Pitch and Chroma

Hélène Papadopoulos

- The tonal content is an important aspect of Western tonal music.
  - Notes, melody.
  - Chords.
- Goal of this lecture:
  1. Describe what the pitch of a sound is.
  2. Chroma: a compact representation of the tonal content of the music signal.

What is the pitch?

- When an instrument produces a sound, the human listener *perceives* a *pitch*.
- Pitch represents the perceived *fundamental frequency* of a sound.
- The pitch is a *subjective quality* of the sound often described as *highness* or *lowness*.
- It is one of the four major auditory attributes of sounds along with loudness, timbre and sound source location.
- In music, the term *note* is used to refer to the relative duration and pitch of a given sound.

We perceive a single pitch but…

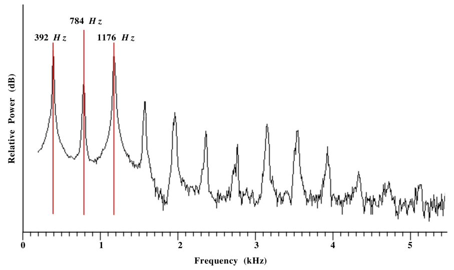
- The sound produced by any musical instrument is far more complex than the single frequency of the desired note.
- When a note is played on a musical instrument, there are always additional tones produced at the same time.
  - The frequency of the desired note is known as the *fundamental frequency (f0)*.
  - The frequencies of the additional tones are known as *overtones*.
- In some instruments (such as gongs, or tam-tams), the overtones are not related to the fundamental frequency in a simple way.
- But in other instruments (like stringed and wind instruments, as well as some percussion instruments such as the marimba, the vibraphone) the overtones are related to the fundamental frequency "*harmonically*".

What is the relation between the harmonics and the fundamental frequency?

- *Harmonics* are overtones that happen to be *simple integer multiples* of the fundamental frequency: $k * f_0$.
  - Example: If a string is plucked and produces a frequency of, say, 110 Hz,
  - multiples of that 110 Hz will also occur at the same time: 220 Hz, 330 Hz, 440 Hz, etc… will all be present, although not all with the same intensity.
- Remark: In general, the harmonics of music sounds do not have frequencies that are exactly multiples of its f_0 . For this reason, they are often called *partials*.
- A musical instrument's fundamental frequency and all of its overtones combine to produce that instrument's *sound spectrum*.

Example of sound spectrum:

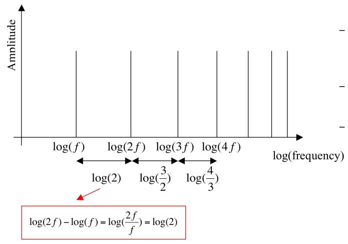
- Sound spectrum from a flute playing the note G4.



- The first peak is the desired G4 pitch (frequency of 392 Hz).
- The second and third peaks are also identified with vertical lines and have frequencies of about 784 = 392*2 Hz and 1,176 = 392*3 Hz.

The harmonics form a “pattern” in the frequency domain:

- Consider a harmonic sound made up of components with equal amplitude frequency f, 2f, 3f,...
- We plot the positions of these frequency components against log frequency:



- The *absolute positions* of the frequency components depend on the frequency of the fundamental,
- But the *relative positions* to each other are constant.
- Perception: difference, Physics: ratio

- They form a *"pattern"* in the frequency domain that is the same for all sounds with harmonic frequency components.

Timbre and sound spectrum

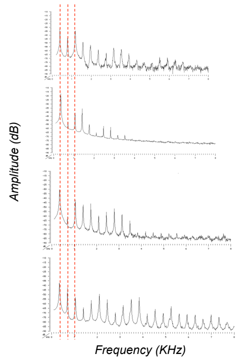
- It is the difference in intensities of the various overtones produced that gives each instrument a characteristic sound quality or *timbre*.

G4 - Modern flute

G4 - Clarinet forte

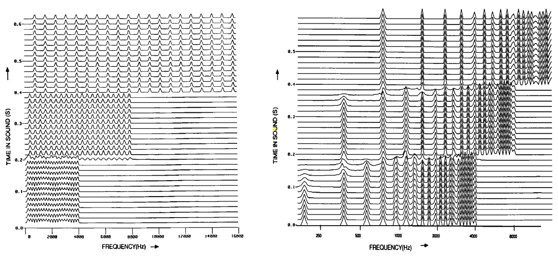
G4 - Clarinet mezzo piano

G4 - Saxophone



Time-frequency representations adapted to music signals:

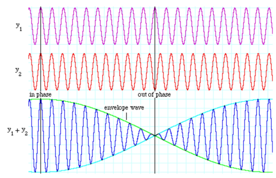
- Constant-Q transform: logarithmic scale.
- Better representation of spectral data from a music signal.



Discrete Fourier Transform (left) and Constant-Q transform (right) of three complex sounds with fundamentals G3 (196 Hz), G4 (392 Hz), and G5 (784 Hz), and each having 20 harmonics with equal amplitude [Brown1991].

Pitch is a perceptual attribute of sound: illustration.

- The frequency of the perceived pitch increases with loudness.
- Interferences beats: consider two tones with similar frequencies, f_1 and f_2, and same amplitude A, sounded in the same space.



$$y_1 = A\cos(2\pi f_1 t)$$

$$y_2 = A\cos(2\pi f_2 t)$$

$$y_1 + y_2 = 2A\left(\cos\left(2\pi \frac{f_1 - f_2}{2} t\right)\right) * \left(\cos\left(2\pi \frac{f_1 + f_2}{2} t\right)\right)$$

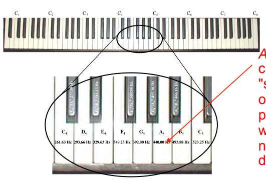
beat frequency $f_1 - f_2$ — perceived frequency

400+400.5 Hz	400 + 401 Hz	400 + 403 Hz	400 + 410 Hz	400 + 420 Hz
1beat/2s	1beat/s	3beats/s	10beats/s	Interference?

