

## 20. Perceptual Fusion and Auditory Perspective

*John Chowning*

### 20.1 Introduction

1. 청각 기관의 민감도: 청각기관은 변화를 측정하는데 있어 민감하게 반응 많이 알려지지는 않았으나 중요한 의미를 가짐.
2. Auditory perspective: Loudness 인식에 나타나는 다양한 통찰력에 의한 원근법
  - 사운드 인식에 관한 근본적인 이해에 의문 제기
  - 이러한 현상을 포괄적으로 이해하기 위해 요구되는 인식의 미묘함
    - 예) Periodic이란 자주 사용되는 용어이나 실제 자연계의 신호는 주기적이지도 않으며 인간의 청각 체계 역시 이를 인식

### 20.2 Perceptual Fusion and Quasi Periodicity

#### 20.2.1 The Limits of Perception

- Acoustic signal의 perfection
  - Acoustics signal에는 청각 체계가 반응할 수 있는 한계를 넘는 완벽한 정도가 존재
  - 연주자와 악기장은 완벽에 도달하기 위해 정교한 악기를 만들고 숙련된 연주에 의해 완벽함에 접근
  - 하지만 acoustic signal이 가지는 완벽함의 정도는 악기장과 연주자의 능력을 초월
- 완벽한 acoustic signal에 대한 인간의 반응
  - 실제 완벽함에 직면했을 때 청각/인식 체계는 일종의 aesthetic rejection 수행
  - 진정한 완벽함은 컴퓨터나 신디사이저 같은 디지털 장치에 의해 만들어진 전자적 소리에만 존재

#### 20.2.2 Periodicity and Quasi Periodicity

- Periodicity와 Quasi Periodicity의 정의
  - Periodicity: 실시간 압력(pressure) 변화가 완벽하게 정규적으로 순환하는 pattern
  - Quasi Periodicity:
    - 실시간 압력 변화가 주거나 압력에서 작은 변화를 가지고 발생하는 pattern. 주기적이지만 감지할 수 없는 변형을 가진 음향파(acoustic wave)는 실제 거의 존재하지 않고 스피커에 의해 만들어짐

• 소리의 일반적인 특징

- 인간이 감지할 수 있는 quasi-periodic wave로 이루어짐
- 청각 기관은 1%의 작은 부분에 나타나는 주기 변화도 감지  
→ quasi-periodicity에 민감

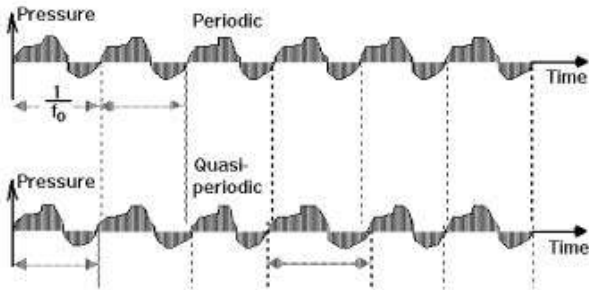


Figure 20.1 Periodicity와 Quasi Periodicity 비교

• Quasi Periodicity의 두 가지 형태

- 자연적 형성과 인공적 형성

자연적 형성: 연주자와 악기 모두에 나타나는 작은 결함에 의해 형성

인공적 형성: 연주자가 의식적으로 vibrato나 tremolo를 첨가함으로써 부가적 변형이 만들어짐

- Vibrato의 형성

가수: ① 공기의 압력이 성대를 지나갈 때 폐에서 공기 압력에 작은 변화가 생성되어, ② 성대 자체의 근육 긴장의 작은 변화에 의해, ③ 성도에 있는 음파와 결합한 성대에서의 교류로부터 야기된 non-linearity를 통해 발생

