Introduction to Audio Content Analysis

Module 7.1: Human Perception of Pitch

alexander lerch

Georgia Tech Center for Music Technology
College of Design

introduction

overview

Georgia Tech Center for Music Technology College of Design

corresponding textbook section

Section 7.1

- **lecture content**
  - pitch as perceptual phenomenon
  - non-linear relation of frequency and pitch
  - frequency content of a simple pitched sound
  - dimensions of pitch perception

- **learning objectives**
  - describe basic properties of models for pitch
  - explain the two dimensions of pitch perception

introduction
overview

corresponding textbook section

Section 7.1

- **lecture content**
  - pitch as perceptual phenomenon
  - non-linear relation of frequency and pitch
  - frequency content of a simple pitched sound
  - dimensions of pitch perception

- **learning objectives**
  - describe basic properties of models for pitch
  - explain the two dimensions of pitch perception

tonal analysis
introduction

- **pitch & pitch-based properties** belong to the most important parameters describing music
  - melody
  - harmony
  - tonality
  - tuning & intonation

- **related ACA tasks**
  - fundamental frequency detection
  - key detection
  - chord detection
  - tuning frequency & temperament estimation

tonal analysis
introduction

- **pitch & pitch-based properties** belong to the most important parameters describing music
  - melody
  - harmony
  - tonality
  - tuning & intonation

- **related ACA tasks**
  - fundamental frequency detection
  - key detection
  - chord detection
  - tuning frequency & temperament estimation

pitch perception
pitch definition

definition (American Standards Association)

pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale[1]

- temporal variations in pitch give rise to a sense of melody
- closely related to frequency, but **subjective**

$\Rightarrow$ assigning a pitch value to a sound means **specifying the frequency of a pure tone having the same subjective pitch** as the sound

[1] ASA, "Acoustical Terminology," American Standards Association (ASA), Standard, 1960.

pitch perception
pitch definition

definition (American Standards Association)

pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale[1]

- temporal variations in pitch give rise to a sense of melody
- closely related to frequency, but **subjective**

$\Rightarrow$ assigning a pitch value to a sound means **specifying the frequency of a pure tone having the same subjective pitch** as the sound

[1] ASA, "Acoustical Terminology," American Standards Association (ASA), Standard, 1960.

definition (American Standards Association)

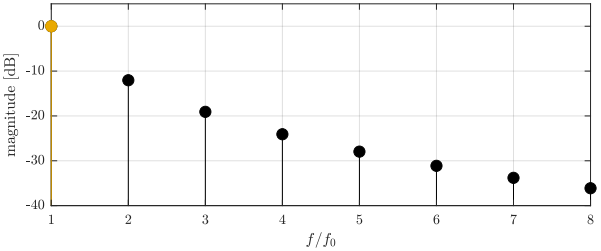
pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale[1]

- temporal variations in pitch give rise to a sense of melody
- closely related to frequency, but **subjective**

$\Rightarrow$ assigning a pitch value to a sound means **specifying the frequency of a pure tone having the same subjective pitch** as the sound

[1] ASA, "Acoustical Terminology," American Standards Association (ASA), Standard, 1960.

pitch perception
fundamental frequency



- *fundamental* frequency is relevant for pitch perception ($f_0, 2f_0, 3f_0, \ldots$)
- higher fundamental frequency $\Rightarrow$ higher pitch (mono-dimensional)

matlab source: plotHarmonics.m

pitch perception
missing fundamental

- basilar membrane location **does not explain** the pitch perception of complex tones

$\Rightarrow$ **virtual pitch**, **residue pitch**

- **example 1**: missing fundamental
  - $f_0 = 120\,\mathrm{Hz}$, 33 harmonics, with(out) bandpass 300-2400 Hz
- **example 2**: missing fundamental
  - speech $f_0 \approx 100\,\mathrm{Hz}$, with(out) bandpass 300-4000 Hz

pitch perception
missing fundamental

- basilar membrane location **does not explain** the pitch perception of complex tones

$\Rightarrow$ **virtual pitch**, **residue pitch**

- **example 1**: missing fundamental
  - $f_0 = 120\,\mathrm{Hz}$, 33 harmonics, with(out) bandpass 300-2400 Hz
- **example 2**: missing fundamental
  - speech $f_0 \approx 100\,\mathrm{Hz}$, with(out) bandpass 300-4000 Hz

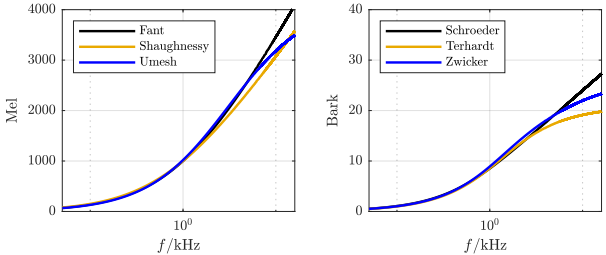
pitch perception
missing fundamental

- basilar membrane location **does not explain** the pitch perception of complex tones

$\Rightarrow$ **virtual pitch**, **residue pitch**

- **example 1**: missing fundamental
  - $f_0 = 120\,\mathrm{Hz}$, 33 harmonics, with(out) bandpass 300-2400 Hz
- **example 2**: missing fundamental
  - speech $f_0 \approx 100\,\mathrm{Hz}$, with(out) bandpass 300-4000 Hz

pitch perception
frequency & pitch



non-linear pitch frequency relation:

