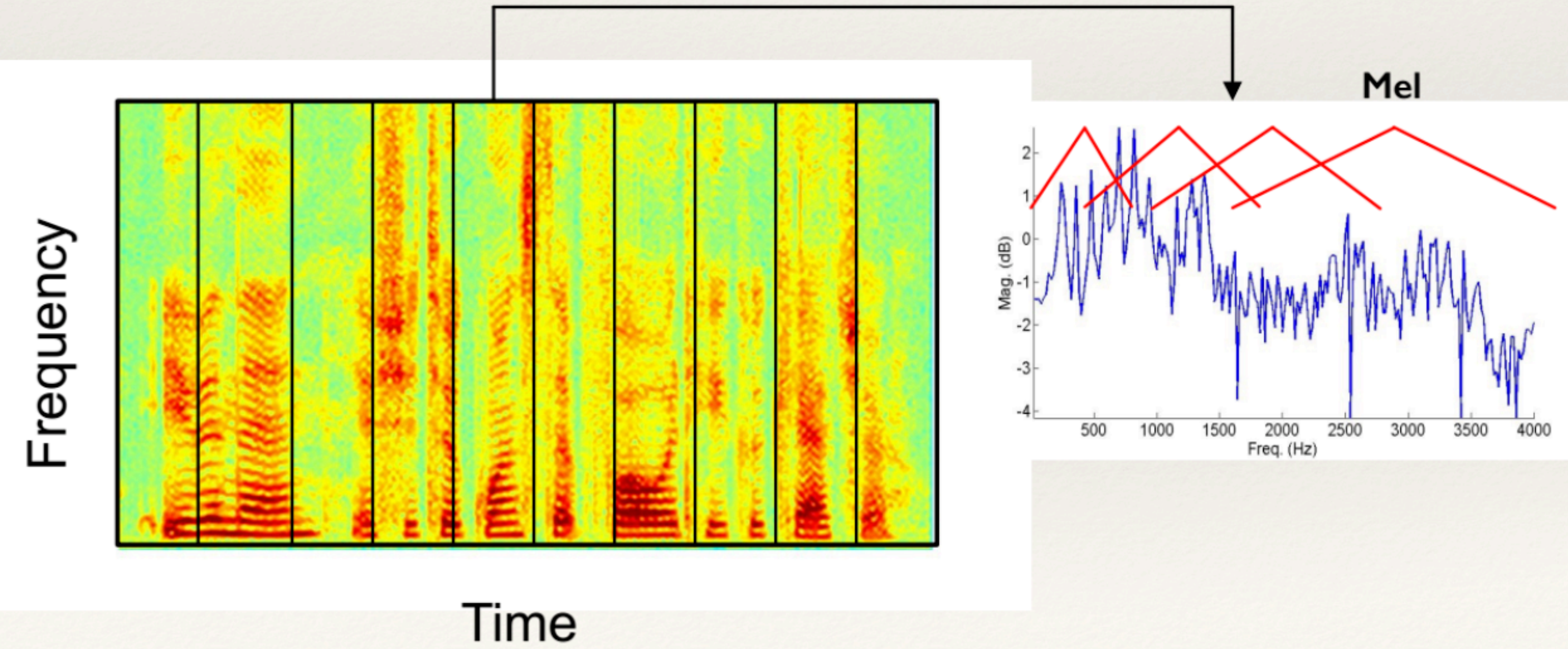
- MFCC coefficients model the spectral energy Distribution in a perceptually meaningful way
- Why do we need?
  - Automatic speech recogonition
  - Speaker Identification
  - Audio classification