현악기: 현의 길이변화에 의해 발생

관악기: 공기 기둥의 길이, 리드나 입술의 긴장에 따라 발생

- Tremolo의 형성

가수: 성대를 통한 breath pressure를 변화시킴으로서 발생

현악기: 활의 압력 변화를 통해 발생

관악기: 공기의 압력 변화에 의해 발생

\*\* Organ pipe와 recorder의 경우 tremolo만 가능

- 두 종류의 modulation 모두(특히 vibrato) 다양한 음향적 음악적 인식적 기능 수행

20.2.3 Source Identification

• 소리의 특징을 만드는 미세한 pitch변화가 가지는 중요성에 대한 실험(1980)

- 방법: 가수 목소리의 단계적 modeling을 통한 실험

15초간 지속되는 소프라노 tone을 세 단계로 합성

단계1 근음의 사인파

단계2 단계1 + 배음 첨가

단계3 단계1 + 단계2 + random pitch variation(비주기적)과 vibrato(주기적) 혼합

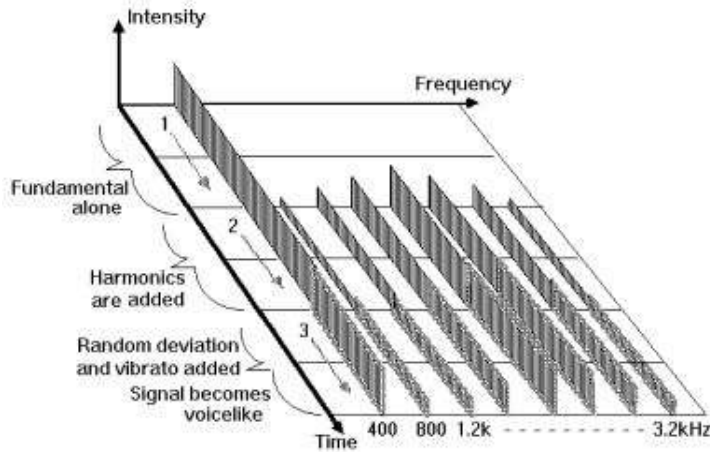


Figure 20.2 세 단계에 의한 소프라노 음색의 modeling 과정에 나타나는 3차원 주파수 그래프

- 결과: 단계1을 제외한 단계별 특징
  - 단계2 - ① 목소리에 필요한 모든 스펙트럼 정보가 존재함에도 불구하고 목소리가 가진 특징이 확실하지 않음(source signature 갖지 못함)
  - ② 배음들이 근음과 하나로 결합되지 못함
  - 단계3 - ① 음성으로 인식 ② 배음 융합(fuse)으로 하나의 소리로 지각
- 결론:
  - ① 파형(이나 신호)이 quasi periodicity 상태에 있을 때 인식적 융합 발생
  - ② 이러한 특징은 모든 소리에 존재하는 것으로 보이며 소리의 종류에 따라 변화 pattern은 다르게 나타남

### 20.2.4 Source Segregation

- Source Segregation
  - 인간이 각 소리를 분리하여 인식 할 수 있는 이유는 청각 기관이 인식적 융합의 기능을 수행하기 때문
- Source Segregation 실험
  - Figure 20.2에 사용된 소리에 500, 600Hz 및 연관 배음 첨가
  - 예상: 400, 500, 600Hz의 pitch를 가진 3화음
  - 결과: 단계 1을 제외한 단계별 특징
    - 단계2 - 400, 500, 600Hz는 근음인 100Hz가 빠진 배음을 형성하기 때문에 3화음으로 잘 들리지 않음
    - 단계3 - 명확하게 3화음으로 들림
  - 결론: 각 그룹의 배음에 조금씩 다른 주기적인 vibrato와 임의적 vibrato의 혼합이 더해짐으로서 각 그룹 결합

### 20.2.5 The Chorus Effect and Spectral Smearing

• Chorus 실험

- 방법: Source segregation 실험에 사용된 세 주파수 및 연관배음을 동일한 주파수로 변경
- 결과: 단계별 특징
  - 단계1과 2 - Loudness만 증가
  - 단계3 - 각 그룹에 다른 pitch variation이 더해지면서 모든 가까운 배음 간에 발생하는 복합적인 random beating생성. 그 결과 spectrum smearing 발생 소리는 분리되며(source segregation), 세 개의 성부가 unison노래(chorus)

