

7주차 1차시 사운드 기본요소의 이해

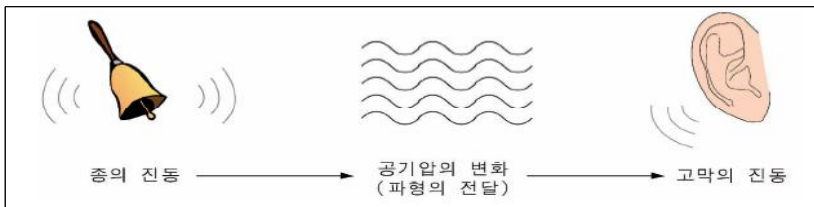
【학습목표】

1. 사운드의 기본 개념 및 사운드의 3요소를 설명할 수 있다.
2. 사운드의 종류와 사운드의 제작 과정을 설명할 수 있다.

학습내용1 : 사운드의 기본요소

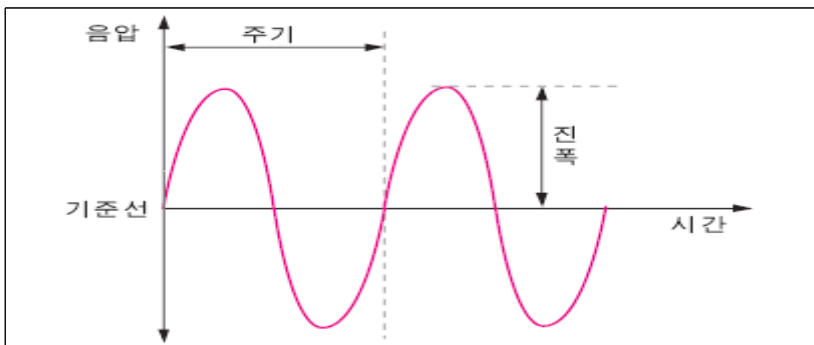
1. 사운드의 기본 개념

- 사운드는 물체가 진동하여 만들어진다.
- 진동이 발생할 때 그 주위의 공기압에 변화가 생기고 공기압의 변화가 파형(Waveform)의 형태로 우리 귀에 전달된다.
- 우리가 보통 사운드를 처리한다는 것은 이러한 파형을 가공, 편집하는 것이다.



