Introduction to **Audio Content Analysis**

module 3.3.3: time-frequency representations — (Log) Mel-spectrogram

alexander lerch

introduction
overview

Georgia | Center for Music
Tech | Technology
College of Design

corresponding textbook section

section 3.3.3

- **lecture content**
  - log mel spectrogram

- **learning objectives**
  - discussing advantages and disadvantages of different time-frequency transforms
  - explaining the computation of the mel spectrogram

introduction
overview

Georgia | Center for Music
Tech | Technology
College of Design

corresponding textbook section

section 3.3.3

- **lecture content**
  - log mel spectrogram

- **learning objectives**
  - discussing advantages and disadvantages of different time-frequency transforms
  - explaining the computation of the mel spectrogram

mel spectrogram
introduction

- the log mel spectrogram is a normal spectrogram with a logarithmic frequency axis

- more musically and perceptually meaningful
- usually less frequency bins
- often used as audio input representation to neural networks

mel spectrogram
computation

Georgia | Center for Music
Tech | Technology
College of Design

1 compute STFT

2 compute magnitude

3 aggregate frequency bins into bands with increasing width (logarithmic frequency)

4 optional: compute the aggregated magnitudes to dB

mel spectrogram
computation

Georgia | Center for Music
Tech | Technology
College of Design

1 compute STFT

2 compute magnitude

3 aggregate frequency bins into bands with increasing width (logarithmic frequency)

4 optional: compute the aggregated magnitudes to dB

mel spectrogram
computation

Georgia | Center for Music
Tech | Technology
College of Design

1. compute STFT

2. compute magnitude

3. aggregate frequency bins into bands with increasing width (logarithmic frequency)

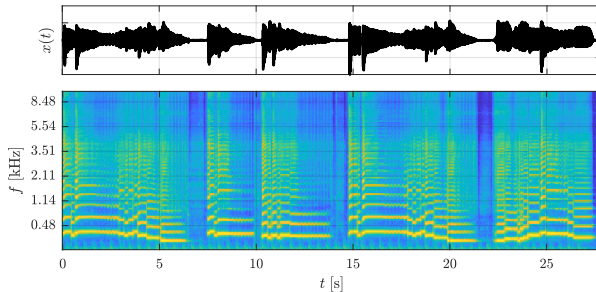4. optional: compute the aggregated magnitudes to dB

mel spectrogram
computation

Georgia | Center for Music
Tech | Technology
College of Design

1 compute STFT

2 compute magnitude

3 aggregate frequency bins into bands with increasing width (logarithmic frequency)

4 optional: compute the aggregated magnitudes to dB

mel spectrogram
example

Georgia | Center for Music
Tech || Technology
College of Design

matlab source: plotLogMelSpecgram.m

summary
lecture content

Georgia | Center for Music
Tech | Technology
College of Design

- **Mel spectrogram**
  - logarithmic frequency axis (musically and perceptually more meaningful)
  - logarithmic magnitude (perceptually more meaningful)
  - usually a more reduced representation than STFT