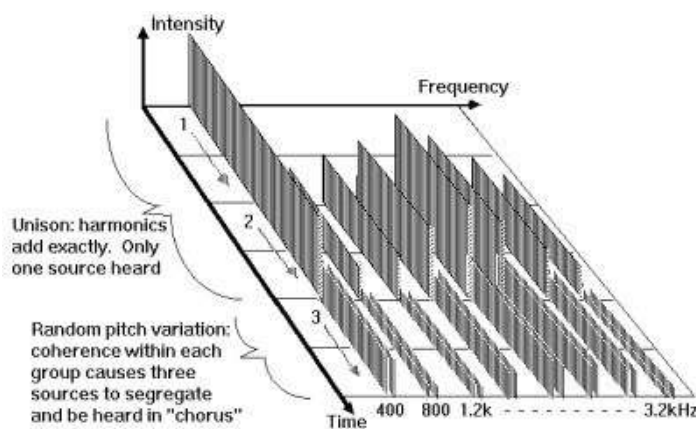


Figure 20.3 세 개의 동일한 pitch에 의한 소프라노 음색의 modeling 과정 그래프

- 결론
  - ① 여러 sound source가 있는 복잡한 음향적 흐름에서(코러스 같은) 각 source는 각각의 독특한 random pattern을 가짐
  - ② 이 pattern은 귀로 파악할 수 있는 정도의 차이를 가짐
  - ③ Quasi periodicity를 가지지만 결함이 없는 source인 경우(사람이 인식하지 못하는 경우) 인식적 융합은 발생하지 않아 다른 소리로 듣지 못하며(source identification 실패), 여러 개의 소리로 인식되지 않음(source segregation 실패) 결과적으로 청각기관은 acoustic imperfection에 의해 소리를 구분

### 20.2.6 More about Vibrato

• Vibrato의 기능

- 다양한 음향적, 음악적, 인식적 기능 수행
- 자연스러운 random pitch variation 보완
- 배음은 resonant envelope에서 주기적으로 움직이며 복잡성을 야기하고 시간차를 둔 amplitude modulation을 발생시키기 때문에 하나의 소리가 가지는 음색적 풍부함은 적은 양의 vibrato에 의해서도 만들어짐

- 음악적 예: 앙상블에서 soloist 부분의 경우, dynamic range가 제한된 악기들은 (solo 악기 소리를 가리는) 앙상블의 소리로부터 솔로 악기 소리를 분리시키기 위해 vibrato 사용
- \*\* 이것은 시각에서 배경에 의해 감춰진 물체가 움직일 때만 분리되는 것과 유사

### 20.2.7 Periodicity and Symmetry

- 주기성과 대칭성에 대한 청각과 시각의 차이
  - 두 기관은 주기성과 대칭성을 같은 방법으로 취급하지만 그 정도에 차이를 둠
  - 시각: 주기성에 나타난 미세한 일탈의 정도를 인식하지 못함
    - 예) Figure 2.4에 나타난 그래프의 경우 주기에 표시된 선이 없기 때문에 즉각적으로 quasi-periodicity를 인식하지 못함
    - (비교: Figure 20.1은 동일한 내용에 선이 있어 인식 가능)

