

Lecture
Information Retrieval for Music and Motion

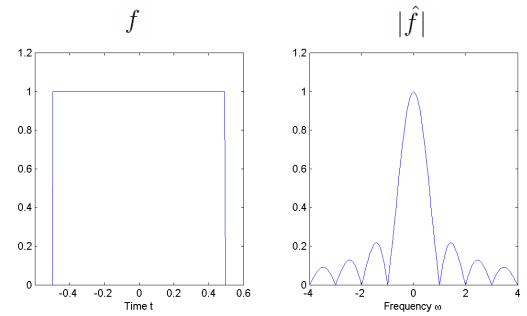
Meinard Müller
Summer Term 2008

Audio Features



Window Functions

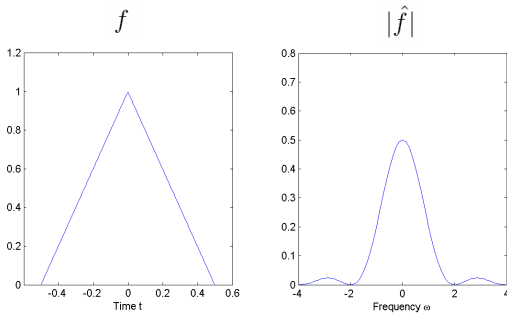
Box window



2

Window Functions

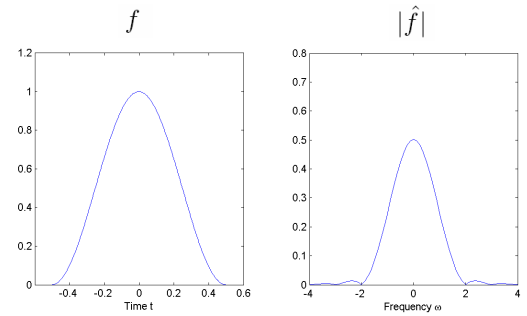
Triangle window



3

Window Functions

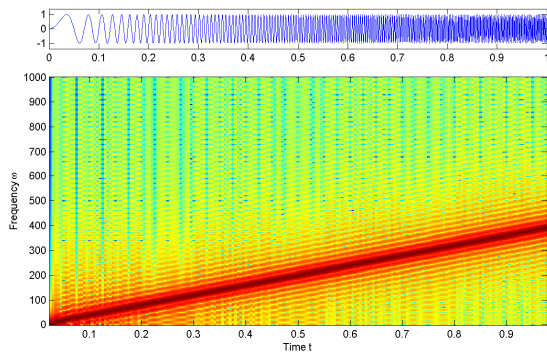
Hanning window



4

Windowed Fourier Transform (WFT)

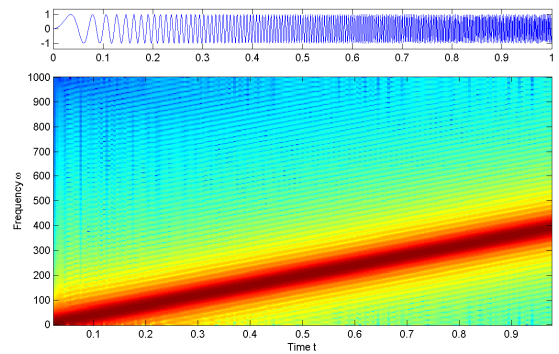
Chirp signal and WFT with box window of length 0.05



5

Windowed Fourier Transform (WFT)

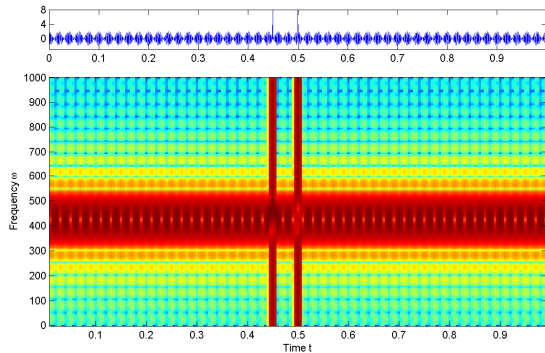
Chirp signal and WFT with hanning window of length 0.05



6

Windowed Fourier Transform (WFT)

