List of neural signal-level features computed

This package computes **neural micro features** called Temporal Acoustic Parmaters (TAP). These features are similar to, but not the same their signal-processing equivalents described below. The neural TAP features are more stable and robust to random perturbances.

For this reason, the descriptions given below apply *only roughly* to the Temporal Acoustic Parmaters (TAP) in this package. The names used for the features have been chosen to be the same as their closest signal-processing based equivalents, to facilitate cross-referencing, and as mnemonics.

1. **Loudness** : Volume or intensity of the audio file
2. **alphaRatio** : Ratio of the summed energy from 50-1000 Hz and 1-5 kHz, i.e., a measure of the relative energy in the two frequency bands 50-1000 Hz and 1-5 kHz.
3. **hammarbergIndex** : Ratio of the strongest energy peak in the 0-2 kHz region to the strongest peak in the 2-5 kHz region. This is a measure of the spectral balance of voice.

4. **slope0-500** : Linear regression slope of the logarithmic power spectrum within the frequency band 0-500 Hz (i.e. within the lower frequency range of 0-500 Hz). Spectral slope is a measure of the spectral distribution of energy in a sound.

5. **slope500-1500** : Linear regression slope of the logarithmic power spectrum within the frequency band 500-1500 Hz (i.e. within the mid frequency range of 500-1500 Hz). This is also a spectral slope like the above.

6. **spectralFlux** : Spectral flux is a measure of the change in the spectral content of a sound over time. Measured as the difference of the spectra of two consecutive frames.

7. **mfcc1** : Dimensions 1-4 of Mel-Frequency Cepstral Coefficients (MFCCs) are related to the spectral shape of the speech signal. Note that 1 is the first dimension, the counting begins from 0. Dimension 0 represents the energy in the signal (and is proportional to the loudness).

   Dimensions 1-4 of MFCCs represent the modulation of the spectral envelope across different frequency bands. Higher values indicate greater modulation in the spectral envelope. mfcc1 represents the spectral shape of the sound in the low-frequency range. Contains information about resonant frequencies of the vocal tract (formants).

8. **mfcc2** : Represents the spectral shape in the mid-frequency range. Influenced by vocal fold vibration.

9. **mfcc3** : Represents the spectral shape in the high-frequency range, which is influenced by the characteristics of the vocal tract's constrictions.

10. **mfcc4** : Represents the spectral tilt, which reflects the overall balance between high and low-frequency energy in the sound. Thus this has information about the overall spectral envelope of the sound.

11. **F0semitoneFrom27.5Hz** : F0 is the fundamental frequency or the first harmonic of a sound wave (or pitch in the case of voice). This entity is actually the pitch, which is expressed as the logarithmic F0 on a semitone frequency scale, starting at 27.5 Hz (semitone 0).

    Each semitone represents a frequency ratio of approximately 1.06. More precisely, on the equal-tempered Western musical scale, each semitone interval is equal to a frequency ratio of the 12th root of 2 (approximately 1.0595). To find the frequency of the F0 semitone, we need to multiply the frequency of the F0 by the ratio of the semitone interval. The F0 semitone is one semitone higher than the F0, so we need to multiply 27.5 Hz by the 12th root of 2.

    Calculating this, we get:

    $F0\ semitone = 27.5\ Hz\ \times (2^{\frac{1}{12}}) \approx 29.14\ Hz$

Thus the frequency of the first F0 semitone from 27.5 Hz is approximately 29.14 Hz.

12. **jitterLocal** : Also known as pitch variation or frequency perturbation. It is a measure of the cycle-to-cycle variations in the fundamental frequency (F0) of a sound. Local jitter refers to the short-term variations in F0 that occur within a single pitch period. Measured as changes in consecutive F0 values.

13. **shimmerLocaldB** : Cycle-to-cycle variability in the amplitude of F0.

14. **HNRdBACF** : HNR dB ACF stands for Harmonic-to-Noise Ratio in decibels (dB) using the Autocorrelation Function (ACF) method.

    It is the ratio between the harmonic components and noise components of the signal. This HNR dB ACF value is reported in decibels. Higher values will indicate more harmonic regularity and lower levels of noise. However, note that different methods of calculating the HNR may yield different results, so this cannot be mixed with other HNR analyses.

15. **logRelF0-H1-H2** : Harmonic difference H1–H2. Ratio of energy of the first F0 harmonic (H1) to the energy of the second F0 harmonic (H2) in the log domain.

16. **logRelF0-H1-A3** : Harmonic difference H1–A3. Ratio of energy of the first F0 harmonic (H1) to the energy of the highest harmonic in the third formant range (A3). Is also a measure of spectral tilt of the signal. For voice signals, it reflects the distribution of energy across the harmonics of the voice.

17. **F1frequency** : Frequency of F1, the first formant, which is the lowest resonant frequency of the vocal tract.

18. **F1bandwidth** : The bandwidth of the first formant (F1 bandwidth). Refers to the range of frequencies that are present in the resonance of the vocal tract at the F1 frequency.

19. **F1amplitudeLogRelF0** : The log of the amplitude of the first formant relative to F0. Refers to a measurement of the relative strength of the first formant (F1) to the fundamental frequency (F0).

20. **F2frequency** : Frequency of F2, the second formant, which is the second resonant frequency of the vocal tract.

21. **F2bandwidth** : The bandwidth of the second formant (F2 bandwidth). Refers to the range of frequencies that are present in the resonance of the vocal tract at the F2 frequency.

22. **F2amplitudeLogRelF0** : The log of the amplitude of the second formant relative to F0. Refers to a measurement of the relative strength of the second formant (F2) to the fundamental frequency (F0).

23. **F3frequency** : Frequency of F3, the third formant, which is the third resonant frequency of the vocal tract.

24. **F3bandwidth** : The bandwidth of the third formant (F3 bandwidth). Refers to the range of frequencies that are present in the resonance of the vocal tract at the F3 frequency.

25. **F3amplitudeLogRelF0** : The log of the amplitude of the first formant relative to F0. Refers to a measurement of the relative strength of the third formant (F3) to the fundamental frequency (F0).