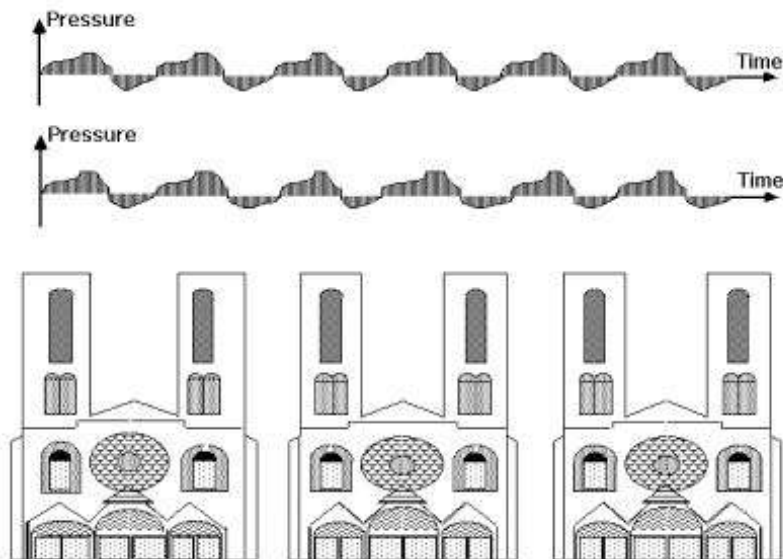


Figure 20.4 시각에서의 주기성과 대칭성 판단

- 청각: 청각 기관은 즉각적으로 주기성에 나타난 일탈의 비율 정도를 인식
- 단기간에 나타나는 주기성과 대칭성 인식
  - 두 기관은 주기와 대칭이 짧은 시간에 인식되는 경우 주요 정보를 정확하게 추출하는데 실패
  - 이것은 청각 기관의 경우 더욱 사실
  - 기계를 이용해 음악을 만드는데 인식 체계에서 요구되는 것에 주의를 기울여야 함
  - 일반적인 악기와 달리 전자 악기는 인식 체계가 의존하는 소리의 본질적인 결함을 갖지 않음

### 20.3 Auditory Perspective

- 청각에서의 원근법
  - 원근법은 시각뿐 아닌 청각에도 나타남
  - 시각의 원근법에서 물체의 크기가 변하는 것처럼 소리의 원근법에서는 소리의 강도 변화
  - 청각 기관은 인식한 정보를 시각 체계와 유사한 방법으로 mapping함
  - 이러한 특징은 스피커에 의해 재생되는 음악에서 더욱 중요
  - 음색정의는 청취자와 소리 재료의 거리가 멀어질수록 감소(시각의 color gradient와 유사)
  
- 두 감각에 나타난 원근법의 상관관계
  - 시각적으로 인식되는 물체가 소리재료일 수 있기 때문에(소리를 가진 물체의 경우) 두 개의 체계가 외부 세계를 파악하는데 있어 sensory mode의 충돌을 피하는 한 가지 방법으로 발전
  - 예) 어머니의 목소리, 사자의 으르렁거림, 빠르게 움직이는 자전거의 접근
    - 이러한 예는 생존과 연결

#### 20.3.1 Loudness

- 일반적인 견해와 실제
  - Loudness는 물리적 강도와 인식간의 관계로 생각되지만 실제 보다 복잡한 지각으로 이루어져 하나 이상의 dimension 포함
  
- Loudness 인식에 있어 나타나는 multi-dimension 관련 실험
  - 방법: 청취자와 두 명의 singer에 의한 실험
    - Singer 한명을 1m 거리, 다른 한 명은 50m거리에 위치
    - 가까운 singer는 *pp*의 tone을 멀리 있는 singer는 *ff*의 tone으로 노래
    - Tone은 동일한 pitch, 음색, duration을 가짐
    - 청취자에게 더 큰 소리를 선택하라고 요구

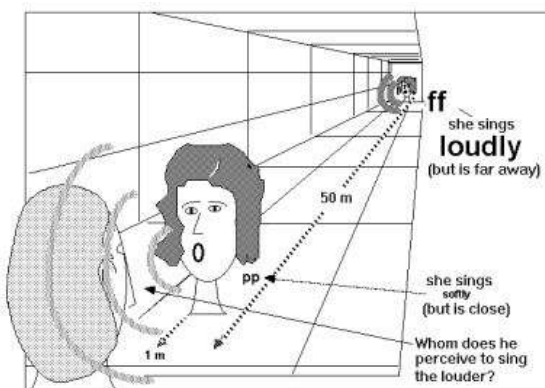


Figure 20.5 Loudness 측정 실험

• 일반적인 예상(direct sound만을 고려)