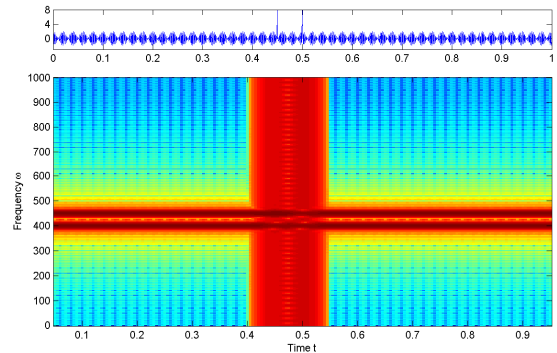
Signal and WFT with hanning window of length 0.02



7

Windowed Fourier Transform (WFT)

Signal and WFT with hanning window of length 0.1



8

Heisenberg Uncertainty Principle

Window function $g \in L^2(\mathbb{R})$ with $\|g\| = 1$

Center

Width

$$t_0 = t_0(g) := \int_{-\infty}^{\infty} t |g(t)|^2 dt \quad T(g) := \left(\int_{-\infty}^{\infty} (t - t_0)^2 |g(t)|^2 dt \right)^{\frac{1}{2}}$$

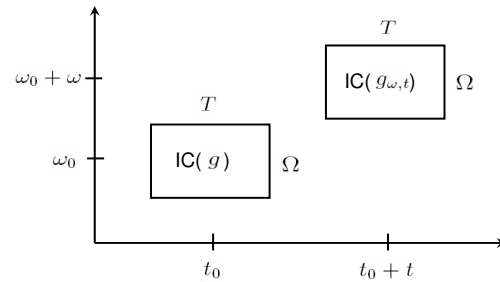
$$\omega_0 = \omega_0(g) := \int_{-\infty}^{\infty} \omega |\hat{g}(\omega)|^2 d\omega \quad \Omega(g) := \left(\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\hat{g}(\omega)|^2 d\omega \right)^{\frac{1}{2}}$$

$$T(g) \cdot \Omega(g) \geq \frac{1}{4\pi}$$

9

Information Cells

$$g_{\omega,t}(u) := e^{2\pi i \omega u} g(u - t) \quad \text{"musical note"}$$



10

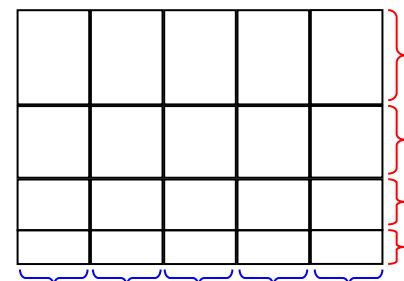
Pitch Features

Note	MIDI pitch	Center frequency	Left boundary	Right boundary	Width
A3	57	220.0	213.7	226.4	12.7
A#3	58	233.1	226.4	239.9	13.5
B3	59	246.9	239.9	254.2	14.3
C4	60	261.6	254.2	269.3	15.1
C#4	61	277.2	269.3	285.3	16.0
D4	62	293.7	285.3	302.3	17.0
D#4	63	311.1	302.3	320.2	18.0
E4	64	329.6	320.2	339.3	19.0
F4	65	349.2	339.3	359.5	20.2
F#4	66	370.0	359.5	380.8	21.4
G4	67	392.0	380.8	403.5	22.6
G#4	68	415.3	403.5	427.5	24.0
A4	69	440.0	427.5	452.9	25.4

