Final project presentation

DETECT FUNDAMENTAL FREQUENCY IN TIME DOMAIN USING AUTOCORRELATION FUNCTION, AVERAGE MAGNITUDE DIFFERENCE FUNCTION AND IN FREQUENCY DOMAIN USING HARMONIC PRODUCT SPECTRUM ON FAST FOURIER TRANSFORM RESULT.

Class id: 1022103.2010.18.10
Develop tools: Python and its built-in library
Team 10:

Dam Quang Tien (team leader)
Nguyen Nghia Thinh
Tran Giang Phong

INTRODUCTION

INTRODUCTION

Physical nature of human voice

- sound signal with specified fundamental frequencies around 90hz to 300hz
- base on the fundamental frequency, computer can predict the nature of the sound source. (male/female speaker, tone of words or sentences)

INTRODUCTION

Approach

- time-domain
  - autocorrelation function
  - average magnitude difference function
- frequency-domain
  - harmonic product spectrum based on the result of FFT functions

TIME-DOMAIN AUTOCORRELATION FUNCTION

Tran Giang Phong

AUTOCORRELATION FUNCTION

- The original Autocorrelation function is defined as:

$$r_{xx}(l) = \lim_{N\to\infty} \frac{1}{(2N+1)} \sum_{n=-N}^{N} x(n)x(n+l)$$

$r_{xx}(l)$ is value of Autocorrelation function with the delay l, $(2N+1)$ is the window length, $x(n)$ is the amplitude of the signal in the timestamp n.

Autocorrelation function has some features:

$r_{xx}(l) = r_{xx}(-l)$

$l = 0$: $|r_{xx}(l)| \leq r_{xx}(0) \; \forall \, l$;

$r_{xx}(0)$ is power of voice signal.

