

22. Storage and Reproduction of Music

John Pierce

22.1 Introduction

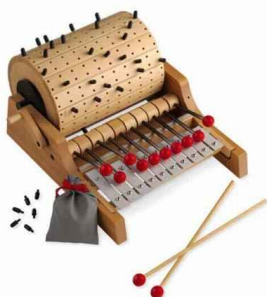
- 음악의 재생산(reproduction)
 - 1) 음악가의 기억에 의한 연주
 - : 연주 당시의 환경과 조건에 의해 연주가 달라질 수 있음, 연주자의 기억의 한계
 - 2) 악보에 의한 연주
 - : 연주자의 곡 해석에 의존, 기억의 한계는 없으나 악보의 비용이 존재
 - 3) 음악을 연주하거나 재생(replay)하는 기계에 의한 연주
 - 전송된 음악(TV, Radio), 녹음된 음악(Tape, CD)
 - : 기계 자체의 비용이 존재하며, 정보를 저장하기 위한 매체에 대한 비용 존재

22.2 Scores

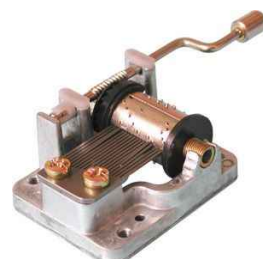
- 숙련된 연주자에게는 매우 훌륭한 음악 재생산의 도구
 - Key Signature, Clef Marks, Note Duration, Time Signature, Tempo, Timber, 작곡가 및 곡제목, 셈여림의 표시 등
 - 연주자의 의도 및 곡 해석에 따라 결과물이 달라짐

22.3 Mechanical Instruments

- 레오나르도 다 빈치에 의해 발명된 자동으로 연주되는 악기가 존재했다고 전해짐
 - : Spinet과 2개의 드럼으로 돌출부가 있는 회전 실린더에 의해 연주됨
- 15세기, 시계탑에 설치된 깔리용(Carllion)들은 자동으로 연주되었음
- 휴대용 풍금(Barrel Organ)은 현재까지도 사용되고 있는 자동연주악기
 - : Haydn, Mozart, Bach등이 휴대용 풍금을 위한 곡을 만들기도 했음
- Music Box는 18세기 말 정도에 발명되었음



Barrel Organ



Music Box

- Player Piano가 자동연주악기 중에서는 가장 음악적
 - : 오늘날 사용하는 MIDI Piano는 녹음된 연주를 통해 연주자의 세밀한 뉘앙스까지도 표현할 수 있으나, 그 곡이 실제로 연주될 당시의 공간감 등은 표현할 수 없음
 - : MIDI Interface의 전송 대역폭은 31,250 bps 정도로, CD의 전송대역폭인 11,412,000 bps에 비해 약 45분의 1의 크기

22.4 Imitation Instruments

- 필자의 정의: 실제 악기의 소리에 효과를 주는데 사용되는 전자 하드웨어나 소프트웨어
- Vocoder: Homer Dudley에 의해 1939년에 발명. 인공적인 음성을 생성해낼 수 있음.

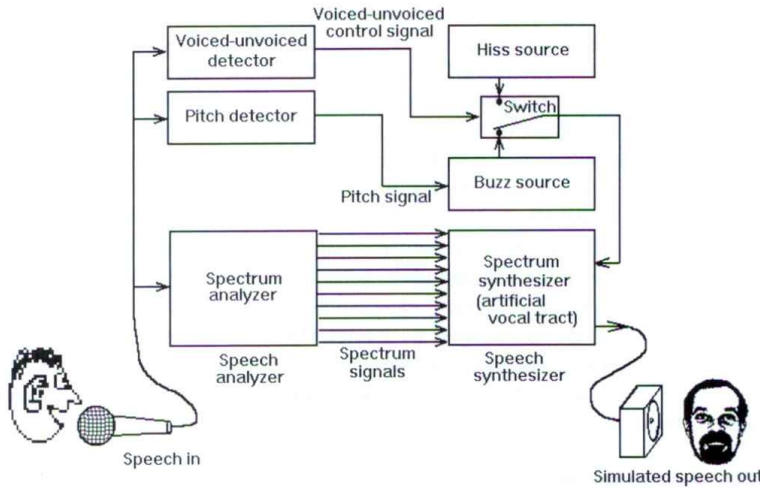


Figure 22.2 The channel vocoder. Filters decompose the voice for coding, transmission, and reconstruction. Systems like this are used for security as well as communication. (Modified after Dudley, 1939.)

- 1) 송신 단말에서 유성, 무성음을 구분하고, 유성음인 경우 음고를 추출
- 2) 소리의 스펙트럼을 적당한 수의 밴드패스 필터를 통해 측정
- 3) 수신 단말에서 밴드패스 필터의 게인을 조절하여 음을 합성
- 4) 필터로 입력되는 신호는 노이즈 또는 음고를 가진 신호로 유, 무성음의 구분에 따라 여기(excited)됨: 이 신호를 바꾸어서 다양한 소리들이 말하는 것처럼 변형될 수 있음 (cross-synthesis)

- Vocoder는 본래 보안시스템을 위해 만들어 졌으며, 실용적으로 활용되는 Vocoder는 대부분 선형예측 Vocoder(LPC)이다.
: 선형예측은 디지털 오디오의 전송량을 최소화하는데 적격으로, 음성 통신 분야에서 널리 사용되고 있음
- Phase Vocoder - 컴퓨터 음악에서 많이 사용됨
- 솔로 악기를 흉내 내거나 조작하는 관계적 시스템의 제안(Serra and Smith 1990)
: 악기의 파형을 Deterministic Part와 Stochastic Part로 분리하여 고려
- Deterministic Part는 소리의 분석을 통한 데이터를 이용하여 재현 가능
- Stochastic Part는 Noise를 통해 재현 가능
: 2 부분의 조합으로 원래의 소리를 재합성(Resynthesis)