11

Pitch Features

Time-frequency representation



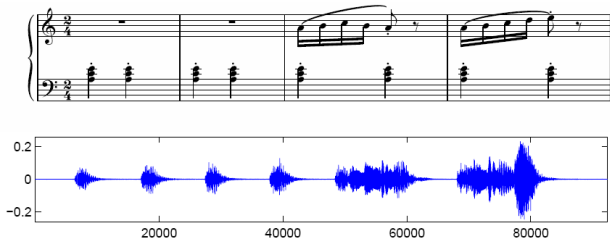
Windowing in the time domain

Windowing in the frequency domain

12

Pitch Features

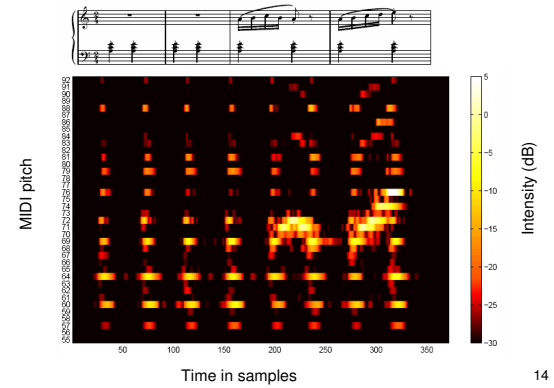
Op. 100, No. 2 by Friedrich Burgmüller



13

Pitch Features

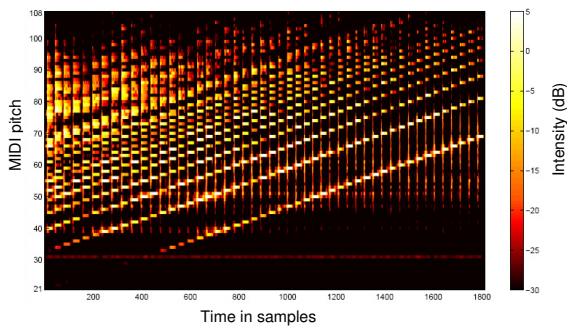
Time-pitch plot



14

Pitch Features

Time-pitch plot

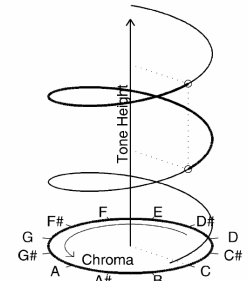
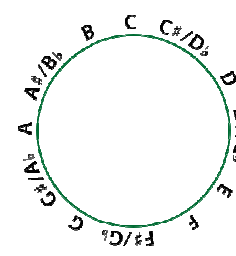


15

Chroma Features

Chromatic circle

Shepard's helix of pitch perception



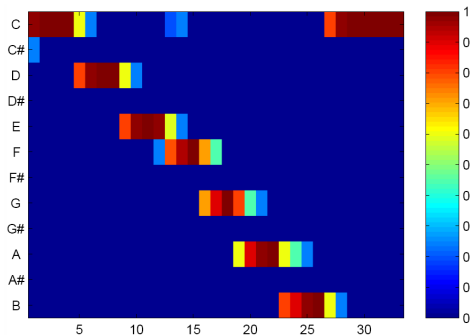
http://en.wikipedia.org/wiki/Pitch_class_space

Bartsch/Wakefield, IEEE Trans. Multimedia, 2005

16

Chroma Features

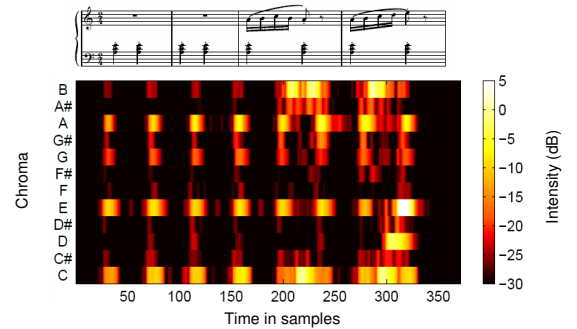
Example: C-Major Scale ▶ ▶



17

Chroma Features

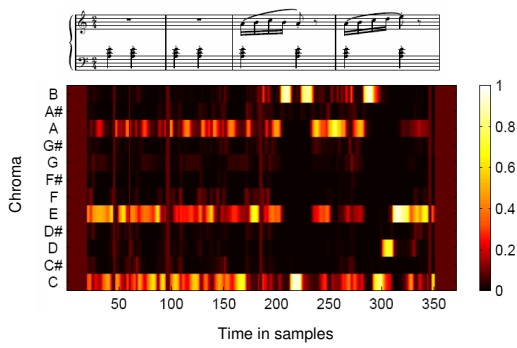
Example: Burgmüller Op. 100, No. 2



18

Chroma Features

Normalization



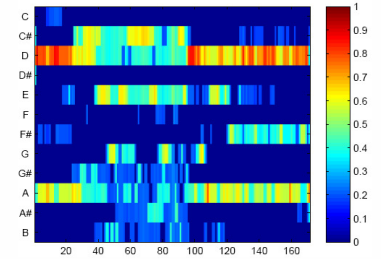
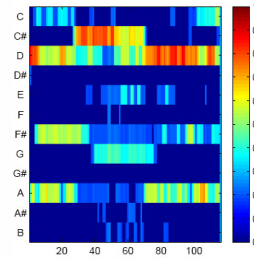
19