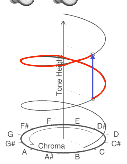
From pitch to notes: the Western tonal music equal-tempered scale

- In music, notes are governed by structural principles. A scale is a series of notes arranged in ascending or descending order.
- Notes having a frequency ratio of 2 correspond to an *octave*. They are perceived as very similar (--> common harmonics).

G3, G4, G5, clarinet



A4 = 440 Hz: current "standard pitch" or "concert pitch" around which the other notes are defined.

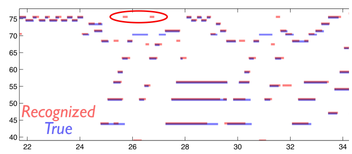


Our perception of pitch is two-dimensional and can be modeled by a helix [Shepard1960]

- In the equal-tempered scale, the octave is divided into 12 notes. The *ratio* of the frequencies of any two adjacent notes is the same : $\frac{f_{k+1}}{f_k} = 2^{1/12}$
- Other scales exist (Pythagorean, just intonation etc.)

Pitch tracking: a challenging task

- Many applications of pitch tracking in speech and audio music analysis:
  - Pitch is largely responsible for inflections in human speech --> Speaker identification.
  - Music transcription.
  - Instrument identification.



- There are many challenging issues in pitch tracking that include:
  - Missing/unclear fundamental frequency: The fundamental frequency that the human ear interprets as the pitch of a sound may be absent in the sound.
  - Noise, set of overtones.
  - Problem of the detection of notes onsets.
- Monophonic pitch tracking is difficult but *polyphonic pitch tracking* is even more complex!
  - Overlap between the harmonics of different pitches (--> octaves).
- The topic of pitch tracking has been well explored, but there is still room for *lots of improvement*.

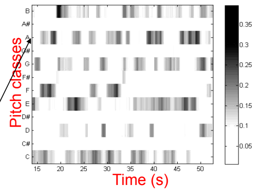
2. Chroma features

What about representing the tonal content of a music signal?

- Pitch tracking is a very popular task in MIR but polyphonic transcription algorithms from an audio signal are still limited and time-consuming.
- Chroma representation, an alternative to transcription.
  - [Fujishima 1999]: Pitch Class Profiles.
  - [Wakefield 1999]: Chroma-based representation.

Chroma vector: Mapping between the spectrum and a 12-dimensional vector representing the intensity of the 12 semitone pitch classes.

Freq. (Hz)	55	110	220	440	880	1760
Note	A1	A2	A3	A4	A5	A6



Chromagram of an extract from *Misery (Please Please Me), Beatles.*

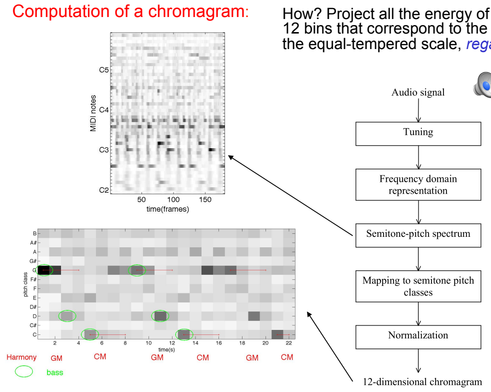
- Applications: estimation of chords and musical keys from audio recordings, audio similarity retrieval tasks.

2. Chroma features

Computation of a chromagram:

How? Project all the energy of the spectrum onto 12 bins that correspond to the 12 pitch classes of the equal-tempered scale, *regardless of octave*.

Love Me Do (Beatles)



Standard reference frequency A4 = 440Hz.

In the mapping, not all the frequencies of the spectrum are considered --> restriction to a frequency region $[f_{min}, f_{max}]$ that is the most relevant for pitch distribution.

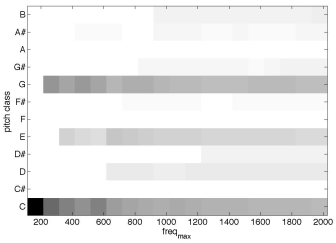
Provides robustness against variations of dynamics.

Chroma features capture both melodic and harmonic accompaniment information.

2. Chroma features

Chroma features do not only grasp the notes played by the instruments:

- The chroma vector of a note played by an instrument contains:
  - The pitch classes corresponding to the fundamental frequency f0 of the perceived pitch (ignoring octave considerations) ,
  - But also includes a mixture of its harmonics.

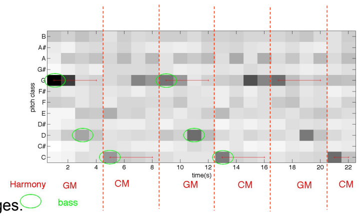


- Chroma feature of a cello C1 note (65.4 Hz) considering various frequency intervals (from f_{min} = 60Hz to various values of f_{max}) for computation.
--> See the harmonics of the C: C-C-G-C-E-G etc.

2. Chroma features

Beat-synchronous chroma features:

Compute one chroma feature per beat:



- Provides invariance to tempo changes.
- The chromagram looks like chords…
- Useful for many tasks:
  - Music similarity and cover song identification when comparing the chord progression of two songs, possibly at different tempi, score alignment etc.
  - All approaches that combine harmonic and metrical information and need to work with features related to the meter.
- Next lecture: how can we transcribe chords?

Summary

- Pitch: very important perceptual attribute of sound, in speech and music.
- Chroma: compact and robust representation of the tonal content of music sounds.
- Next lecture: how can we transcribe chords?