乐　　统

Content-Based Music Recommender Systems

1

2

3

4

5

6

1

-
-
-

1

-
-

2

“ ”

2

3

-
  -
-
  1.
     -
     -
  2.
     -

3

- I_{ui} $I=\{$ i_1 ... $i_n\}$ I_{ua}
-

3

-
  - Metadata-based Approaches
  - Content-based Recommendation Approaches
  - Hybrid Recommendation Approaches

3.1
Metadata-based Approaches

-
-
-
-
  -
  - Web
  -
  -

3.2
Content-based Recommendation Approaches

-
-
  -
  - item-to-item

3.2
Content-based Recommendation Approaches

-
  -
  -
-
  -
  - SÖ NI

Và

3.2
Content-based Recommendation Approaches

•

•

3.3
Hybrid Recommendation Approaches

.

.

.

4
Music Signal Processing Fundamentals

- 1
  - f_s 1 s Hz
  - PCM
  -
    - 8KHZ 11kHz 22kHz 44.1KHz 96KHZ 192KHz

4 Music Signal Processing Fundamentals

- 2 Spectrum Analysis
  - 1 Discrete Fourier Transform (DFT)
    -
$$X[k] = \mathbf{DFT}(x[n])] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi nk/N} \qquad k = 0, 1, \ldots, N-1. \qquad (2.1)$$
  - 2 inverse Discrete Fourier Transform

    iDFT
    -
$$x[n] = \mathbf{iDFT}(X[k]) = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j2\pi nk/N} \qquad n = 0, 1, \ldots, N-1. \qquad (2.2)$$

4
Music Signal Processing Fundamentals

- 2 Spectrum Analysis
  - 3 magnitude spectrum

$$|X[k]| = \sqrt{\mathrm{Re}(X[k])^2 + \mathrm{Im}(X[k])^2} \tag{2.3}$$

  - 4 phase spectrum ϕ[k].

