

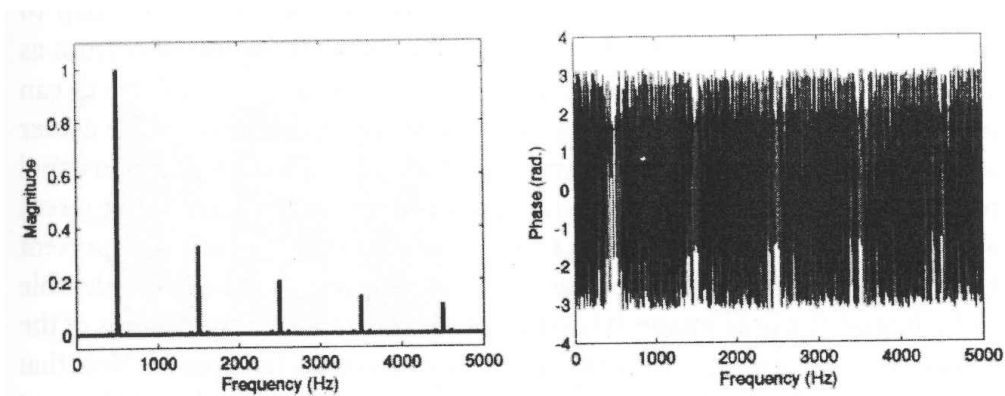
## 2. Visualization of Musical Signals

### 2.1. Introduction

- 시간-주파수 분석법(Spectrogram)의 소개
- 인간의 시각 체계와 청각 체계는 매우 다르기 때문에, 음악 또는 음향을 시각화하는 것은 중요한 난제
- 인간 시각 체계의 공간 지각력은 우수하지만, 색상을 판단하는 빛의 스펙트럼에 대한 지각은 공간 지각력에 비해 매우 열등함
  - 색상을 감지하기 위해 삼원색을 기준으로 하는 3종류의 색상 감지기가 사용되며, 삼원색의 정도만 정확하게 조절한다면 인간이 느끼기에는 똑같은 색상을 쉽게 만들어 낼 수 있음
- 인간 청각 체계의 소리 스펙트럼에 대한 분석, 지각력은 매우 우수하지만, 시각 체계와는 반대로 공간 지각력이 매우 열등함
- 음악을 시각화하려는 다양한 시도가 존재했음
  - Stephne Malinoski, Kaper and Tipei: 음악을 이산적인 사건으로 인식하여 분석
  - Jonathan Foote, Matthew Cooper, Martin Wattenberg: 동기 및 음악의 스펙트럼으로부터 파생하는 자기유사성(Self-Similarity)을 이용한 분석을 시도

### 2.2. Time-Domain Representation

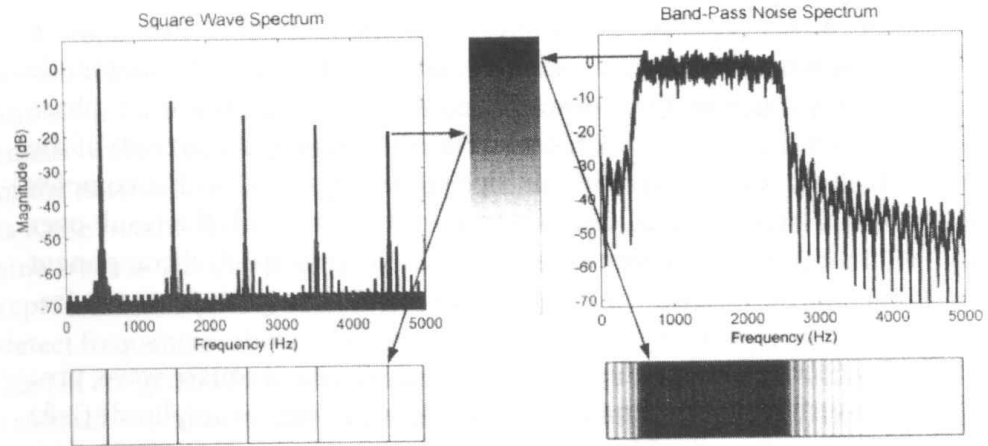
- 시간에 따른 음량 변화를 나타내는 그래프로 소리를 시각화하는 방법
  - 파동의 모형, 즉 파형(waveform)에 대한 세밀한 관찰이 필요할 때 유용함
  - 시각적 인식과 청각적 인식사이의 상관관계가 미약함; 시작과 끝이 즉각적이고 주파수가 높은 소리를 이 방식으로 시각화 시켰을 경우, 투박한 사각형이 나타날 뿐이라 의미를 알 수 없게 됨.