- 역제곱의 법칙: 어떤 지점에서 강도는 역제곱의 법칙에 따라 거리의 제곱으로 감소 ( $1/d^2$ )
- 역제곱의 법칙에 따라 두 거리(1m와 50m)의 강도가 동일한 경우, 1m의 강도가 1이라면 50m의 강도는  $1/50^2$
- 동일한 거리에서 *pp*의 강도가 1이라면 *ff*의 강도는 128
- 결론: 강도만을 고려한 경우 1:128/2500으로 1m의 *pp*가 더 크게 들림
  
- 하지만 청취자는 50m의 소리가 더 크다고 반응

20.3.2 Spectral Cue

• Pitch와 Energy 변화에 따른 스펙트럼 변화(Natural sound)

- Pitch가 높아질 경우: 배음의 개수는 줄어들고 스펙트럼의 모양도 변화
- Dynamics가 증가하는 경우: 스펙트럼 모양 변화
- Figure 20.6: pitch, dynamics, distance에 따른 배음요소의 강도 및 스펙트럼 변화

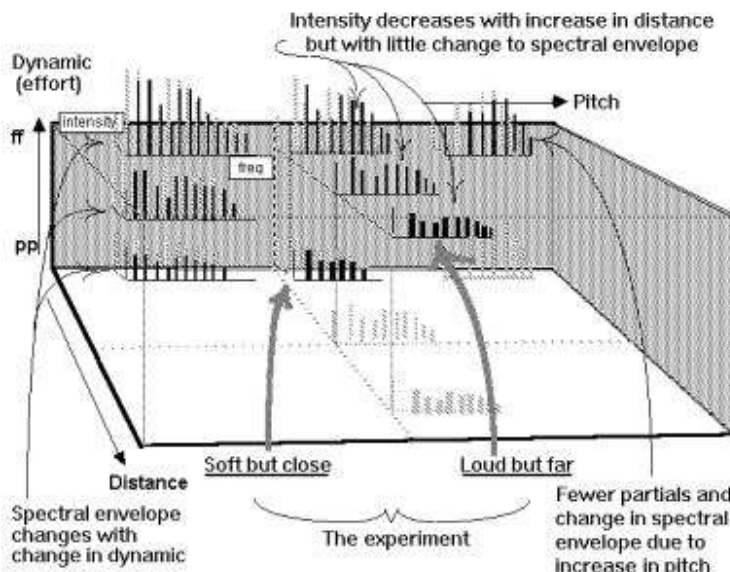


Figure 20.6 Pitch, dynamics, distance 변화에 따른 배음의 특징

- 배음의 변화에도 불구하고 하나의 음색으로 인식하는 것은(즉 음색의 연속성) ① perceptual fusion과 source segregation 그리고 ② 공간에서 tone의 위치를 인식하기 때문
- 청취자는 강도에 따른 음색의 질적 차이를 인식하기 때문에 loudness인식에서 스펙트럼은 중요한 역할을 함
  
- 하지만 실험에 사용된 두 개의 tone을 스피커로 바꾸어도(spectral difference가 없는 경우에도) 결과는 동일

### 20.3.3 Distance Cue and Reverberation

- **Direct signal:** 소리 재료에서 청취자에게까지 중단되지 않고 전달된 신호.
- **Reverberant signal:** 소리 재료에서 청취자에게까지 간접적으로 도달한 신호. 일반적으로 몇 만 정도의 반향으로 이루어짐. 하나의 공간 안에서 여러 표면으로부터 반사되어 청취자에게 전달된 신호

• 두 신호의 강도

- Direct sound의 경우 청취자와 사운드 소스의 거리가 증가하면 역 제곱의 법칙에 따라 강도 감소
- Reverberant energy의 경우 청취자와의 거리와 상관없이 변화량 거의 없음  
 ∴ 청취자는 direct sound의 강도와 reverberant energy의 강도 관계에 의해 소리 재료의 거리 판단

