MFCC-Plotter
— A graphical tool to interactively analyse MFCCs —

Frederick Kukla & Vanessa Reichel

Institute for Phonetics and Speech Processing, LMU Munich

MFCCs are multidimensional and complex numeric data which are hard for humans to comprehend. Even though they are the most popular feature to use for automatic speech recognition and speaker recognition [1, 2], the existing visualisation methods are abstract and non intuitive [3]. Therefore, we developed a visualisation tool. It provides interactive ways to explore and analyse MFCC data. Users can choose between various MFCC calculation options such as *Praat* or the python-library *librosa*. The *MFCC-Plotter* provides two different normalisation techniques: either speaker or phoneme normalisation. Furthermore, users can select a phoneme and the Plotter will display its corresponding MFCCs.

Figure 1 shows the developed tool *MFCC-Plotter*. The default visualisation is a combination of a radar plot and a boxplot-like structure. The plot has 13 sections, each referring to one MFCC dimension. Besides the default display, three other statistical visualisations are available. These include a histogram, an ECDF-plot as well as a boxplot.

The *MFCC-Plotter* can help students to get a better understanding of MFCCs in general, distributions of multidimensional data and normalisation methods. It could be used in teaching as well as in self-studying contexts. The tool can also be deployed to get an overview of the data or for troubleshooting.

Figure 1: *MFCC-Plotter*; left side: 13 MFCCs of /E/ calculated with *Praat*, phoneme normalised; right side: 13 MFCCs of /E/ calculated with *Praat*, speaker normalised

References

[1] Nidhi Desai, Kinnal Dhameliya, and Vijayendra Desai. Feature extraction and classification techniques for speech recognition: A review. *International Journal of Emerging Technology and Advanced Engineering*, 3(12):367–371, 2013.

[2] Tomi Kinnunen and Haizhou Li. An overview of text-independent speaker recognition: From features to supervectors. *Speech communication*, 52(1):12–40, 2010.

[3] OpenGenus IQ. Mfcc (mel frequency cepstral coefficients) for audio format, 2022.