22.5 Listening to Sounds

- Pop Music의 경우 최종 스테레오 채널의 소리는 다수의 모노 채널 마이크로부터 녹음된 소리들로 구성되어 있음

- 클래식 음악은 일반적으로 스테레오 혹은 그 이상의 채널의 마이크로부터 녹음된 소리로 최종 스테레오 채널이 구성됨
- 재생 장치의 중요한 조건
 - 충분한 음량 레벨에 이르기까지의 전달함수가 선형을 이룰 것
 - 주파수 응답이 평탄할 것
- 녹음된 소리 그대로를 재생할 수 있는가가 중요한 것이 아니라, 녹음된 원래의 소리를 사람이 구분하지 못하는 범위 내에서 보다 적은 정보량으로 표현할 수 있는가가 중요함

22.6 What We Cannot Hear

- CD의 16Bit로 표현 혹은 저장이 가능한 음량의 범위는 90dB 내외임
- Equal Loudness Curve와 Bit Rate Reduction

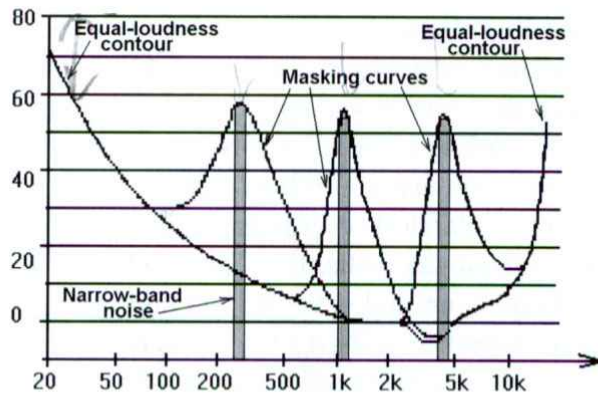


Figure 22.5 An equal loudness curve and the threshold of hearing for three masking curves. (Adapted from Zwicker, 1990.)

- 1000Hz의 가청한계음량은 30Hz의 가청한계음량보다 약 60dB작음
 - : 즉, 30Hz의 음들은 가청한계로부터 최고음량까지의 범위를 표현하기 위해 6bit를 사용하며, 1000Hz는 16bit를 사용해야 함
 - : 이러한 원리로 초저음역, 초고음역의 표현을 위한 디지털 데이터를 감소시킬 수 있음
- Masking Effect
 - 스펙트럼 내에서 강한 에너지를 갖는 Peak를 중심으로 그 주변의 낮은 음량을 갖는 소리들은 잘 들리지 않게 되는 효과
 - : 그림 22.5의 250Hz, 1000Hz, 4000Hz의 Peak는 그 주변의 가청한계음량을 다소 상승시킴
 - Masking Effect를 이용한 Bit Reduction의 예
 - : 700Hz의 범위를 갖는 32개의 주파수 하위대역(subband)로 분할한 뒤,
 - 1) 하위대역의 에너지가 Masking Threshold보다 낮아지게 되면 Bit를 할당하지 않음
 - 2) Masking Threshold 이하의 음량을 가지는 양자화 노이즈를 생성하는 만큼의 Bit를 할당함

; 일반적으로 90dB를 위해서는 16Bit가 필요하며, 이 경우 양자화 노이즈는 90dB (6N)보다 클 수 없음. 이러한 이유로 어떠한 소리를 5bit로 인코딩하게 되면 최대 30dB의 양자화 노이즈가 발생하지만, 이 양자화 노이즈의 음량이 Masking Threshold 보다 하위에 존재한다면 무시할 수 있음

- Time Domain내에서의 Masking: Backward Masking, Forward Masking
 - 큰 신호 앞, 뒤의 작은 신호는 Masking됨 (각각 Backward, Forward)
 - : Backward Masking은 매우 짧은 시간의 범위(수 ms)내에서만 유효
 - : Forward Masking은 다소 긴 시간의 범위(수백 ms)내에서 유효
 - 5~6 bit per sample 수준의 감소에서도 거의 인식할 수 없으며, 보다 적은 수의 bit를 사용함에도 비교적 우수한 음향을 얻을 수 있음(왜곡이 발생하는 프로세스이지만, 그 왜곡을 사람이 인식할 수 없음)

22.7 In Retrospect

- 음악 작곡의 기록은 연주자 혹은 기계적, 전자적 수단(Player Piano, DiscClavier)로 하여금 곡을 재창조하도록 함
- 인간의 음성과 비건반악기류들의 음합성기법을 통한 재현이 만족스러운 수준은 아니지만, 음악적으로 유용한 결과들로 인해 널리 사용되고 있음
- 대안은 음향을 포착한 후 그것을 어떻게든 재생하는 것: 아날로그 및 디지털 레코딩
- 클래식 음악의 스테레오 레코딩을 들을 때, 우리는 녹음에 사용된 마이크의 위치로부터 녹음되는 소리를 듣게 되며, 이것이 스테레오 레코딩 및 레코드의 한계임
- 스테레오 혹은 멀티채널 레코딩의 과제는 파형을 재생하는데 필요한 데이터의 양을 줄이는 것: 정확한 원래 파형의 재생이 목적이 아니라, 우리의 귀가 원본과 구분할 수 없는 새로운 파형을 만들어내는 것이 목적