- 시간에 따라 음색이 변화하는 경우가 많은 Electro-acoustic Music을 위한 효과적인 시각화 방법이라 할 수 없음

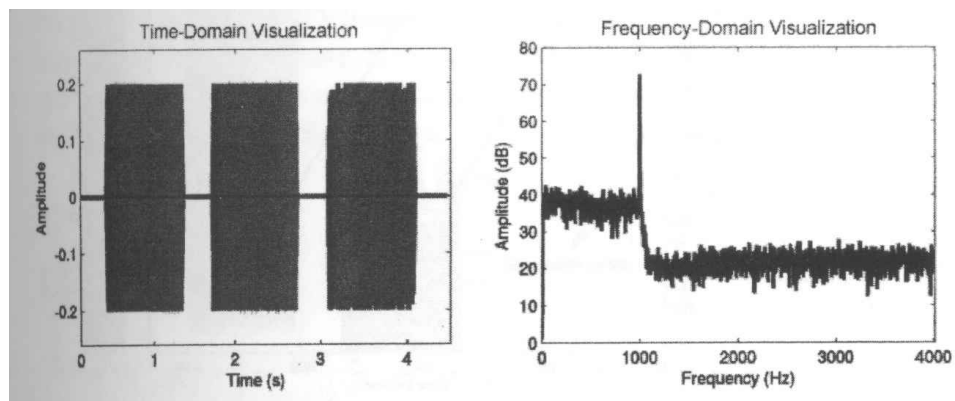
### 2.3. Frequency-Domain Representation

- Fourier Transform에 기초한 시각화 방법
  - 음향적 신호의 시각화에 있어 Time-Domain 방식보다 적합함; 시각적 인식과 청각적 인식 사이의 상관관계를 개선
  - 분석한 소리의 전체적 배음구조 즉, 음색은 파악할 수 있으나 시간에 따라 변화하는 음색에 대한 묘사는 불가능; 이 문제는 FFT 변환이 가지고 있는 자체적 한계로 인해 발생

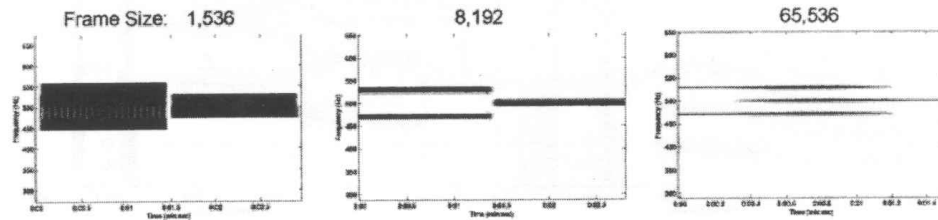


## 2.4. Time-Frequency Representation: Spectrogram

- STFT(Short-Time Fourier Transform) 분석과 시간 통합에 기초한 시각화 방법
  - FFT 변환이 갖고 있었던 태생적 한계를 극복하여, 실시간 스펙트럼 분석을 가능하게 했음. 전술한 2가지 방법에 비해 인간의 인식적인 면에 있어 우수함