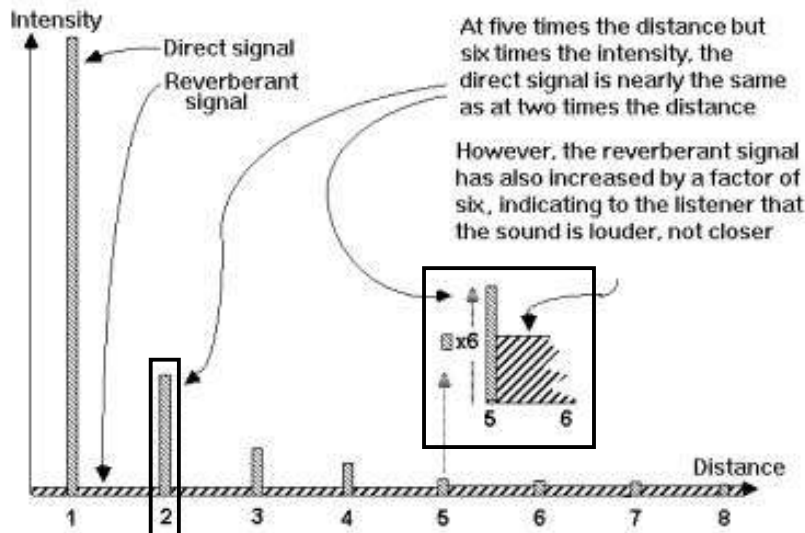


Figure 20.7 Direct signal과 Reverberant signal의 거리에 따른 강도 비교

예) 5의 거리에서(20.7) 2가 가진 강도의 6배를 가진 소리가 만들어질 경우 청취자는 두 소리에서 동일한 direct signal 인식. 하지만 reverberant 신호의 경우 신호의 강도에 따라 5의 거리에 있는 소리의 reverberant energy는 여섯 배로 증가. 이런 이유로 청취자는 5와 2의 거리에 있는 소리 재료를 혼돈하지 않음

- Reverberant signal은 청취자가 소리 재료의 거리를 가늠하는 단서 제공

• Loudness constancy

- 일정한 강도를 가진 소리는 청취자와의 거리가 증가하더라도 일정한 강도로 인식
- 스펙트럼 변화가 없을 때 이러한 loudness constancy는 reverberant energy에 의해 생성



- 시각에서의 size constancy:  
Size constancy - 크기를 판단하는데 있어 망막에 인식된 상의 크기와 직접 인식된 크기가 항상 일치하는 것은 아님
- Loudness constancy와 size constancy의 차이:  
시각의 경우 짧은 시간 안에 실제 크기가 증가하거나 감소하는 것으로 생각되는 상 (image)은 많지 않음.  
청각의 경우 짧은 시간 안에 다양한 loudness변화 발생 가능
- 청각에서의 원근법은 시각과의 관계에 대한 은유가 아닌 공간 인식에 관한 일반적 법칙
- 청각의 원근법은 강도, 스펙트럼, 거리 등을 포함한 다양한 요소의 물리적 연관성에 의해 만들어짐
- 또한 loudness 인식은 chorus effect, vibrato depth, vibrato rate에 영향 받음

## 20.4 Conclusion

- 컴퓨터를 이용함으로써 loudness는 (intensity외의)다양한 dimension에서 더욱 정교하게 또 더욱 음악적으로 제어 가능
- Loudness의 이러한 dimension이 음악적으로 지나치게 강조될 수는 없기 때문에 컴퓨터나 synthesizer 이용에서 이러한 내용이 아직까지 일반적이지 않음
- 자연의 소리와 유사한 소리를 만들기 위해서가 아닌 소리의 범위를 넓히기 위해
- 마술적인 음악과 음향의 경계를 제안하도록 하기 위해  
(우리의 일반적인 음향 경험 외적 부분)
- 우리의 인식/인지 체계에 적합한 방법으로 표현하기 위해