$$\varphi[k] = \arctan \frac{\mathrm{Im}(X[k])}{\mathrm{Re}(X[k])} \tag{2.4}$$

4
Music Signal Processing Fundamentals

$$X[k]_{dB} = 20 \log_{10} (|X[k]|) \tag{2.5}$$

4
Music Signal Processing Fundamentals

- 2.1 x[n]
- 22,000 Hz
  11 K Hz
- MIR
- MIR DFT

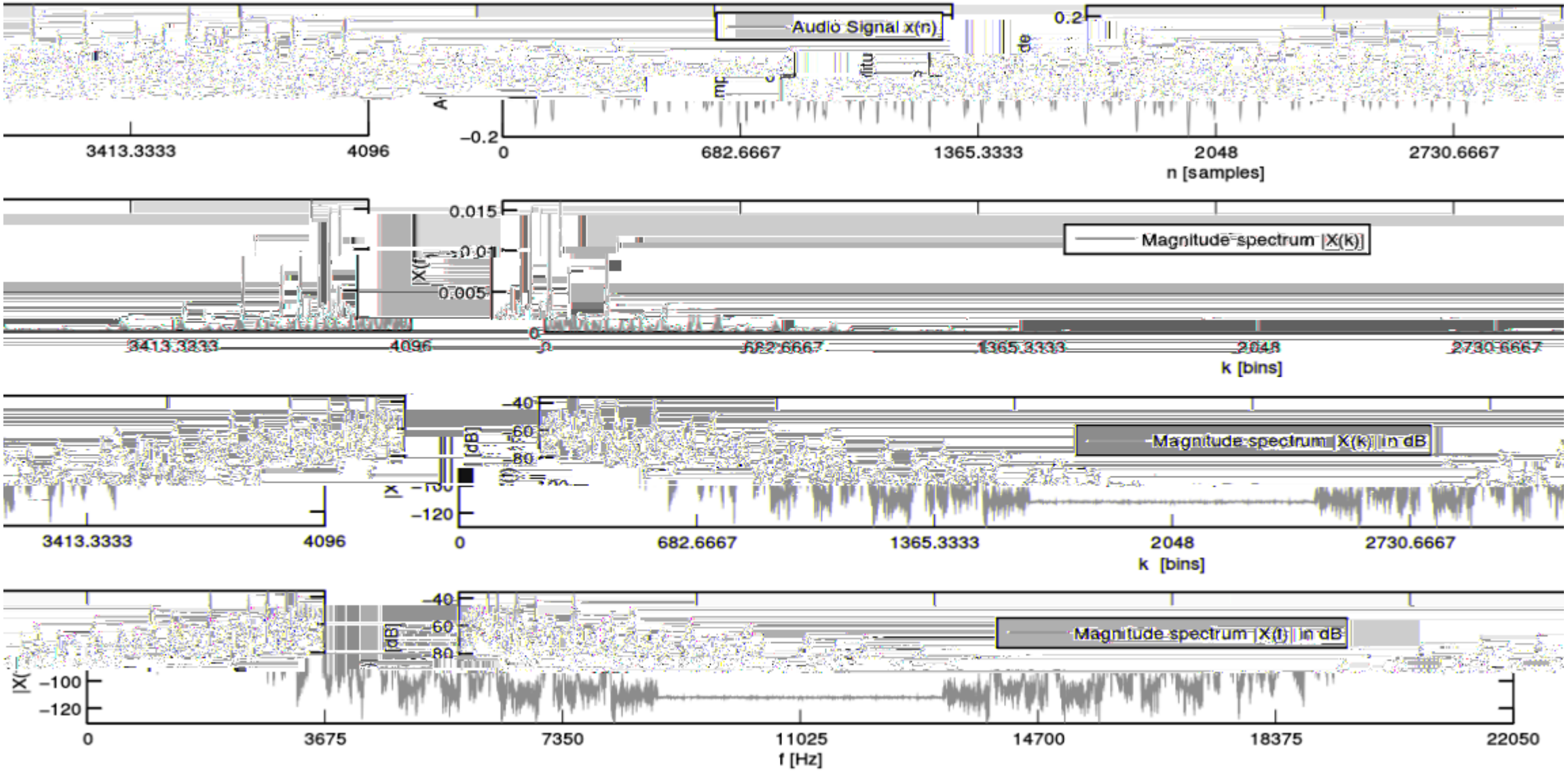


Figure 2.1: Visualization of an audio signal, the magnitude spectrum and the magnitude spectrum in dB.

4
Music Signal Processing Fundamentals

- 3 Auditory Scales
  -
  -
  - DFT
    Mel Scale Boek ERB Cent
    FHz

4 Music Signal Processing Fundamentals

- 3 Auditory Scales
  - 1 Mel-Scale

$$f_{\text{mel}} = 2595 \log_{10}\left(\frac{f_{\text{Hz}}}{700} + 1\right)$$

  - 2 Bark-Scale

$$f_{\text{bark}} = 13 \arctan(0.00076 f_{\text{Hz}}) + 3.5 \arctan((f_{\text{Hz}/7500})^2)$$