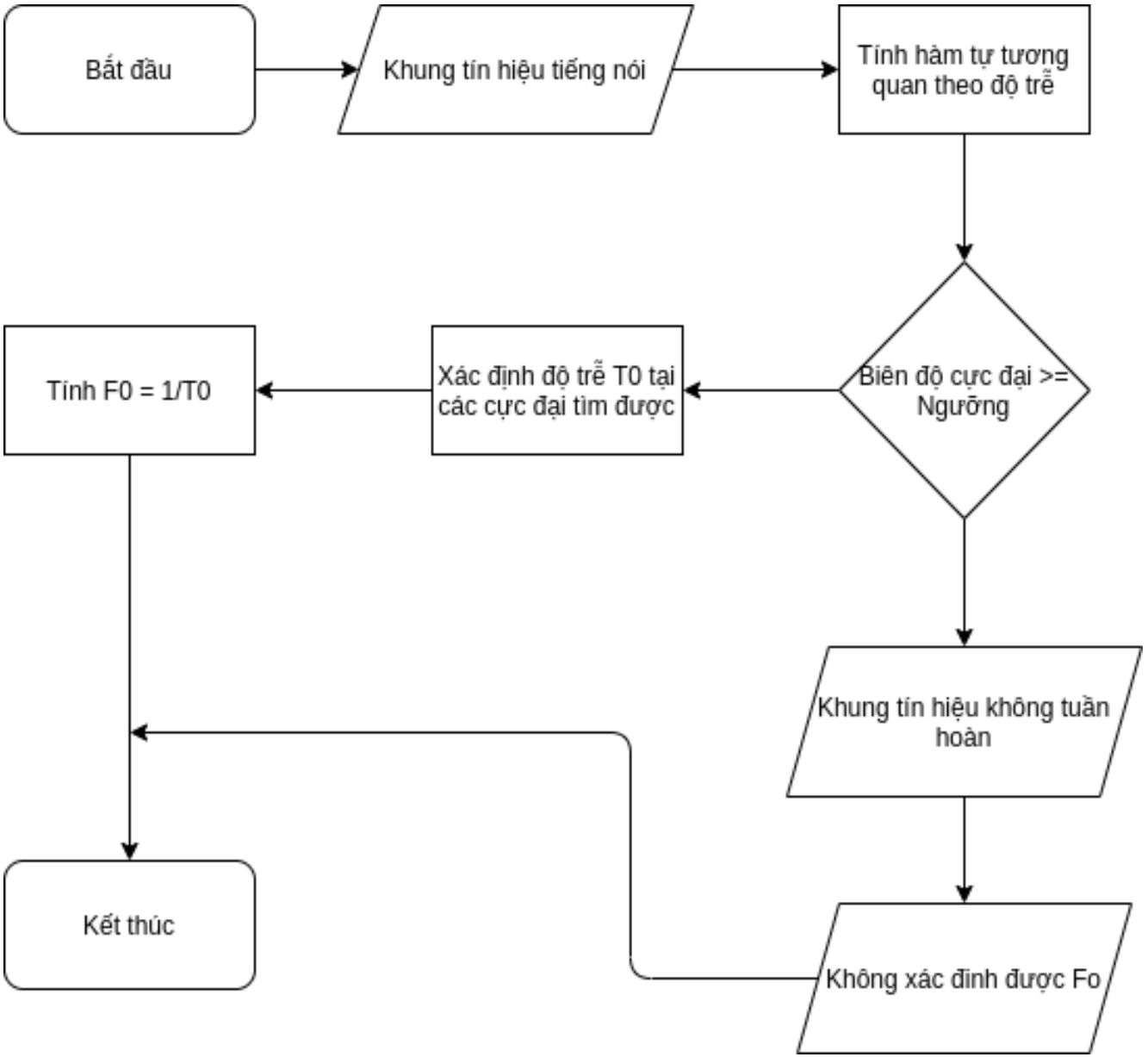
AUTOCORRELATION FUNCTION IN FRAME ANALYSE

$$r_t(\tau) = \sum_{j=t+1}^{t+W} X_j X_{j+\tau}$$

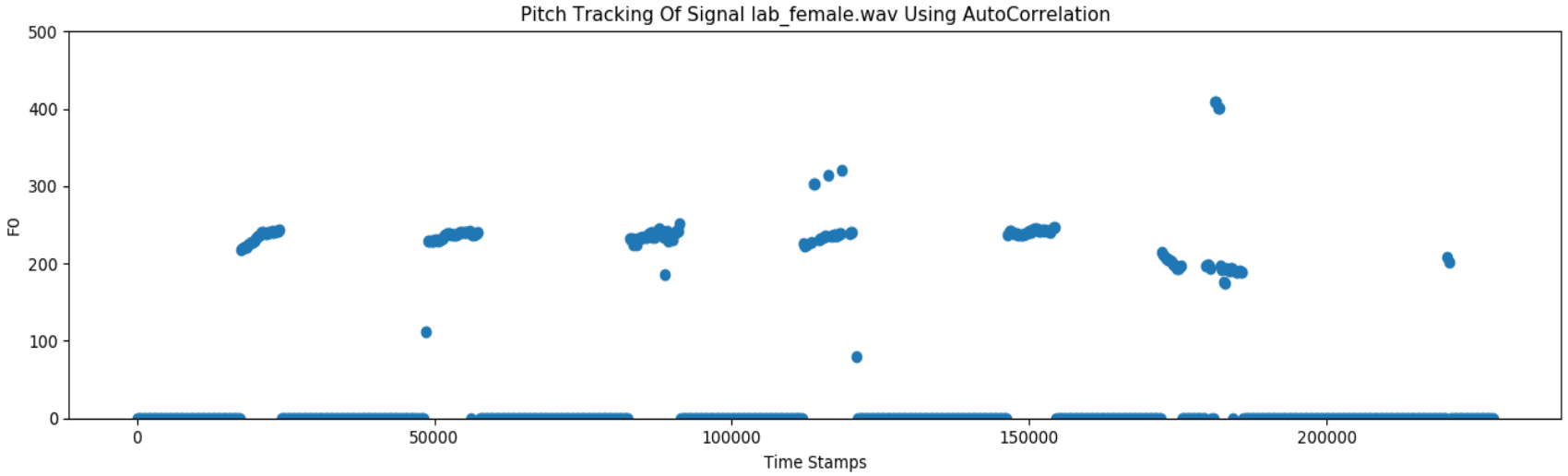
X_j is the amplitude of the signal in the timestamp j, $r_t(\tau)$) is value of Autocorrelation function with the delay τ in the frame t, and W is the length of the signal frame

