



Music Seat 중간발표

## Scope

---

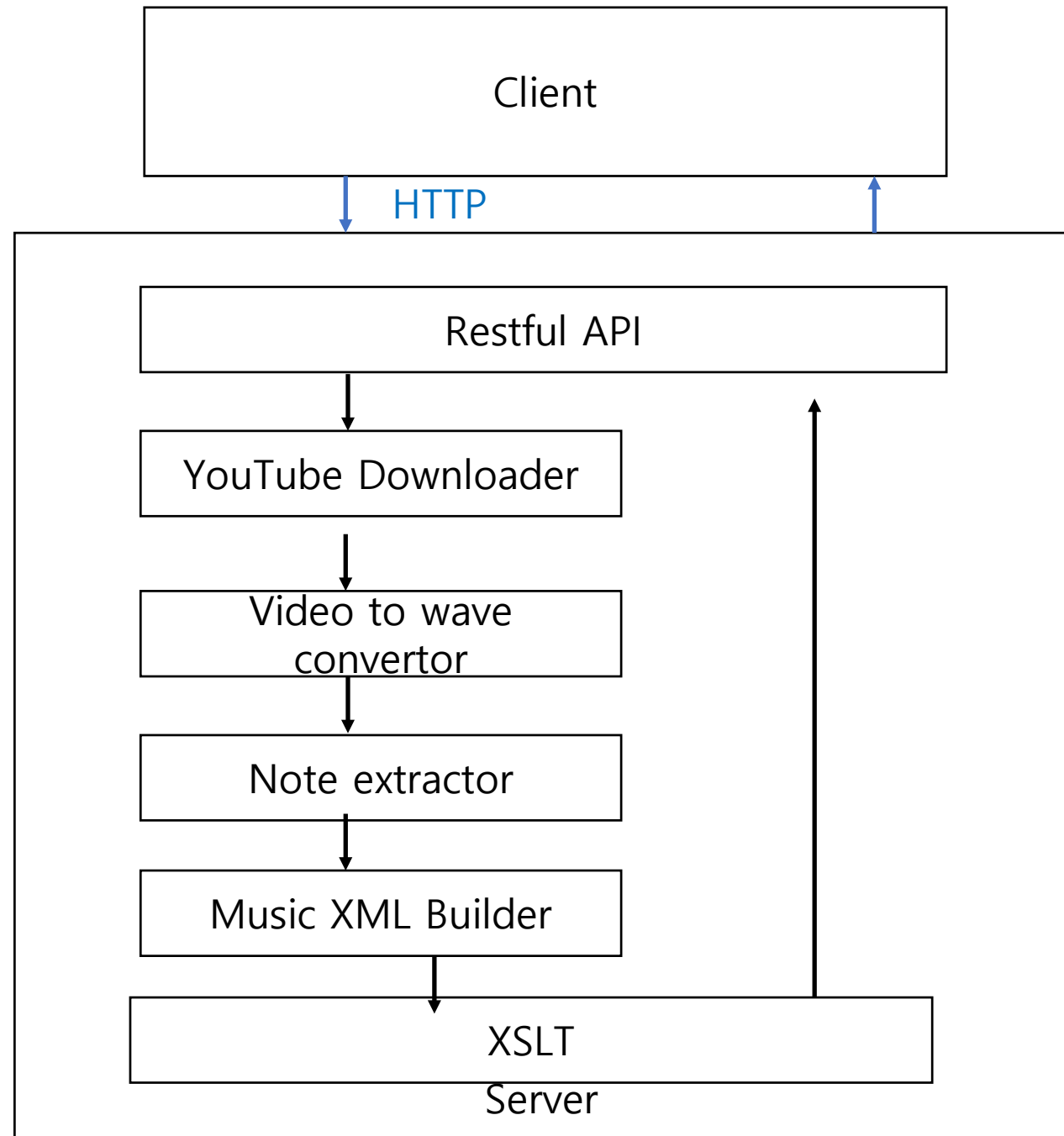
1. 화음이 없다고 가정.
2. 한 가지의 단순한 멜로디  
악기 Ex) 피아노 독주
3. 대략적으로 멜로디를  
측정한다.

# Technical term

---

1. 유튜브 URL에서 mp4 동영상을 추출한 뒤 wav 음성 파일로 변환
2. wav 파일을 읽어서 초당 44k 샘플링진행 음성 데이터를 4.4k씩 분할
3. 4.4k씩 분할된 음성 데이터를 각각 Fourier transform을 진행
4. Fourier transform 결과에서 amplitude 와 frequency를 추출