4
Music Signal Processing Fundamentals

- 3 Auditory Scales
  - 3 ERB-Scale

$$BW_{\mathrm{Hz}} = 24.7\,(0.00437 f_c + 1)$$

  - 4 Cent-Scale

$$\Delta f_{\mathrm{cent}} = 1200 \log_2 (\frac{f_a}{f_b}).$$

4
Music Signal Processing Fundamentals

- 4

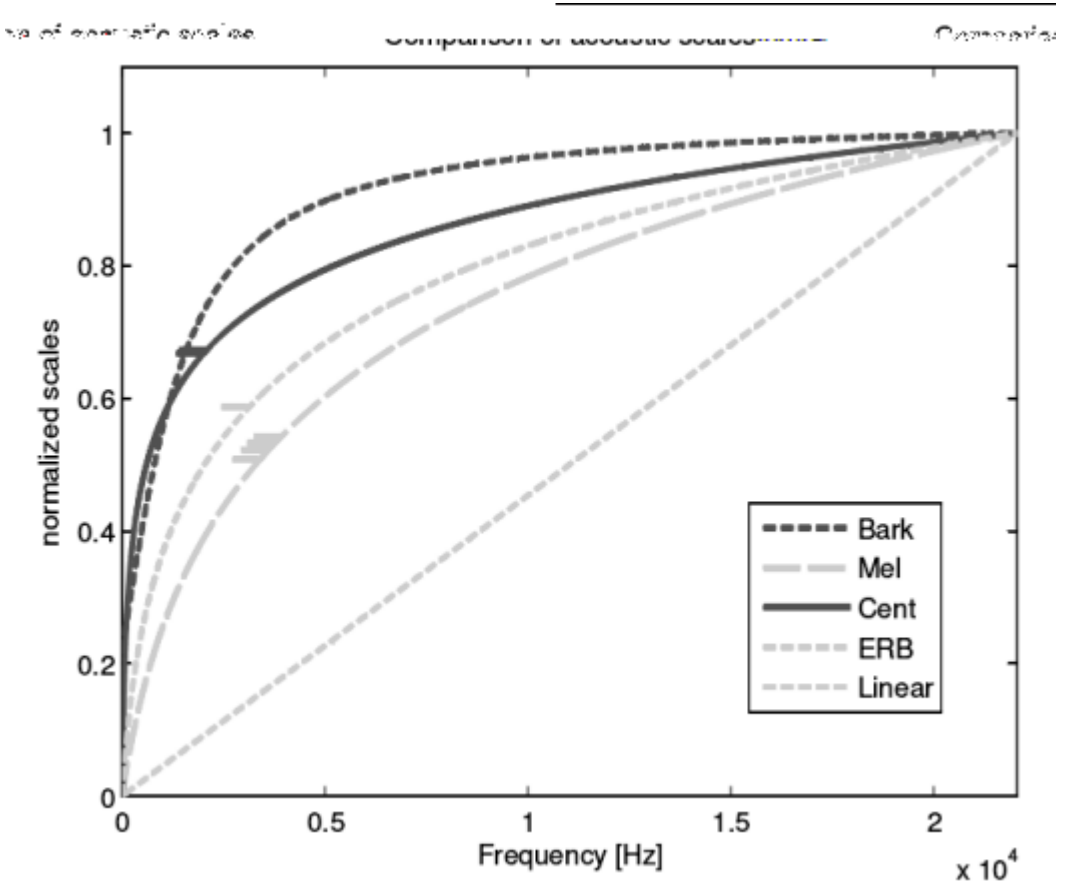


Figure 2.2: Comparison of auditory scales.

4
Music Signal Processing Fundamentals

- 5
  - x(t)