Chroma Features

Example: Bach Toccata

Koopman

Ruebsam



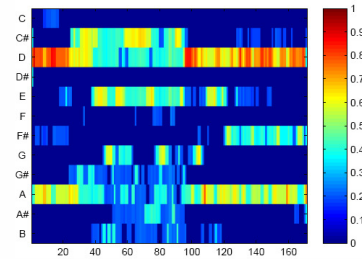
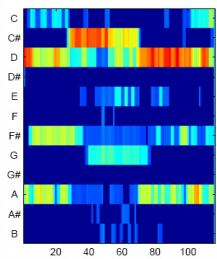
20

Chroma Features

Example: Bach Toccata

Koopman

Ruebsam



Feature resolution: 10 Hz

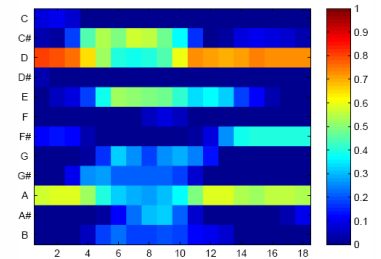
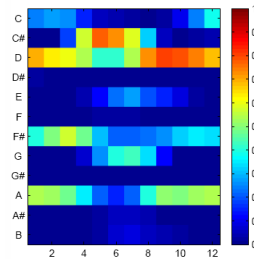
21

Chroma Features

Example: Bach Toccata

Koopman

Ruebsam



Feature resolution: 1 Hz

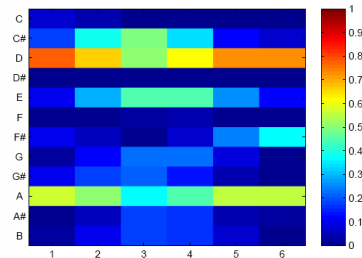
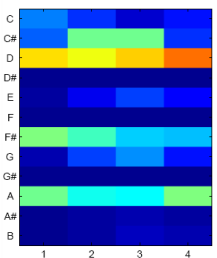
22

Chroma Features

Example: Bach Toccata

Koopman

Ruebsam



Feature resolution: 0.33 Hz

23

Chroma Features

WAV Chroma (10 Hz) GENS (1 Hz)

???

???

???

24

Chroma Features

	WAV	Chroma (10 Hz)	CENS (1 Hz)
Beethoven's Fifth (Bernstein)	▶	▶	▶
???	▶	▶	▶
???	▶	▶	▶

25

Chroma Features

	WAV	Chroma (10 Hz)	CENS (1 Hz)
Beethoven's Fifth (Bernstein)	▶	▶	▶
Beethoven's Fifth (Piano/Sherbakov)	▶	▶	▶
???	▶	▶	▶

26

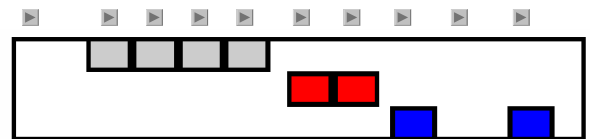
Chroma Features

	WAV	Chroma (10 Hz)	CENS (1 Hz)
Beethoven's Fifth (Bernstein)	▶	▶	▶
Beethoven's Fifth (Piano/Sherbakov)	▶	▶	▶
Brahms Hungarian Dance No. 5	▶	▶	▶

27

Chroma Features

Example: Zager & Evans "In The Year 2525"

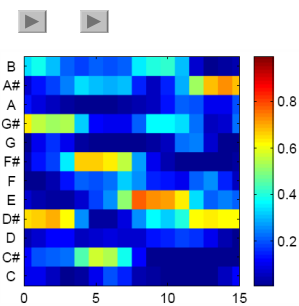


How to deal with transpositions?

28

Chroma Features

Example: Zager & Evans "In The Year 2525"

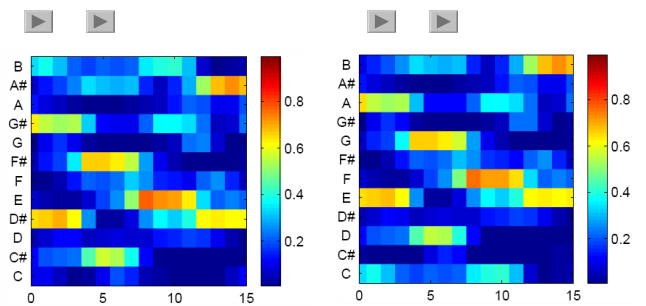


Original: (v^1, \dots, v^N)

29

Chroma Features

Example: Zager & Evans "In The Year 2525"



Original: (v^1, \dots, v^N)

Shifted: $(\sigma(v^1), \dots, \sigma(v^N))$

30