

Alberto Andrés Valdés González.

Degree: Mathematical Engineer.

Work position: ML-Engineer.

Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com

Location: Santiago, Chile.

MFCC

The mel frequency cepstral coefficients (MFCCs) of an audio signal are a small set of features (usually about 10–20) which describe the overall shape of the spectral envelope. MFCCs were frequently used for voice recognition (see for instance [Muda et al., 2010]) before the task started to be performed using deep learning (see for instance [Yu and Deng, 2016]). It was also used to describe “timbre”.

The two dimensions we use is **Time** and **Amplitude**.

The extraction of MFCCs from an audio input is summarised as follows. A log power spectrum is evaluated from the audio input. The power spectrum is resampled on the mel scale. The MFCC output is the Discrete Cosine Transform of the resampled spectrum. A significant dimensionality reduction comes from the resampling to the 16-band mel filter bank.