Mel Frequency Cepstral Coefficients

- Implementation steps



Mel Frequency Cepstral Coefficients



Mel Frequency Cepstral Coefficients

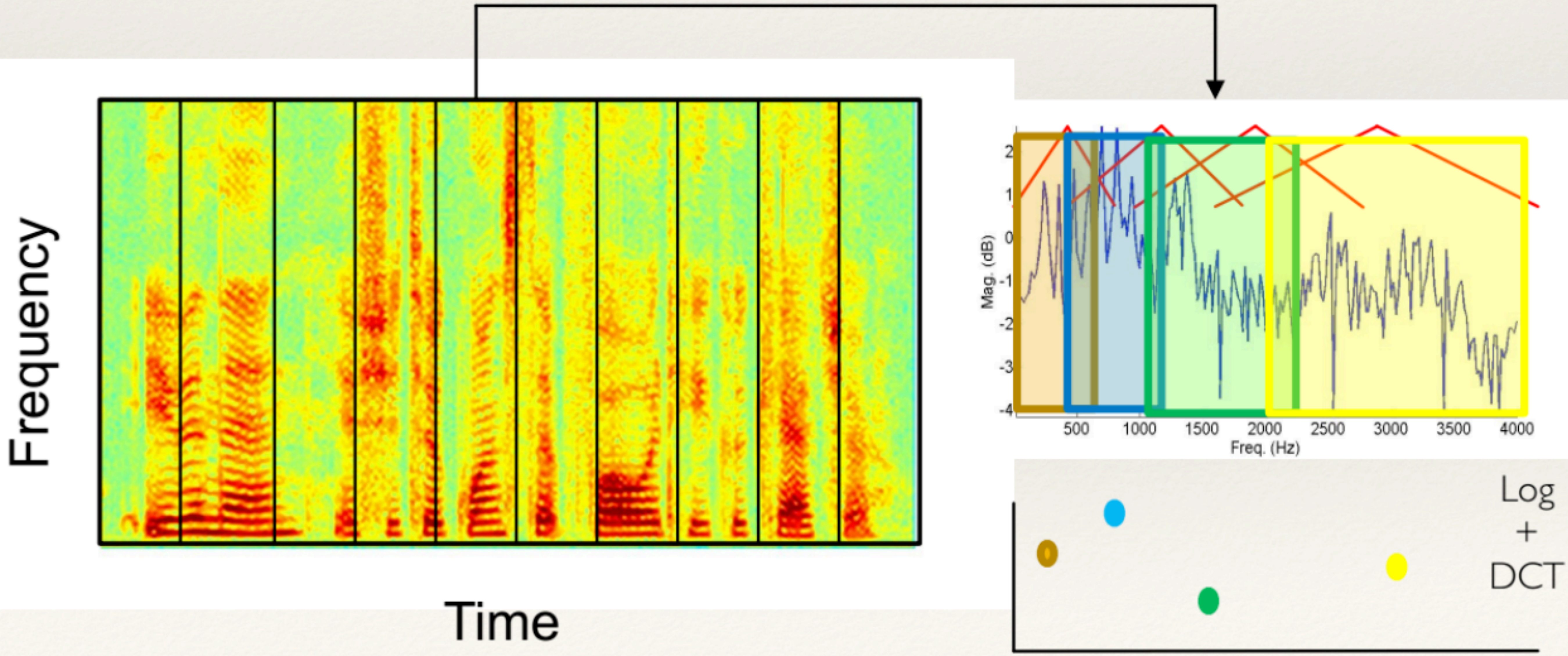


Image Processing