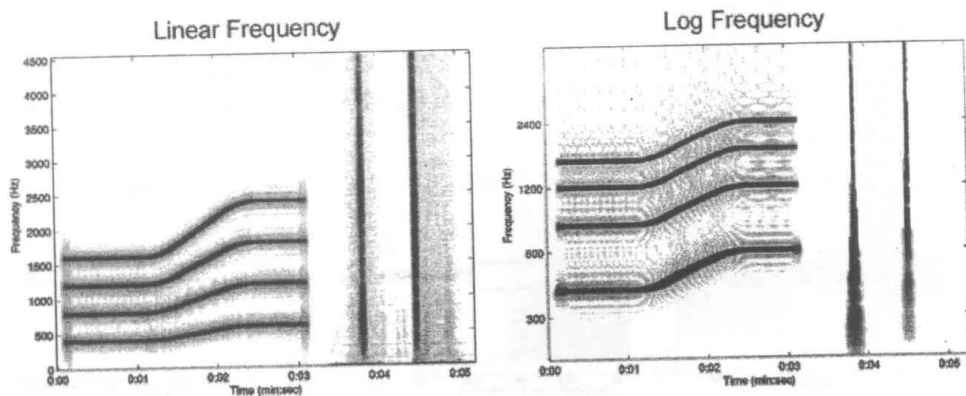
- 불확정성의 원칙에 의한 한계
  - 시간 해상도와 주파수 해상도 사이의 Trade-Off가 존재; Window Size가 작을수록 소리의 시간적 변화의 반영도가 높으나, 분석된 주파수가 정확하지 않게 나타남. Window Size가 클수록 분석된 주파수는 정확하게 나타나지만, 소리의 시간적 변화의 반영도가 미비함
  - 적절한 Window Size를 정하는데 원칙이 없으며, Case-by-case로 최적화 된 값을 찾아야 함



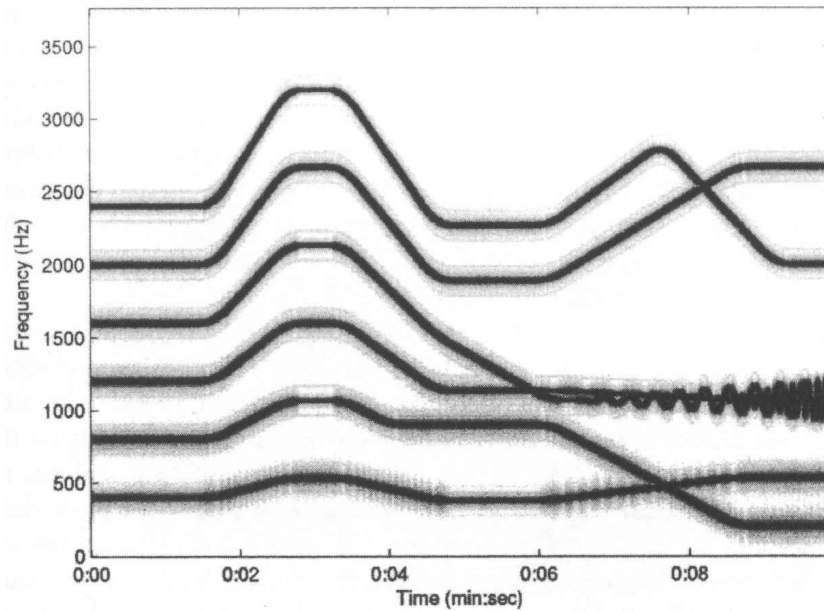
- 비슷한 접근에 약간의 보완을 추가한 방법인 Wigner Distribution, Modal Distribution이 있으나 널리 사용되지는 않음
- 이를 통한 분석 및 시각화가 심리음향학적으로 반드시 부합하는 것은 아니며, 분석된 결과 혹은 시각화를 다시 해석하는 것 역시 신호처리 및 심리음향학적 개념에 대한 이해가 필요