TFR

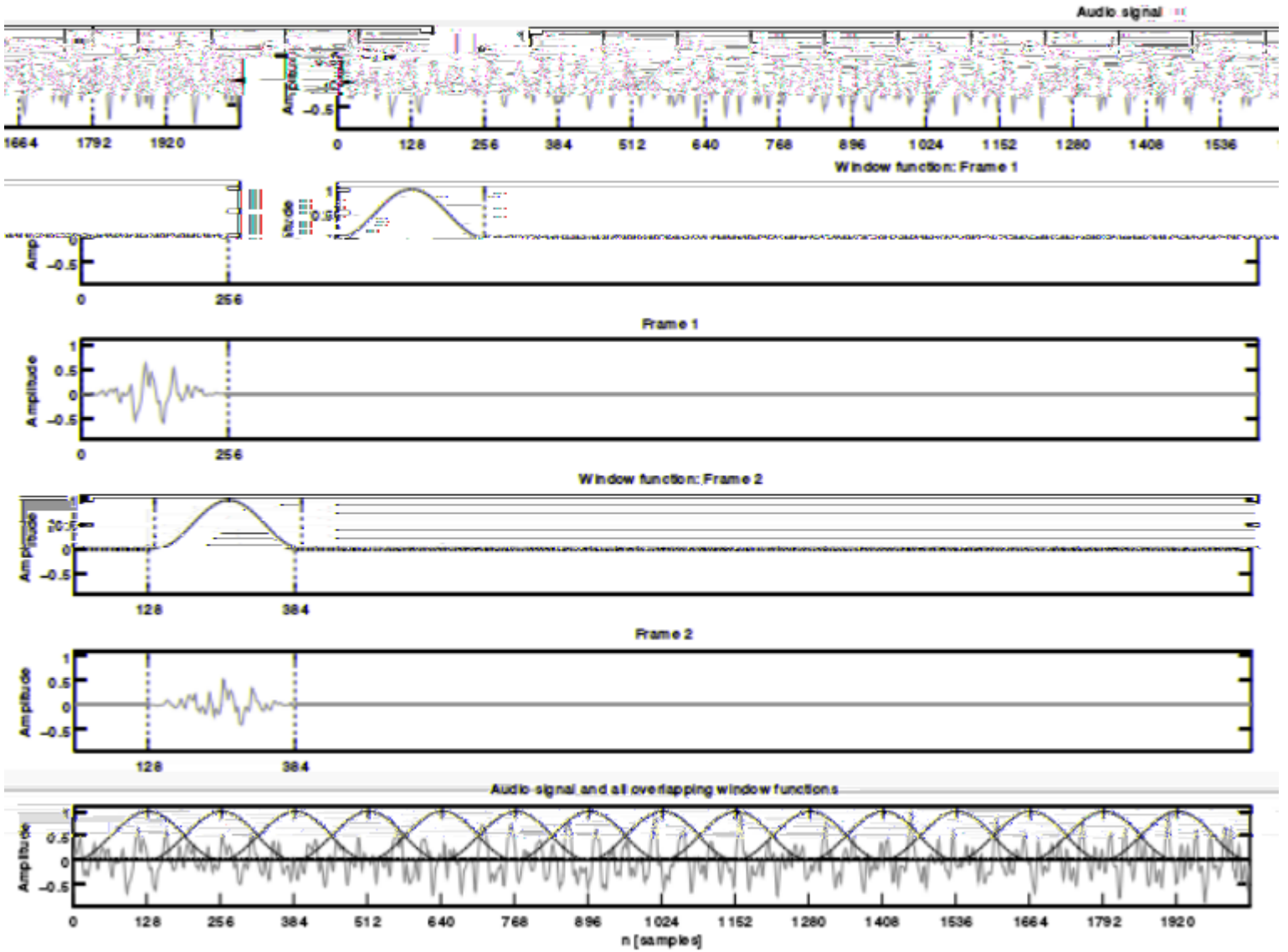


Figure 2.4: Windowing of an audio signal using a Hanning window function .

4
Music Signal Processing Fundamentals

- STFT W[n] m

STFT

$$\mathbf{STFT}\{x[n]\} \equiv X[m,k] = \sum_{n=0}^{N-1} x[n]w[n-m]e^{-j\omega kn}$$

4
Music Signal Processing Fundamentals

- 2
- “ ” W[n]

x[n] x[k] W[n]

W[k]

$$x[n] \overset{\mathcal{F}}{\leftrightarrow} X[k]$$

$$w[n] \overset{\mathcal{F}}{\leftrightarrow} W[k]$$

- y[n]

$$y[n] = x[n]w[n]$$

4
Music Signal Processing Fundamentals

- Y[n]