Image Capture and Representation

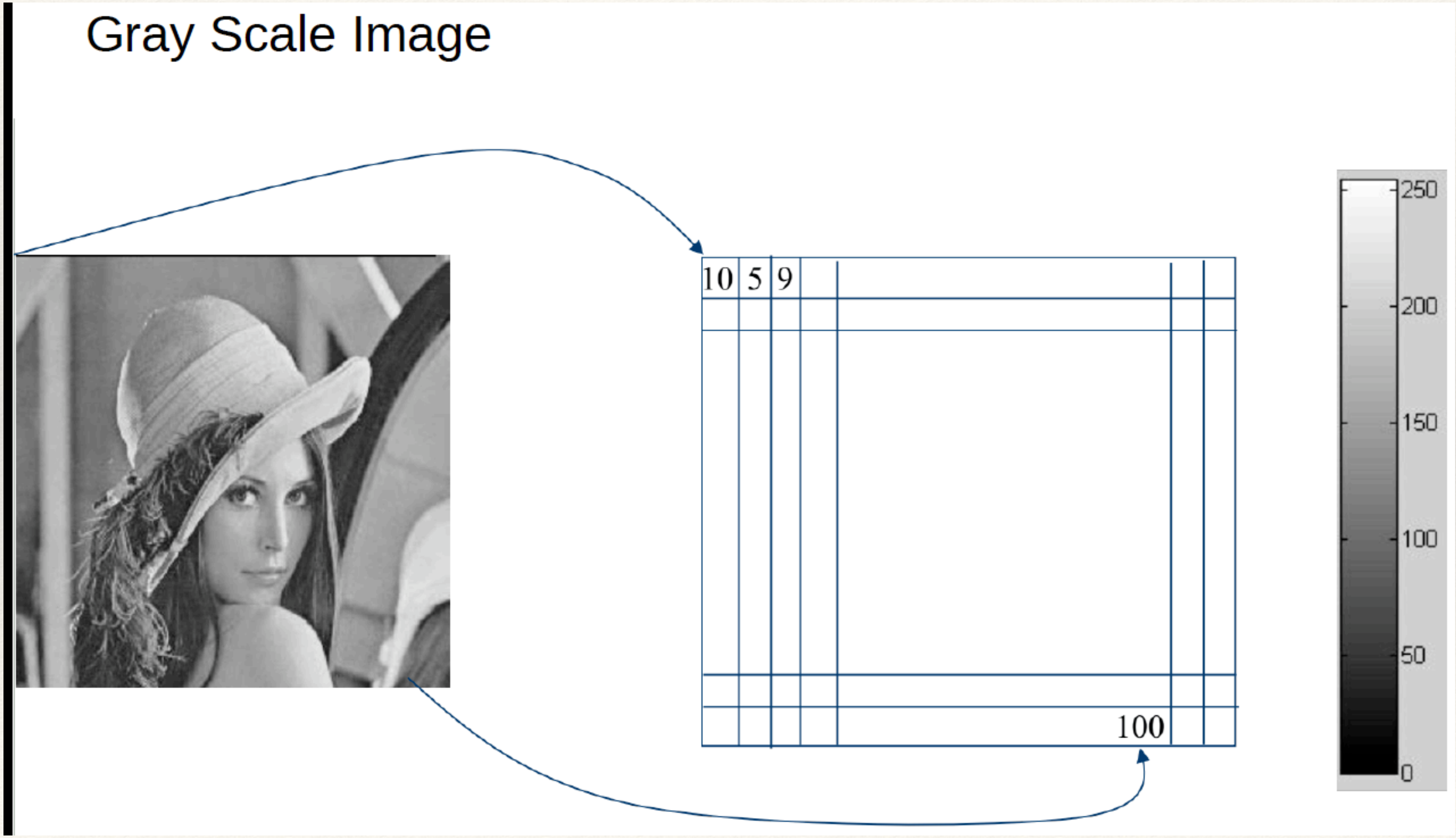
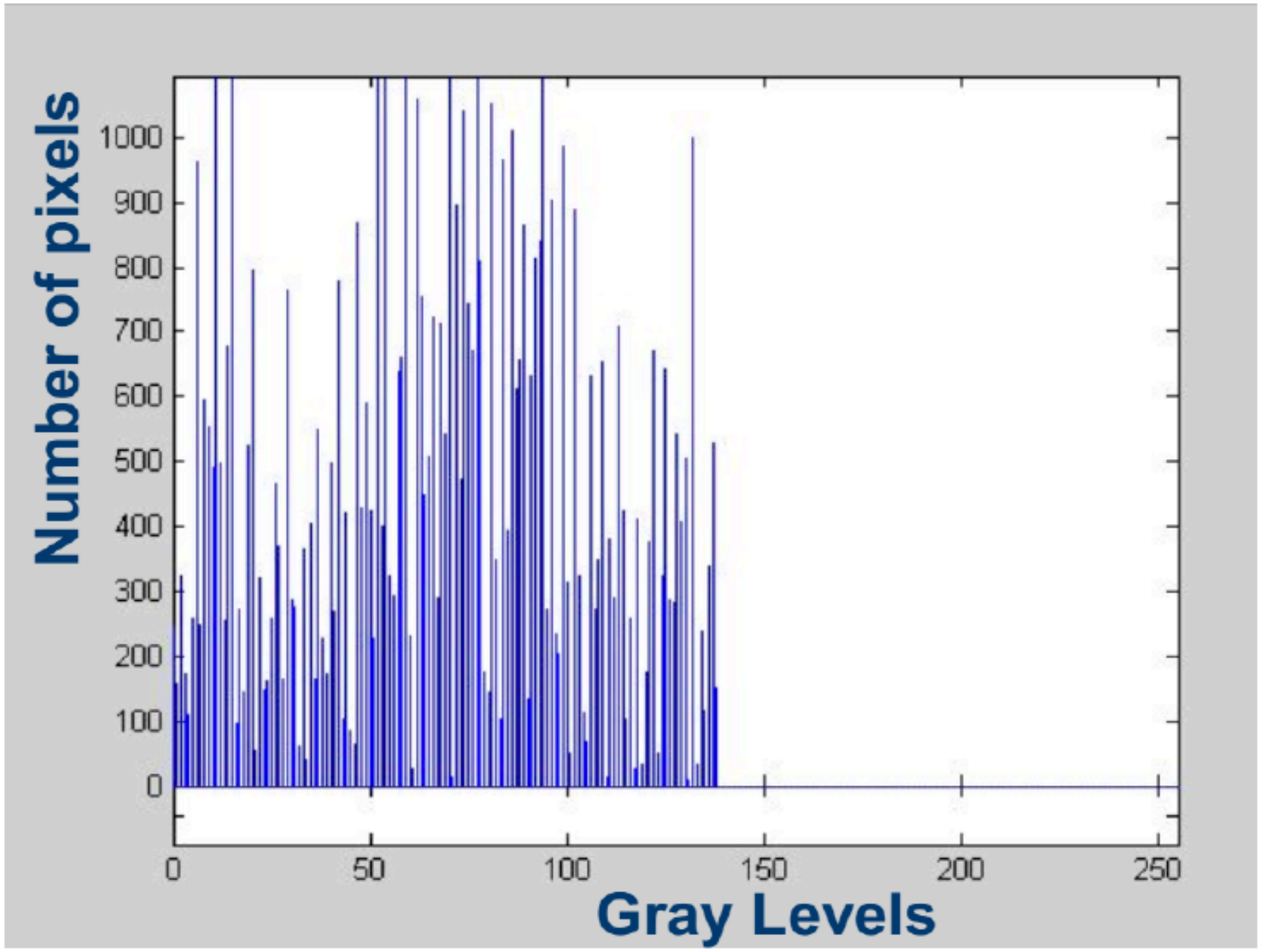