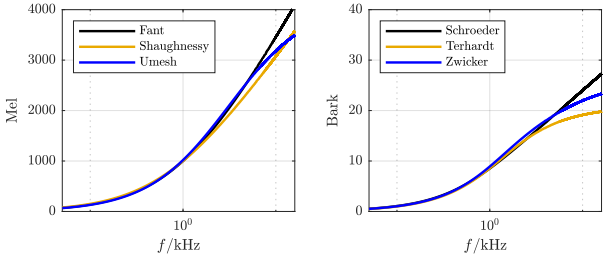
- perceptual pitch distance $\neq$ frequency distance

$\Rightarrow$ *models* for psycho-acoustic/physiological data

- *Mel* scale (equal pitch distance)
- *Bark* scale (critical band width)
- physiological frequency location (basilar membrane)

matlab source: plotMelBark.m

pitch perception
frequency & pitch



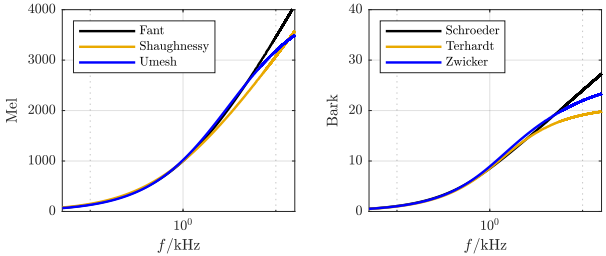
Fant: $\mathfrak{m}_{\mathrm{F}}(f) = 1000 \cdot \log_2\left(1 + \frac{f}{1000\,\mathrm{Hz}}\right)$

O'Shaughnessy: $\mathfrak{m}_{\mathrm{S}}(f) = 2595 \cdot \log_{10}\left(1 + \frac{f}{700\,\mathrm{Hz}}\right)$

$\mathfrak{m}_{\mathrm{S}}(f) = 1127 \cdot \log\left(1 + \frac{f}{700\,\mathrm{Hz}}\right)$

matlab source: plotMelBark.m

pitch perception
frequency & pitch



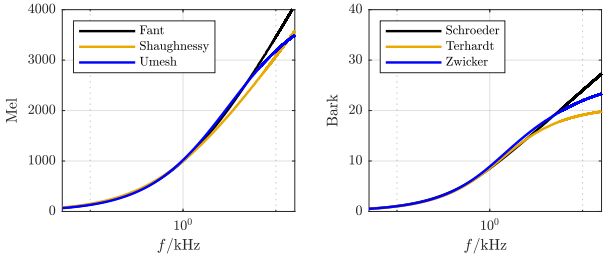
Fant: $\mathfrak{m}_{\mathrm{F}}(f) = 1000 \cdot \log_2\left(1 + \frac{f}{1000\,\mathrm{Hz}}\right)$

O'Shaughnessy: $\mathfrak{m}_{\mathrm{S}}(f) = 2595 \cdot \log_{10}\left(1 + \frac{f}{700\,\mathrm{Hz}}\right)$

$\mathfrak{m}_{\mathrm{S}}(f) = 1127 \cdot \log\left(1 + \frac{f}{700\,\mathrm{Hz}}\right)$

matlab source: plotMelBark.m

pitch perception
frequency & pitch



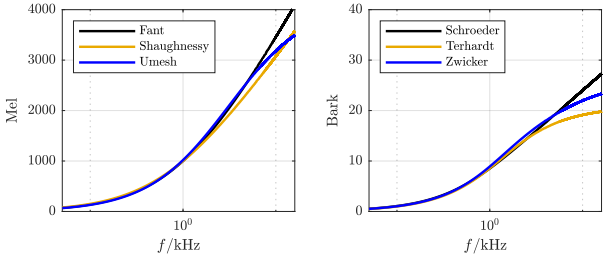
Schröder: $\mathfrak{z}_{\mathrm{S}}(f) = 7 \cdot \operatorname{arcsinh}\left(\frac{f}{650\,\mathrm{Hz}}\right)$

Terhardt: $\mathfrak{z}_{\mathrm{T}}(f) = 13.3 \cdot \arctan\left(0.75 \cdot \frac{f}{1000\,\mathrm{Hz}}\right)$

Zwicker: $\mathfrak{z}_{Z}(f) = 13 \cdot \operatorname{atan}\left(0.76 \cdot \frac{f}{1000\,\mathrm{Hz}}\right) + 3.5 \cdot \operatorname{atan}\left(\frac{f}{7500\,\mathrm{Hz}}\right)$

matlab source: plotMelBark.m

pitch perception
frequency & pitch



ERB: $\mathfrak{e}(f) = 9.26 \log\left(1 + \frac{f}{228.7}\right)$

Cochlear Map: $\mathfrak{x}(f) = \frac{1}{2.1} \log_{10}\left(\frac{f}{165.4} + 1\right)$

matlab source: plotMelBark.m

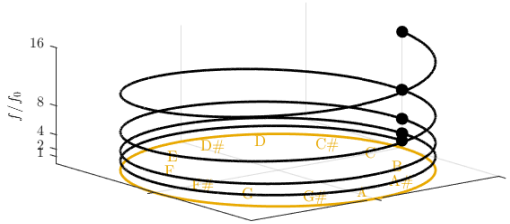
2 dimensions of musical pitch

- **tone height**: monotonic relationship to frequency (increasing frequency ⇒ increasing pitch)
- **tone chroma**: two tones separated by octave sound similar (same *pitch class*)

pitch perception
pitch dimensions

2 dimensions of musical pitch

- **tone height**: monotonic relationship to frequency (increasing frequency ⇒ increasing pitch)
- **tone chroma**: two tones separated by octave sound similar (same *pitch class*)



matlab source: plotPitchHelix.m

summary
lecture content

- **pitch**
  - **subjective** phenomenon
  - **non-linear** monotonic relationship to frequency (tone height increases with fundamental frequency)
  - pitch grouping based on powers of two: tone **chroma perception**