AUTOCORRELATION FUNCTION

Diagram



AUTOCORRELATION FUNCTION



Result of the autocorrelation function of the lab_female.wav

TIME-DOMAIN AVERAGE MAGNITUDE DIFFERENCE FUNCTION

Nguyen Nghia Thinh

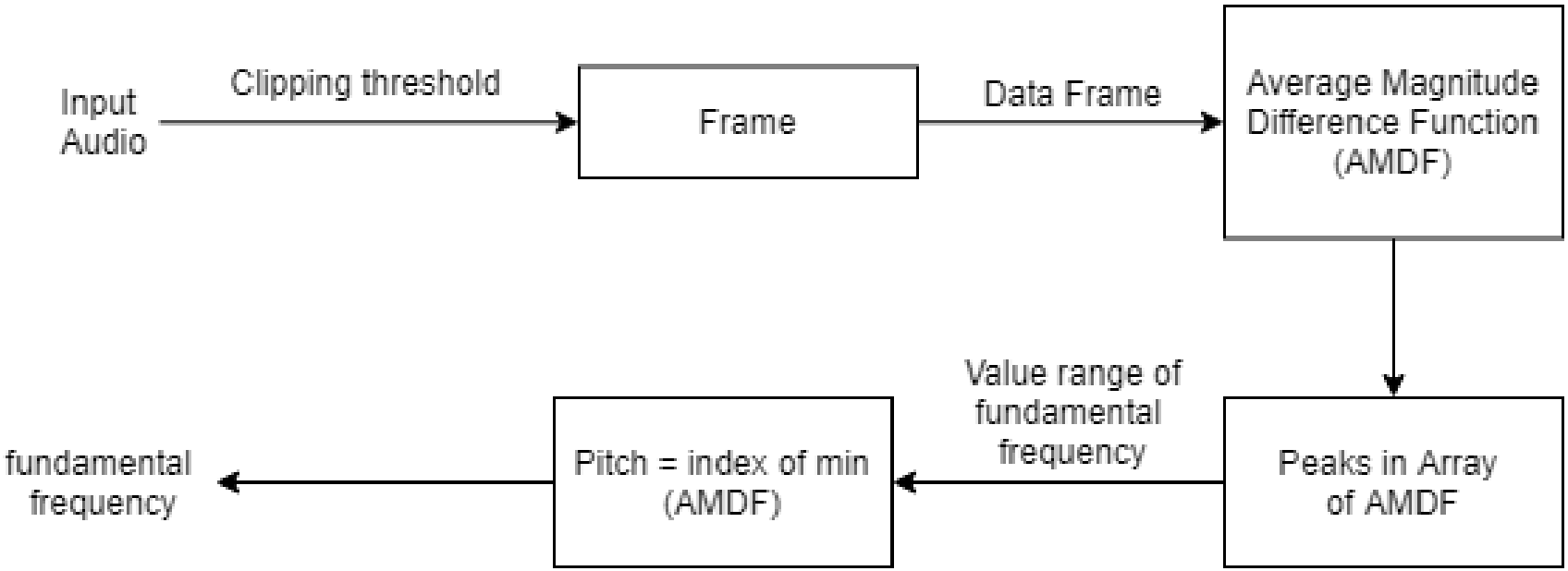
AVERAGE MAGNITUDE DIFFERENCE FUNCTION

- Period detection technology for weak characteristic signals is very important in the fields of speech signal processing, mechanical engineering, etc.
- Average magnitude difference function (AMDF) is a widely used method to extract the period of periodic signal for its low computational complexity and high accuracy.
- The original AMDF is defined as:

$$D(\tau) = \frac{1}{N-1-\tau} \sum_{n=0}^{N-1-\tau} |x(n) - x(n+\tau)|$$

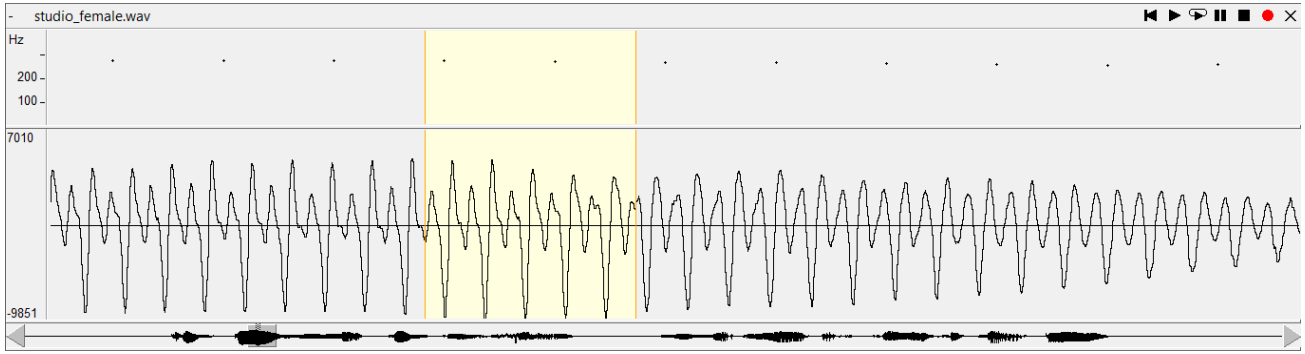
$$T_P = \arg\min_{\tau} \left(D(\tau) \right)$$

