

# Menura: Realtime pitch detection in C++

Valentina Visintini

Practical Course "Advanced Software Development with Modern C++"

Summer Term 2018

Institute for Computer Science

Ludwig-Maximilians-Universität München

## 1 Introduction

A typical task of digital signal processing is to identify the pitch of an audio signal. To achieve this in realtime, C++ ...

## 2 A Brief Outline of Acoustics

In music, a *pitch* is the position of a single sound within the complete range of sound, which is perceived according to the frequency of vibration of the sound waves. Human listeners perceive higher frequencies as higher pitches, hence a lower pitch is indicative of a lower frequency. A pitch detector aims to find the fundamental frequency of an audio signal, i. e. the lowest frequency of a periodic waveform, to determine what note is being played. Since the fundamental is also perceived as the loudest, the ear identifies it as the specific pitch of the musical tone

Sound waves cause a local pressure deviation from the ambient atmospheric pressure, which is known as sound pressure and perceived as *loudness*. Even though it can be ordered on a scale extending from quiet to loud, loudness cannot be measured objectively, as its perception changes individually, whereas the actual sound pressure can be measured using a microphone and expressed in decibel (dB). The sound pressure level  $L_p$  is defined by a logarithmic measure of the effective pressure of a sound  $p$  relative to the standard reference sound pressure  $p_0$  in air or water.

$$L_p = 20 * \log_{10}\left(\frac{p}{p_0}\right)\text{dB} \quad (1)$$

## 3 Frequency Spectrum Analysis

The fundamental frequency of an audio signal can be obtained by a fast Fourier transform (FFT). This algorithm samples a signal over a period of time and divides it into its frequency components, converting it from the time domain to a representation in the frequency domain.

$$\sum_{n=0}^{N-1} x_n e^{-i2\pi k \sqrt{-1}/N} \quad (2)$$

In order to detect the pitches of an audio signal, it is split up in equal sample windows and each window is analyzed separately. The size of the windows, i. e. the FFT size, affects the resolution of the resulting spectra: the larger the size, the higher it is. The resulting output is divided into bins which contain the magnitude information for the respective frequency.

The magnitude of each FFT bin  $i$  can be expressed like this:

$$mag(i) = \sqrt{(re(i)^2 + im(i)^2)} \quad (3)$$

To obtain the loudness of each frequency, the magnitude must be transformed to decibel:

$$A(i) = 20 * \log_{10}(mag(i)) \text{dB} \quad (4)$$

Starting from this, the frequency with the largest magnitude can be identified as the fundamental frequency of the analyzed window.

## 4 Modeling

To realize a pitch detector, an audio signal must be converted to a sequence of samples first.

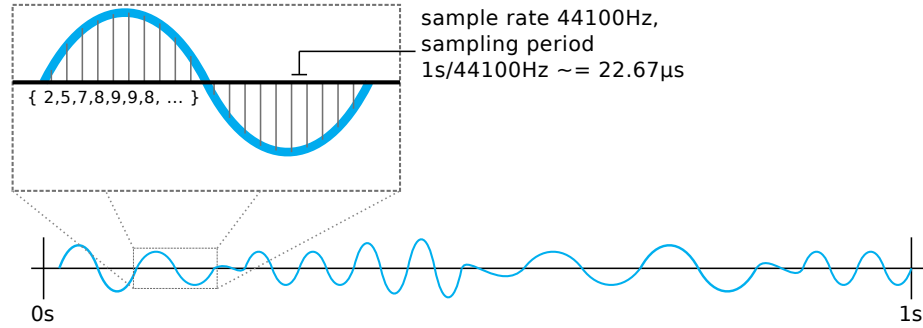


Fig. 1: Audio sampling

Find corresponding key on piano:

$$n = 1200 \cdot \log_2 \left( \frac{b}{a} \right) \quad (5)$$

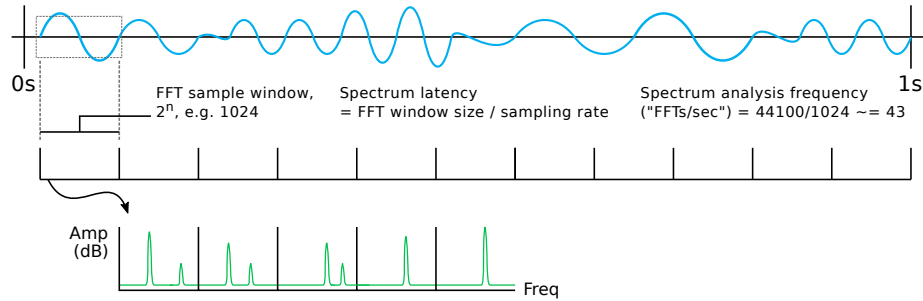


Fig. 2: Frequency Spectrum Analysis & Signal Transformation: Time Domain to Frequency Domain

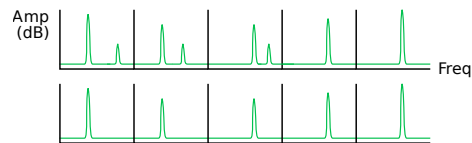


Fig. 3: Pruning

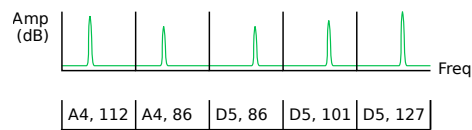


Fig. 4: Classification

Calculate deviation:

$$d_f = 49 + 12 \cdot \log_2 \left( \frac{f}{440 \text{ Hz}} \right) \quad (6)$$

## 5 Implementation

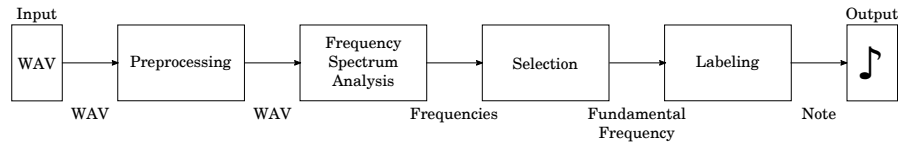


Fig. 5: System diagram