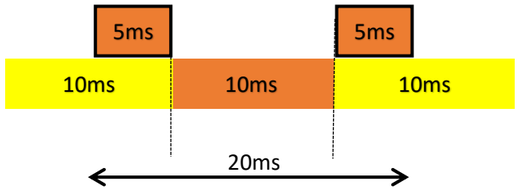
Mel Frequency Cepstral Coefficients

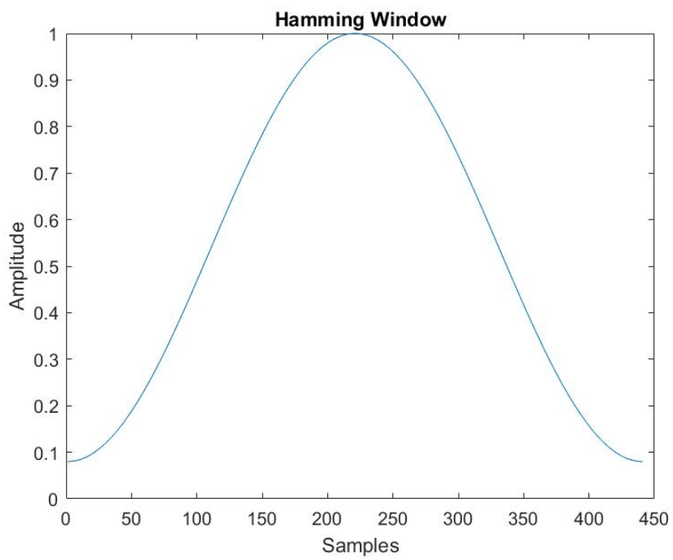
SEGMENTATION

Short sequences of samples called *frames* are extracted from the signal. In our demonstration, the frame is:



∴ Effective Length of a frame = 20ms
With 50% overlap with neighbouring frames.
The goal is to work on excerpts that are short enough to be assumed *stationary* signals.

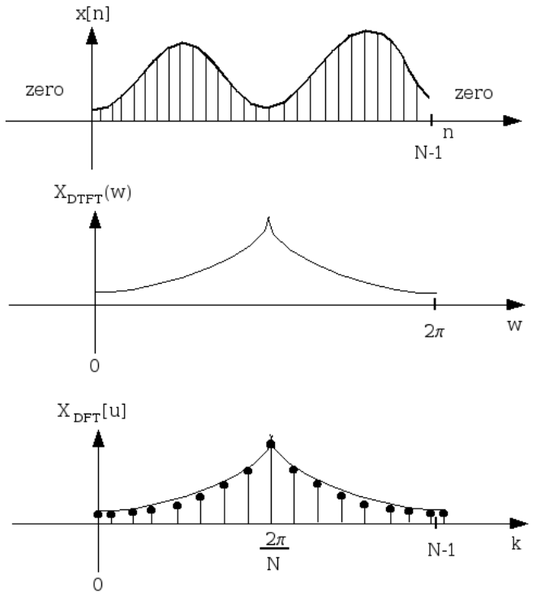
HAMMING WINDOW



Each frame is multiplied with a Hamming window to minimize spectral leakage of energy into neighbouring frequency bins. The Hamming Window is visualized as shown alongside.

DISCRETE FOURIER TRANSFORM

A DFT is applied to each frame, to obtain the distribution of energy into frequencies. Calculation of DFT of a signal is visualised as:



DFT of each frame is computed with the help of FFT technique to speed up the process.

To a certain extent, the DFT might also separate the pitch of a note from the instrument it is played with.

MAGNITUDE SPECTRUM

The DFT of the frames may be a complex numbered value, implying that it has a phase associated with it. By, taking the magnitude of the DFT we ignore the phase value and try to model with the spectral energy density.

This is a critical step as it may result in loss of some of the information.