$$Y[k] = X[k] \star W[k]$$

-

4
Music Signal Processing Fundamentals

- ω ω

-

Hanning Kaiser

- MIR

$$w[n] = 0.5(1 - cos(\frac{2\pi(n-1)}{N})), \quad 0 \leq n < N$$

4
Music Signal Processing Fundamentals

•

STFT

STFT

ΔT K

ΔT =KTs Ts STFT Δ f

Δf Fs/K

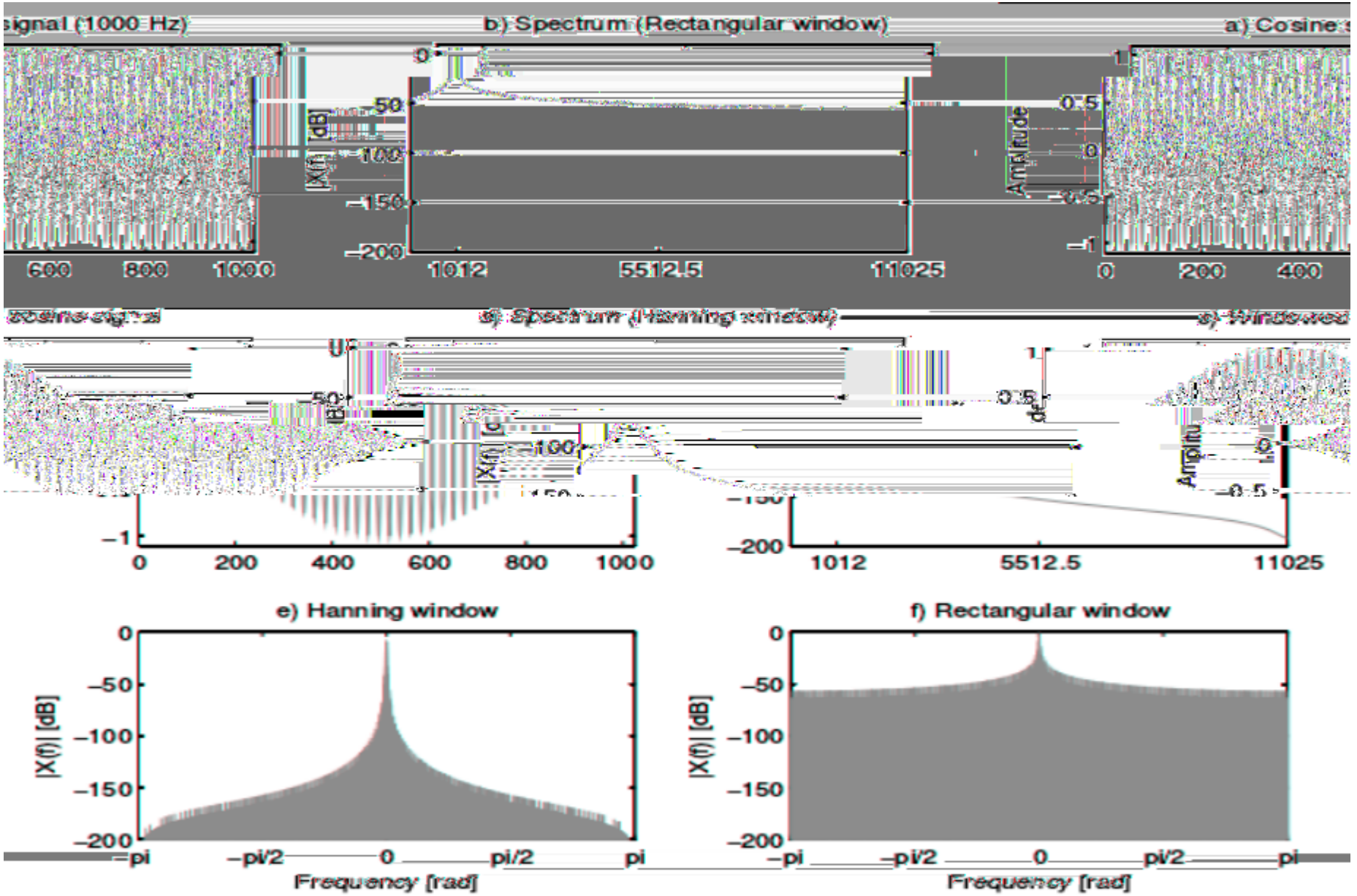


Figure 2.5: Windowing of an audio signal using a Hanning window function

.

4
Music Signal Processing Fundamentals

- 6 (Audio Normalization)
-

x[n]

