

Introduction to Audio Content Analysis

Module 3.4.3: Time-Frequency Representations — (Log) Mel-Spectrogram

alexander lerch

introduction

overview

corresponding textbook section

Section 3.4.3

■ 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

corresponding textbook section

Section 3.4.3

■ 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

- 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

- 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

- 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

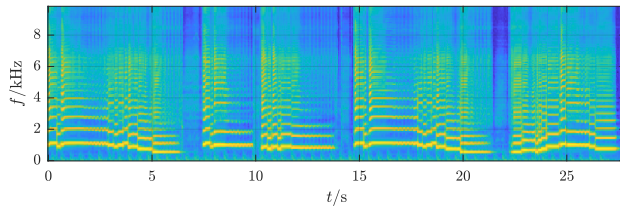
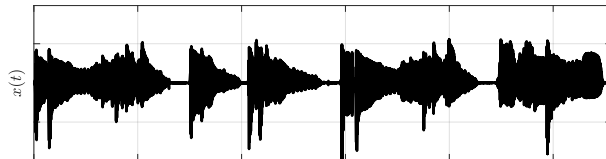
mel spectrogram

computation

- 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



summary

lecture content

■ Mel spectrogram

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