Image Capture and Representation

Histogram



- Histogram captures the distribution of gray levels in the image.
- How frequently each gray level occurs in the image

Image Filtering

- Image filtering: compute function of local neighborhood at each position

  Really important!
  - Enhance images
  - Denoise, resize, increase contrast, etc.
  - Extract information from images
  - Texture, edges, distinctive points, etc.
  - Detect patterns
  - Template matching

Image Filtering

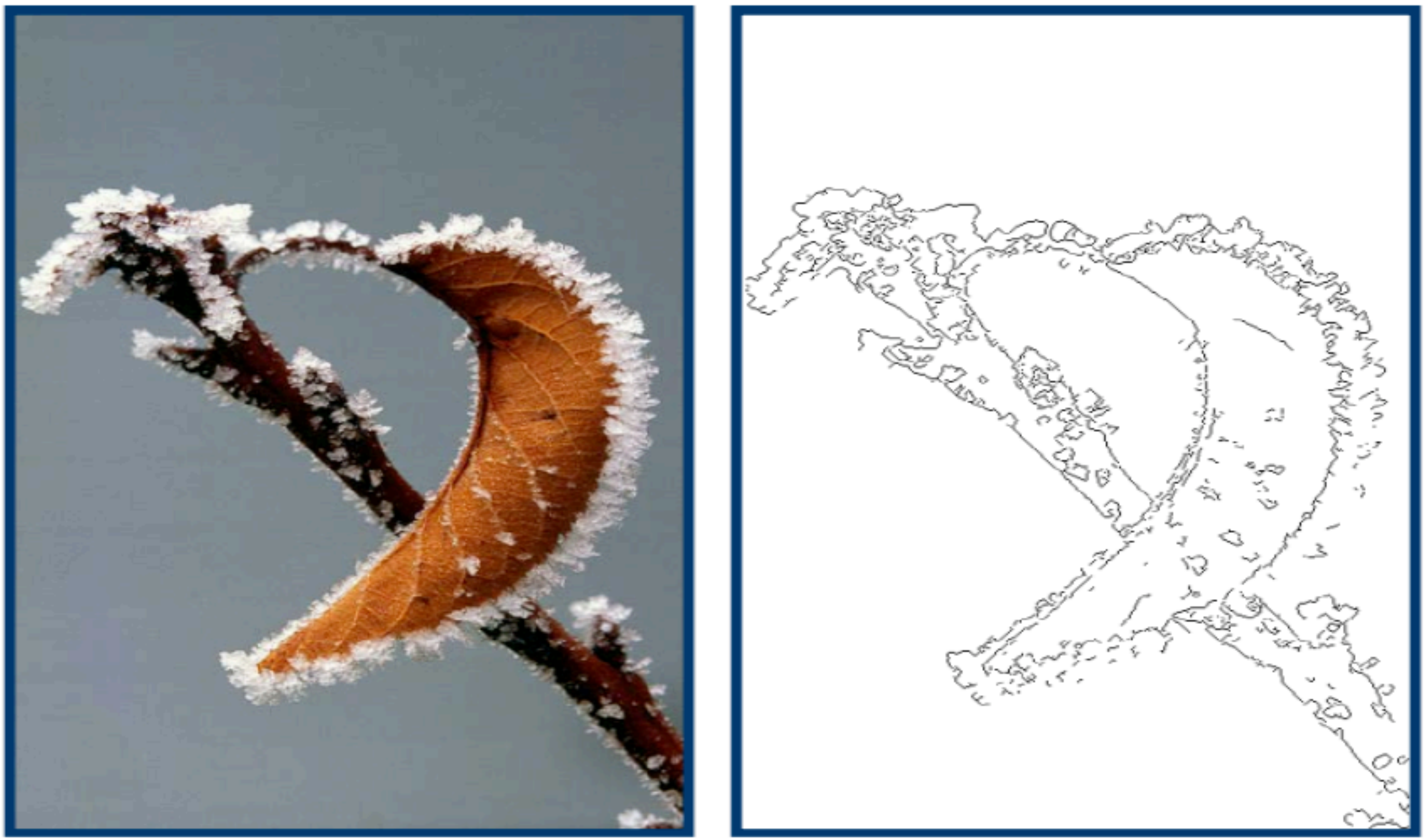
Given function	$f(x, y)$
Gradient vector	$\nabla f(x, y) = \begin{bmatrix} \dfrac{\partial f(x, y)}{\partial x} \\ \dfrac{\partial f(x, y)}{\partial y} \end{bmatrix} = \begin{bmatrix} f_x \\ f_y \end{bmatrix}$
Gradient magnitude	$\lvert \nabla f(x, y) \rvert = \sqrt{f_x^2 + f_y^2}$
Gradient direction	$\theta = \tan^{-1} \dfrac{f_x}{f_y}$

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x) - f(x - \Delta x)}{\Delta x} = f'(x)$$

$$\frac{df}{dx} = \frac{f(x) - f(x-1)}{1} = f'(x)$$

$$\frac{df}{dx} = f(x) - f(x-1) = f'(x)$$

Edge Detection Example



Convolution Operation in Images

Convolution

$$f * h = \sum_{k} \sum_{l} f(k,l)\,h(-k,-l)$$

f = Image

h = Kernel

h

h_1	h_2	h_3
h_4	h_5	h_6
h_7	h_8	h_9

$X - flip$

h_7	h_8	h_9
h_4	h_5	h_6
h_1	h_2	h_3

$Y - flip$

f

f_1	f_2	f_3
f_4	f_5	f_6
f_7	f_8	f_9

*

h_9	h_8	h_7
h_6	h_5	h_4
h_3	h_2	h_1

$$f * h = f_1 h_9 + f_2 h_8 + f_3 h_7 + f_4 h_6 + f_5 h_5 + f_6 h_4 + f_7 h_3 + f_8 h_2 + f_9 h_1$$

Matrix Derivatives (Appendix C, PRML, Bishop)

Dimensionality Reduction - PCA (Chapter 12.1, PRML, Bishop)

Principal Component Analysis

- Reducing the data $\mathbf{x}_n$ of dimension D to lower dimension $M < D$
- Projecting the data into subspace which preserves maximum data variance
  - Maximize variance in projected space
- Equivalent formulated as minimizing the error between the original and projected data points.