A

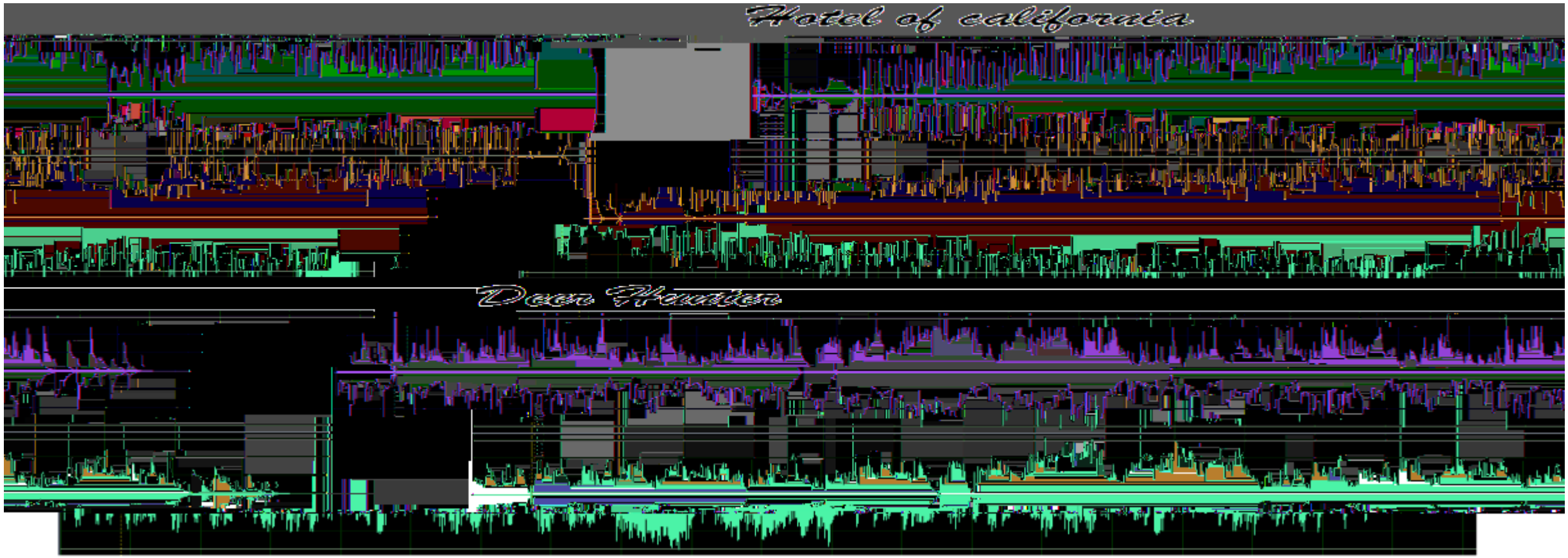
$$\hat{x}[n] = ax[n]$$

- x[k] a

$$|\hat{X}[k]| = a|X[k]|$$

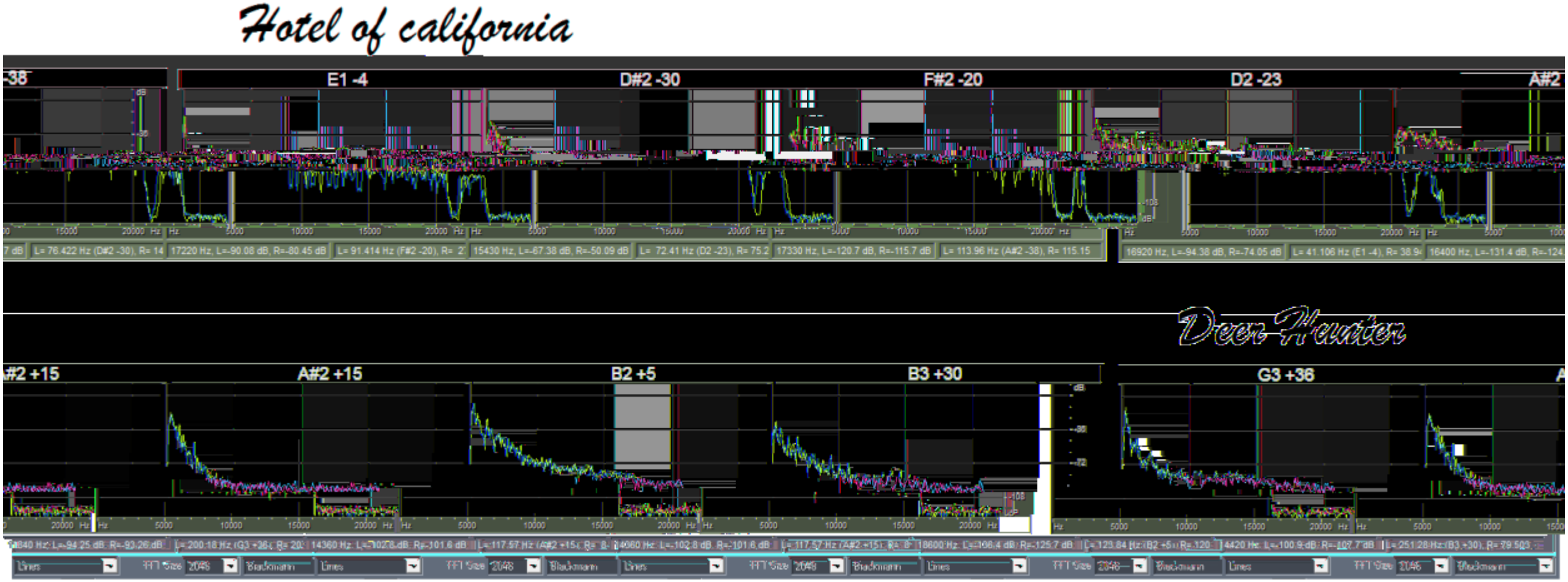
5

.



5

•


Hotel of california
E1 -4
D#2 -30
F#2 -20
D2 -23
A#2
Deer Hunter
A#2 +15
A#2 +15
B2 +5
B3 +30
G3 +36

5

- “ ”
    1.
    2.
    3.
        1.
        2.
    4.

6

-

-