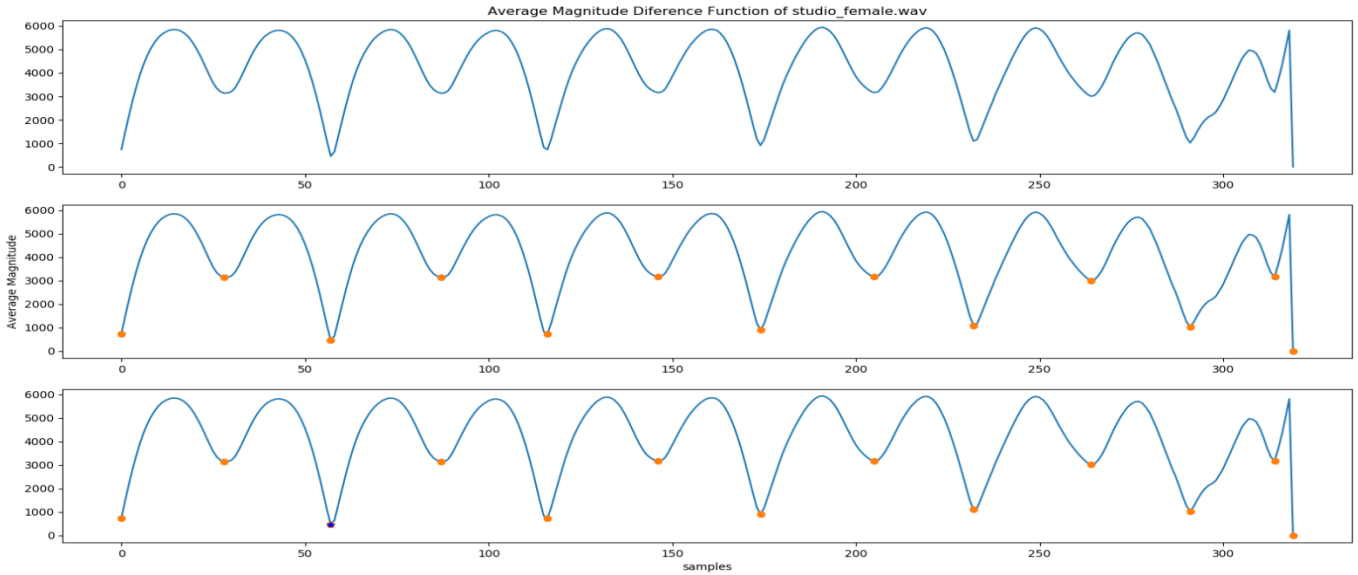
AVERAGE MAGNITUDE DIFFERENCE FUNCTION



FRAME USING AMDF

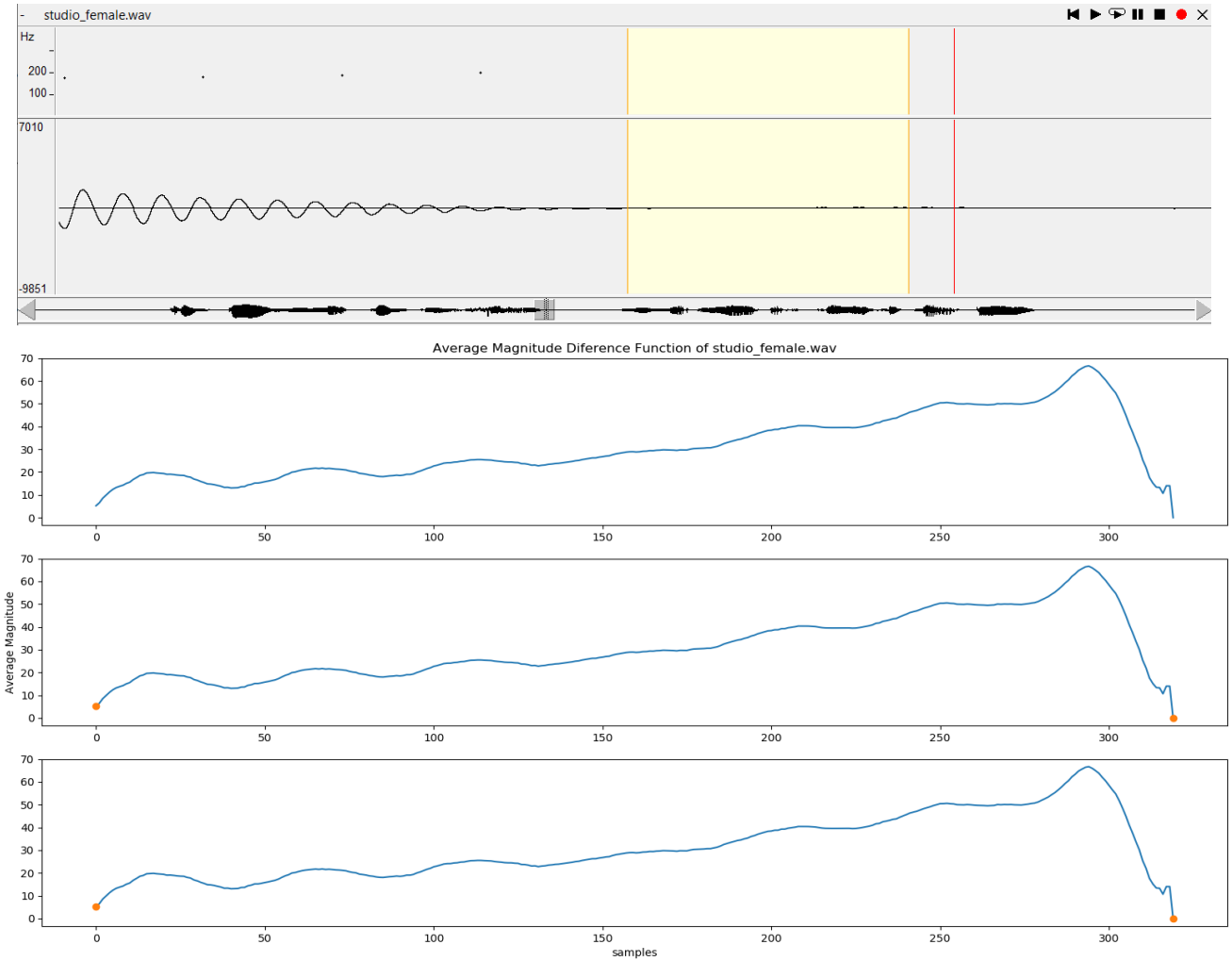
Pick a frame in original signal with time samples from 0.902s to 0.922s