## 2.5. Spectrogram Examples

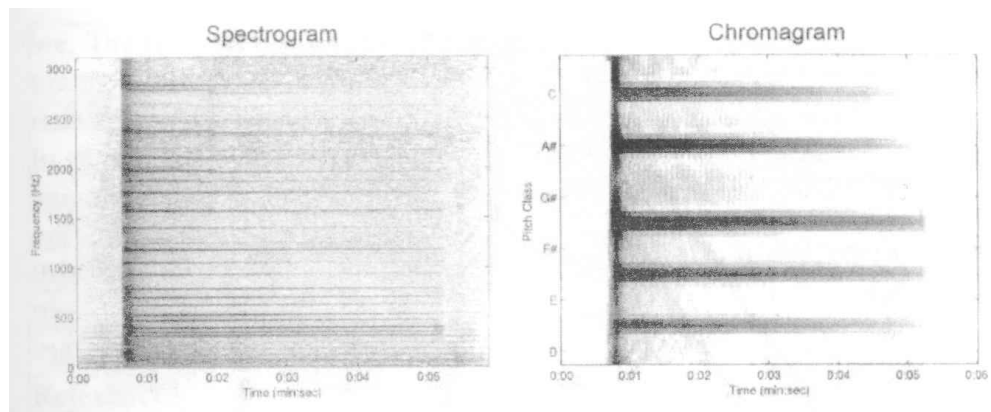
- Color Map, Color Scheme의 사용
  - 유용할 수도 있으나 시각 인식적 측면에서 오해를 불러일으킬 수 있음
  - VU Level Meter의 Color Scheme을 사용하는 것은 음량이 큰 배음을 강조하는데 효과적
- Linear Frequency VS Log. Frequency (Contant-Q Transform)
  - Spectrogram은 기본적으로 STFT에 기반을 두고 있기 때문에, 주파수 대역을 균등하게 나눈(linear) Frequency Bin들을 사용하여 소리를 분석함
  - Logarithm(Exponential)에 기초한 주파수 분석을 통해 얻는 Spectrogram은 옥타브 간의 거리를 동일하게 되도록 주파수 대역을 분할함으로써 시각 인식적 측면에서 보다 우수함; 저역에서 보다 자세한 주파수 해상도를 얻을 수 있으나, 그 역으로 저역에서의 시간에 대한 분해능은 다소 떨어짐



- 청각 체계에 대한 흥미로운 실험
  - 처음으로 피실험자에게 들리는 독립적인 순음은 2차 배음(2nd Partial); 다른 배음들과는 달리 하향하는 것을 가장 먼저 멈추기 때문
  - 또한 4.3초 부근에서 4번째 배음이 독립적으로 구분됨: 다른 배음들과는 달리 멈추지 않고 계속 하향하고 있기 때문



- 녹음된 실제 소리의 Spectrogram은 다소 복잡한 경우가 일반적이거나, 특정 주파수의 크기가 강하게 나타나는 Formant를 찾을 수 있음
  - 특히 Formant의 경우 '모음'의 발음을 구분하는 데 결정적인 단서로 사용할 수 있음
- Chromagram
  - 여러 옥타브를 하나의 범위로 반복 적용(wrapping)한 방식의 스펙트로그램
  - 모든 상위 배음이 옥타브 관계를 갖는 것은 아니므로 주파수 범위를 정함에 있어 주의 를 요함



- Spectrogram을 사용한 시각적 작품: 역발상의 예

## 2.6. 결론 및 요약

- 스펙트로그램은 Electroacoustic Music을 시각화함에 있어 매우 유용한 방법
- 스펙트로그램은 FFT에 의존하기에 Frame Size, Step Size등의 변수에 대한 세심

한 설정이 필요

**참고문헌**

n/a

**작성자**

최홍찬 (박사과정, 5기)

melvin@dgu.edu