Dokumentation SMP Projekt

Peter Wienken, Aaron Weis

19. Januar 2026

Inhaltsverzeichnis

1 Audio Import

2 FFT Analyse

3 FLUX Detection

3.1 Calculation of Logarithmic Magnitude Spectrum

The logarithmic magnitude spectrum is calculated using the formula:

$$L(k) = \log_{10}(|X(k)|) \tag{1}$$

It transforms the linear magnitude spectrum into a logarithmic scale, which is more aligned with human auditory perception.

3.2 Calculation of Spectral Flux

The spectral flux is calculated as the difference between the logarithmic magnitude spectra of consecutive frames:

$$F(n) = \sum_{k=0}^{N-1} L_n(k) - L_{n-1}(k) \tag{2}$$

For each frame n, the spectral flux $F(n)$ is computed by summing the squared differences of the logarithmic magnitude spectra $L_n(k)$ and $L_{n-1}(k)$ across all frequency bins k. The frames are typically stored in a 2D array where each column represents a frame and each row represents a frequency bin. The difference between consecutive frames can be represented as:

$$\Delta L(k) = L_n(k) - L_{n-1}(k)$$

Example of a 2D array representing frames and frequency bins:

$$\mathbf{L} = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 3 & 6 \\ 1 & 2 & 3 & 6 \end{bmatrix}$$

The resulting difference between consecutive frames is:

$$\Delta\mathbf{L} = \begin{bmatrix} 2 & 3 & 6 \\ 2 & 3 & 6 \\ 2 & 3 & 6 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & 1 & 3 \end{bmatrix}$$

Then, the spectral flux is computed by summing these differences across all frequency bins for each frame:

$$F(n) = \sum_k |\Delta L(k)|$$

$$\mathbf{F} = \begin{bmatrix} 1+1+3 & 1+1+3 & 1+1+3 \end{bmatrix}^T = \begin{bmatrix} 5 & 5 & 5 \end{bmatrix}^T$$

The Code to calculate the spectral flux from the differences is as follows:

```
flux = np.sum((calculated_FLUX[:, 1:] - calculated_FLUX[:, :-1]))
```

3.3 Half-Wave Rectification

Half-wave rectification is applied to the spectral flux to retain only positive values, which correspond to increases in spectral energy. This is done by setting all negative values to zero:

$$F_{rect}(n) = \max(0, F(n)) \tag{3}$$

This step ensures that only significant increases in spectral energy are considered for further analysis.

For example, given the spectral flux values:

$$\Delta L = \begin{bmatrix} -2 & 2 & -4 \\ 3 & 1 & 5 \\ -1 & 4 & -6 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ 3 & 1 & 5 \\ 0 & 4 & 0 \end{bmatrix}$$

And Summed up:

$$F = \begin{bmatrix} 0+3+0 & 2+1+4 & 0+5+0 \end{bmatrix}^T = \begin{bmatrix} 3 & 7 & 5 \end{bmatrix}^T$$

4 Peak Detection

5 Fazit

Hier steht das Fazit.