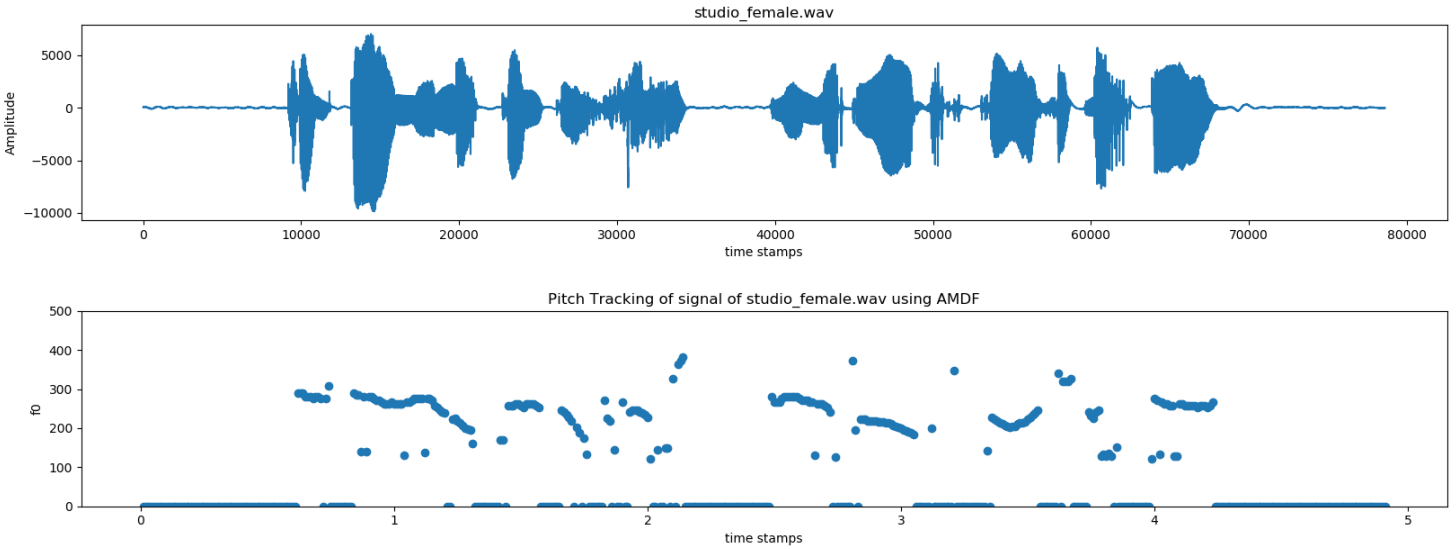


FRAME USING AMDF

Non-voice frame



RESULT



CONCLUTION:

- The results are accurate in a studio environment.
- There are some inaccuracies due to noise, and it is difficult to determine cyclic time and non-cyclic time.
- It is necessary to filter by the following methods: median filter, Top-Hat.

FREQUENCY-DOMAIN: HARMONIC PRODUCT SPECTRUM ON FFT RESULT

Dam Quang Tien

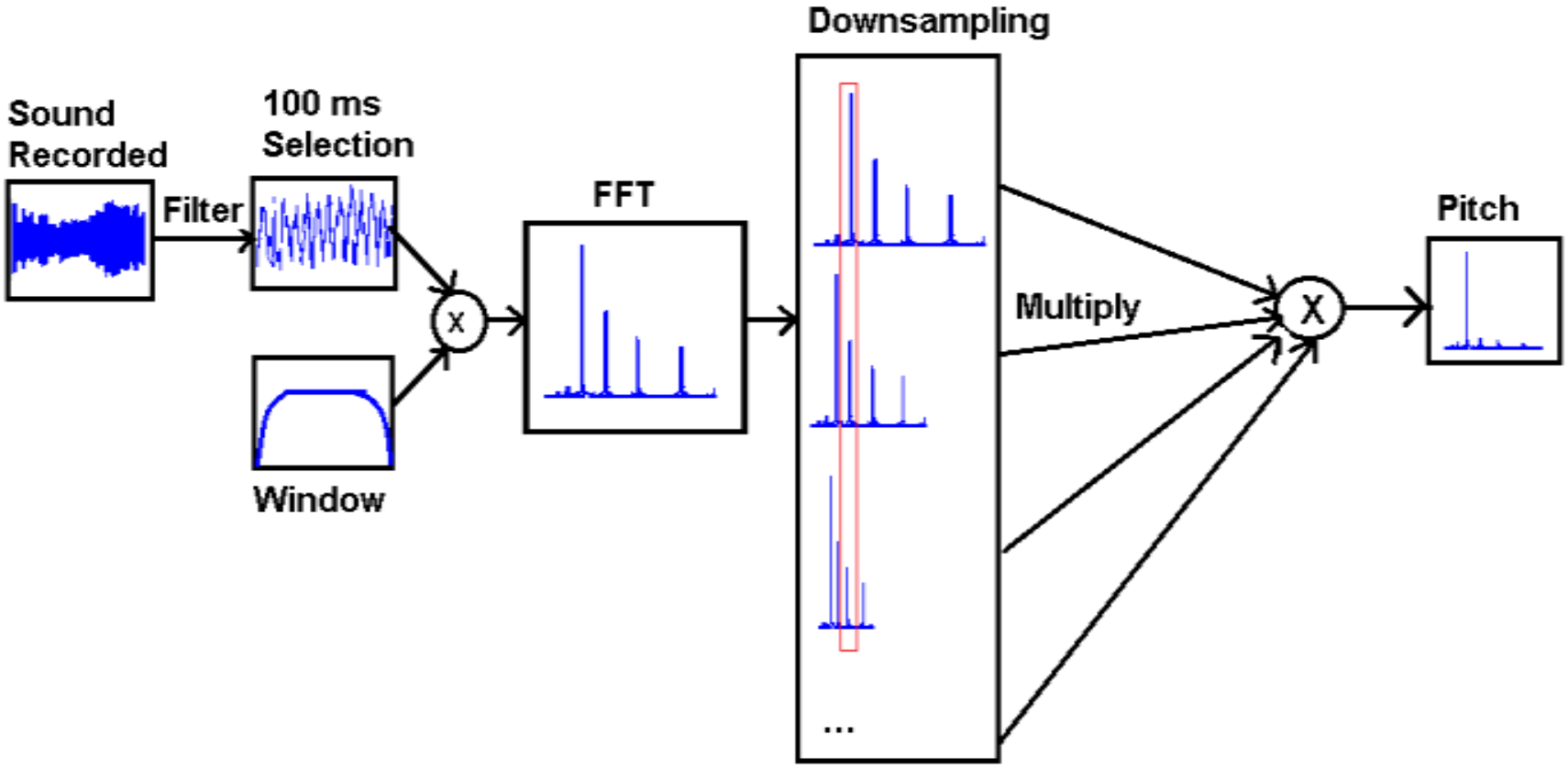
HARMONIC PRODUCT SPECTRUM ON FFT RESULT

Theory

$$Y(e^{j\omega}) = \prod_{r=1}^{R} \left| X\left(e^{j\omega r}\right) \right|^2$$