# Architecture



# Client

---

Angular JS 사용

<https://github.com/kookmin-sw/2018-cap1-14/tree/client>



# Sever

---

Flask

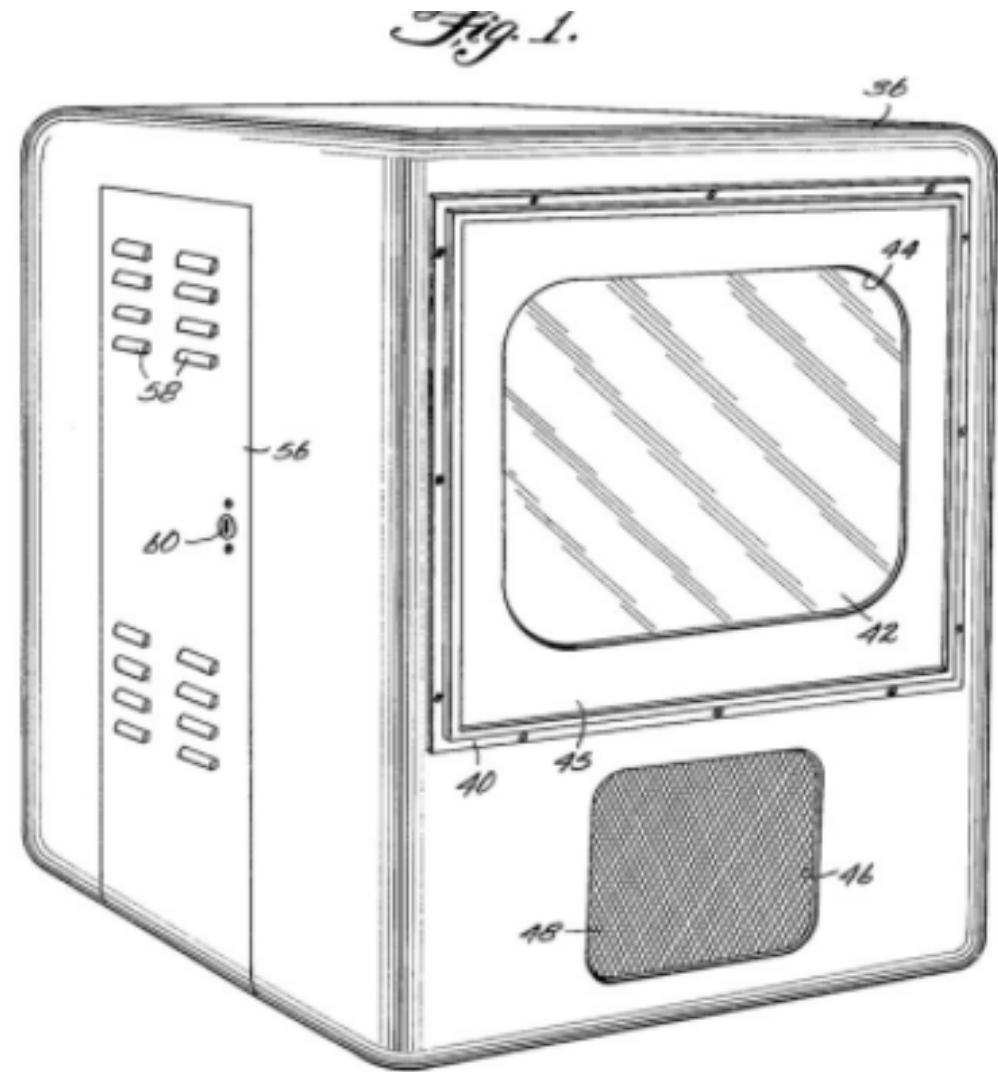
<https://github.com/kookmin-sw/2018-cap1-14/tree/server>



# YouTube URL에서 WAV 음원을 추출

PyTube

<https://github.com/nficano/pytube>



# 문제점

- PyTube를 사용하면 YouTube의 URL에서 즉시 음원을 추출할 수 없는 문제가 발생

# 해결방안

- 동영상으로 추출한 후에 MP3형식 파일로 저장한 후 FFMPEG을 이용하여 WAV파일로 변환한 뒤 파일을 읽어오는 형식으로 문제 해결



분할된 데이터를  
각각 Fourier  
transform을 진행

---

1. 사람이 연주한다는 가정하에, 1초에 10번 이상 연주를 할 수 없다고 가정했다.
2. 음원을 0.1초 간격으로 분할하여 각각 FFT를 진행하였다.
3. 그리고 각각의 결과를 sequential하게 분석하였다.

# Fourier transform 결과에서 Amplitude 와 Frequency를 추출

---

1. Scipy의 fftpack의 fft함수를 통하여 WAV파일을 fast Fourier transform한다.
2. 해당 진행 후 산출된 결과값의 범위를 반분하고 절대값을 취해준다.
3. Numpy의 angle함수를 통하여 일반화 해준다.
4. 해당 부분의 음원에 포함된 모든 주파수의 dB가 측정된다.
5. 400dB이상의 가장 첫번째 주파수를 기준으로 추출한다.

# 문제점

- 악기의 음파의 형태가 Sin함수 형태의 파동이 아닌 다중주파수가 혼합된 형태로 나타나서 주파수의 정확한 측정이 어려운 문제가 발생.

# 해결방안

- FFT변환 후에 많은 주파수 중에서 대략 400dB 이상의 강도의 첫번째 주파수를 기준으로 측정하면 정확한 음정을 알 수 있다는 것을 계속적인 실험을 통해서 찾아냄.
- 향후에는 HPS 알고리즘을 적용하여 더욱 더 정확한 측정을 할 수 있게 할 예정.

지금 집중하고  
있는 분야

---

- 음정추출에서 음정을 바탕으로 Music XML로 변환하는 것에 집중할 계획입니다.

## 현재 한계점

---

- 두개 이상의 음을 동시에 연주하면 동시에 측정(인식) 불가.
- 현재 인식할 수 있는 음역대의 옥타브 개수가 1개.
- #, b을 인식 불가하다.
- 단순한 한가지 멜로디 악기만 인식가능.

## 목표의 상향조정 가능성

---

1. 화음이 인식 및 측정이 안되는 것을 가능하게 한다.
  1. 더욱더 많은 주파수를 분석하여 HPS알고리즘을 도입하여 인식 및 측정 가능케 한다.
  2. 딥러닝을 사용하여 더욱더 정확히 측정한다.
2. 인식하는 음역대의 옥타브 개수를 증가시킨다.
  1. 옥타브의 데이터를 추가한다.
3. #, b을 인식 불가하다.
  1. 데이터의 추가.
  2. HPS알고리즘 도입하여 더 정밀한 주파수를 측정하여 음정사이의 오차범위를 줄여 세분화한다.