$$p = \underset{\omega}{argmax}\, Y\{(\omega)\}$$

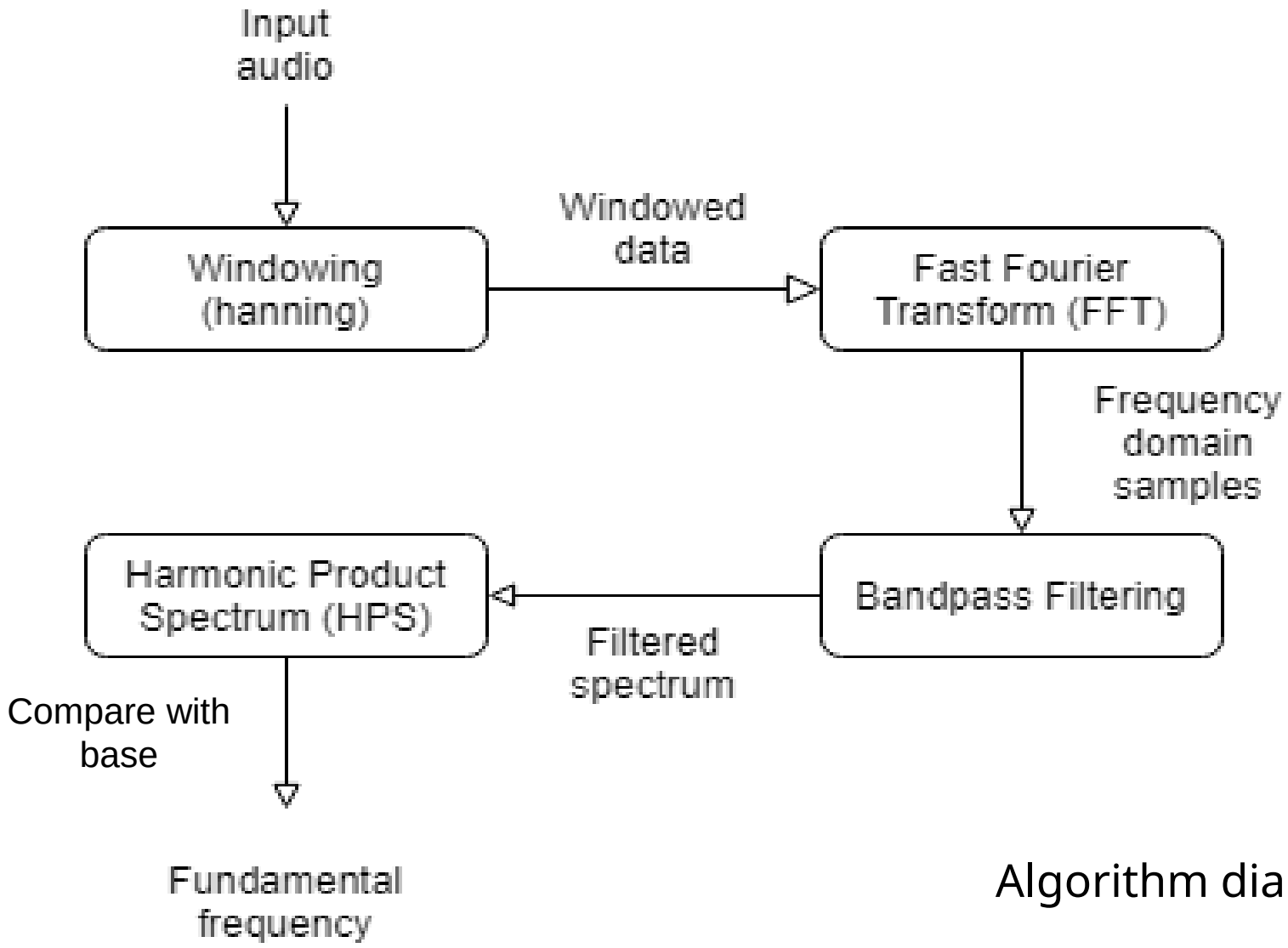
HARMONIC PRODUCT SPECTRUM ON FFT RESULT



Fundamental frequency detection step using HPS

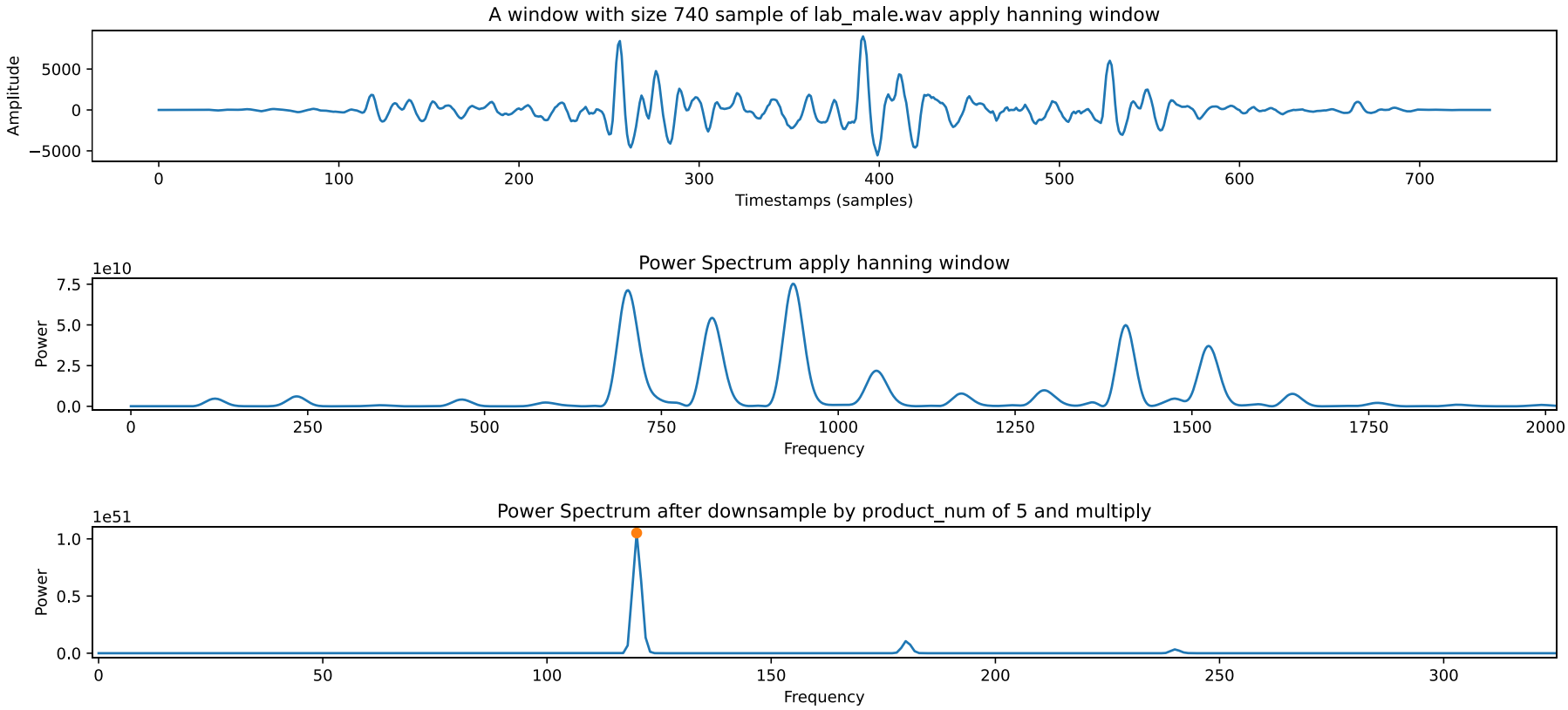
Credit: Vo, Thanh & Sawada, Hideyuki. (2017). [2]

HARMONIC PRODUCT SPECTRUM ON FFT RESULT

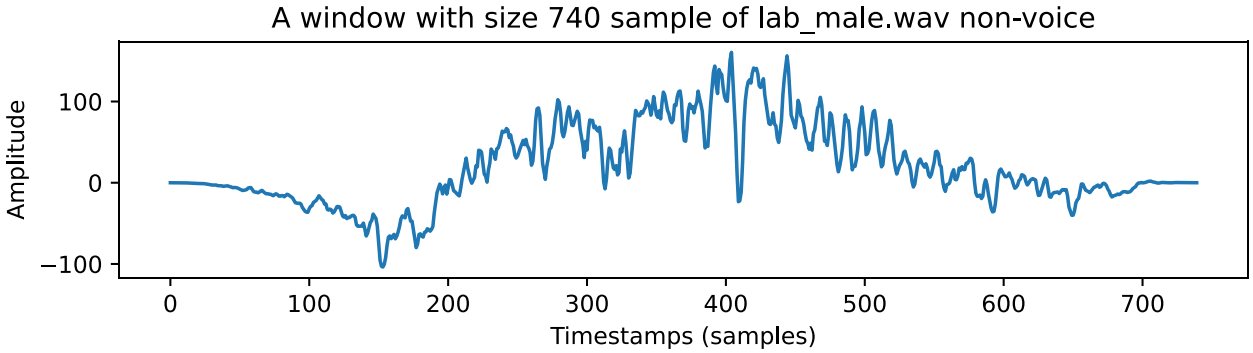


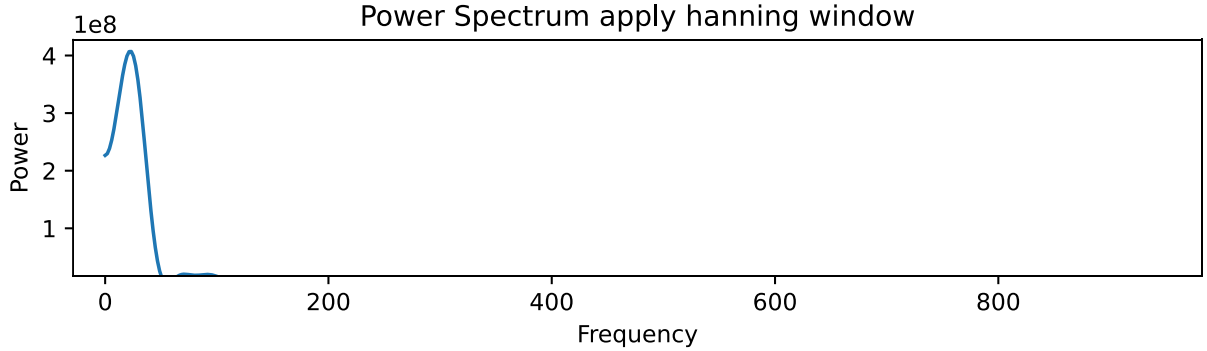
Algorithm diagram

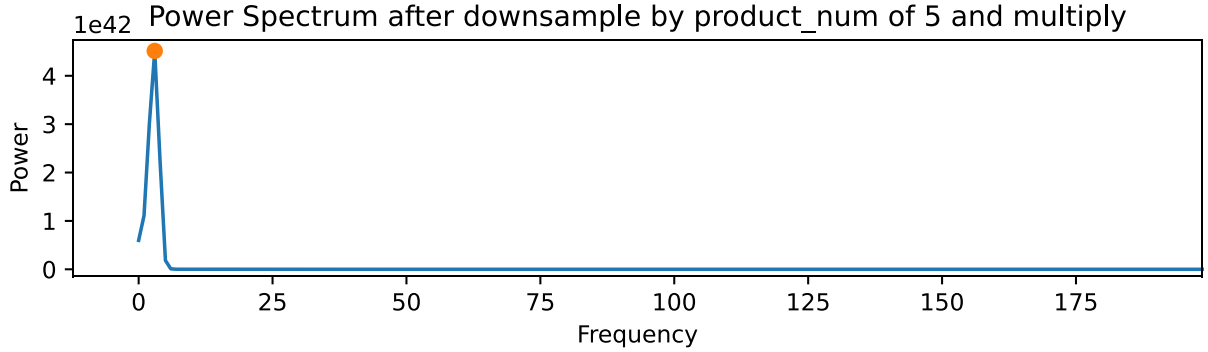
HARMONIC PRODUCT SPECTRUM ON FFT RESULT



The algorithm return the fundamental frequency at around 120Hz which is accurate in this case.

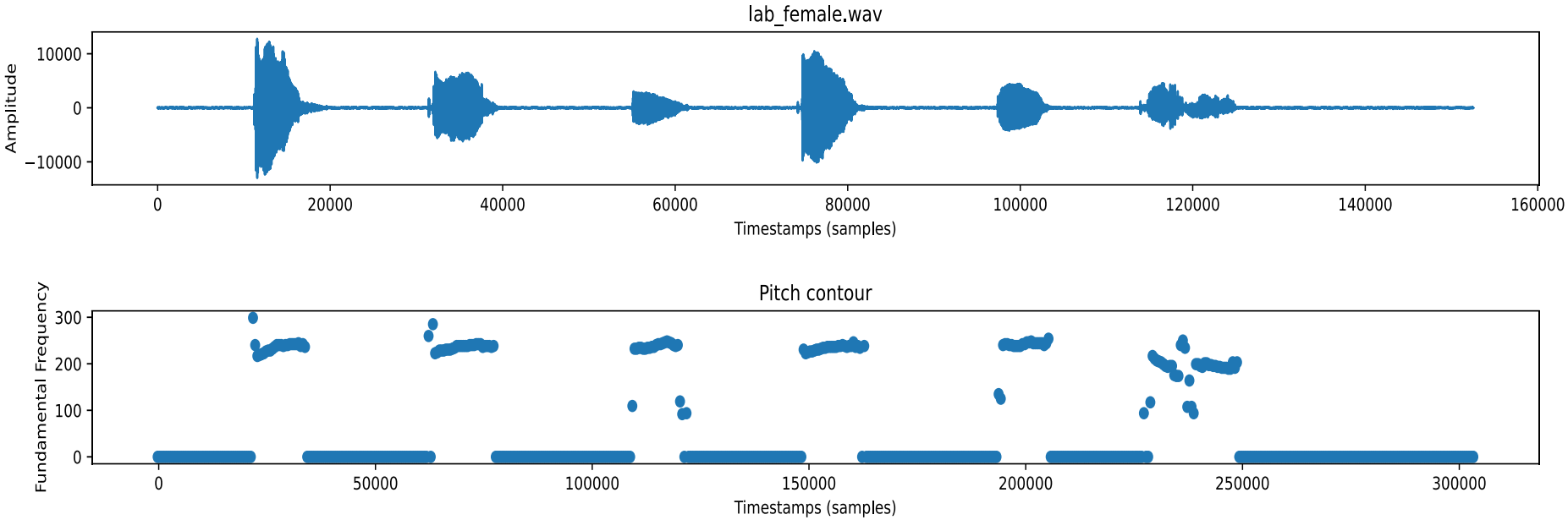






Max power of the nonvoice part (1e42) is relatively low to its counterpart of the voice part (1e51).

HARMONIC PRODUCT SPECTRUM ON FFT RESULT



Conclusion:

- Cheap in complexity and easy to install
- Can be use to detect f0 in real-time

RUNTIME COMPARISON

Method	Runtime
Autocorrelation	50.0409693s
Average Magnitude Difference Function	20.7626348s
Harmonic Product Spectrum on FFT result	0.5441494s

Ran on the same system using the signal file “**studio_male.wav**”, length 08s.

CONCLUSION

The autocorrelation function is expensive in computation.

The AMDF is not as expensive as the autocorrelation and can be use to detect fundamental frequency in near-realtime, but it’s sensitive in noisy environment.

Harmonic product spectrum is inexpensive in computation and can be use to detect fundamental frequency in realtime.