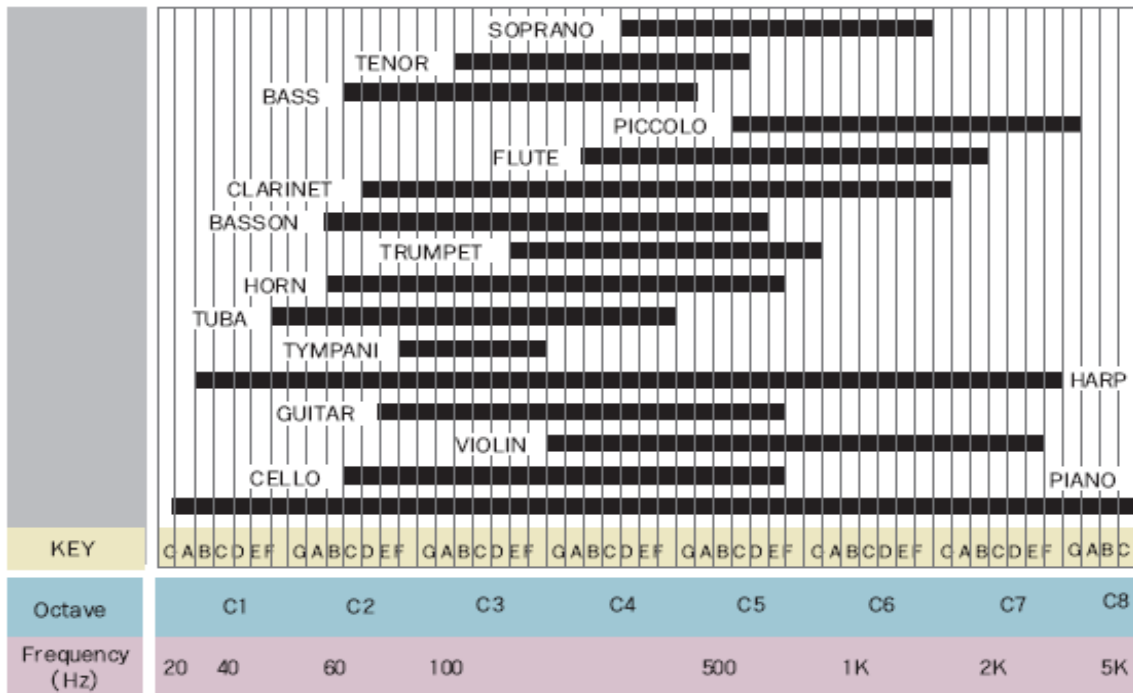
2. 사운드의 기본 요소

- 사운드의 기본 요소==주파수(Frequency), 진폭(Amplitude), 음색(Tone Color)
- 사운드 파형은 일정한 시간 간격마다 동일한 모양으로 반복되는데, 동일한 부분을 사이클 (Cycle)이라 함
- 한 사이클이 걸리는 시간을 주기(Period)라고 한다.
- 사운드는 요소 3가지
 - 음의 높낮이와 관련이 있는 주파수(Frequency)
 - 음의 크기와 관련이 있는 진폭(Amplitude), 그리고
 - 각 음의 특성인 음색(Tone Color)의 세 가지 요소로 구성



가. 주파수(Frequency)

- 주파수는 초당 사운드 파형의 반복 횟수를 의미하며 주기와 역수 값의 관계가 있으며 소리의 높낮이를 결정한다.
- 주파수가 크면 고음이고 작으면 저음이 된다.
- 측정 단위로는 Hz가 사용되며 이것은 초당 진동 횟수를 의미한다.
- 일반적으로 사람이 낼 수 있는 주파수 대는 약 100Hz ~ 6KHz 이다.
- 가청 주파수대는 약 20Hz ~ 20KHz 사이이고, 사람의 청각은 1KHz ~ 6KHz에서 가장 민감하다.
- 사람의 가청 주파수대를 오디오(Audio)라고 구분하며, 오디오를 처리하는 것이 사운드를 처리하는 것이라 생각하는 사람도 있다.
- 각 악기들의 주파수 대역은 아래 그림과 같다.



나. 진폭(Amplitude)

- 사운드 파형의 기준선에서 최고점까지의 거리를 의미하며 소리의 크기와 관련이 있다.
- 소리의 크기는 에너지로 측정하지 않고 음압의 변화로 나타낸다.
- 음압을 표현하는 단위로 bar라는 것이 있다.
- 사람의 귀는 아주 작은 음압의 차이라도 민감하게 감지할 수 있기 때문에 바(bar)의 백만분의 1인 μbar 를 사용한다.
- 사람의 귀가 음파를 소리로 감지할 수 있는 최소 음압은 0.0002 μbar 이며 최대는 200 bar이다.
- 사람의 청각은 음압에 비례해서 소리의 크기를 느끼지 못한다.
- 소리의 크기가 1에서 2로 변했을 때와 2에서 4로 변했을 때 동일한 소리크기의 변화가 있는 것으로 느낌
- 1에서 2, 2에서 3으로 변했을 때에는 같은 1만큼의 변화이지만 2에서 3으로 변화한 쪽이 더 작은 변화인 것으로 느낀다.
- 이처럼 소리에 대한 감각은 변화의 차이에 관계있는 것이 아니라, 변화의 비율 더 정확히는 로그 비(Logarithm Ratio)에 따라 정해진다.
- 소리의 크기(음압)를 표시할 때에는 소리의 크기에 로그값을 취한 벨(bel)로 표시한다.
- 사용 시에는 벨 값을 10배한 데시벨(dB: decibel)을 주로 사용한다.

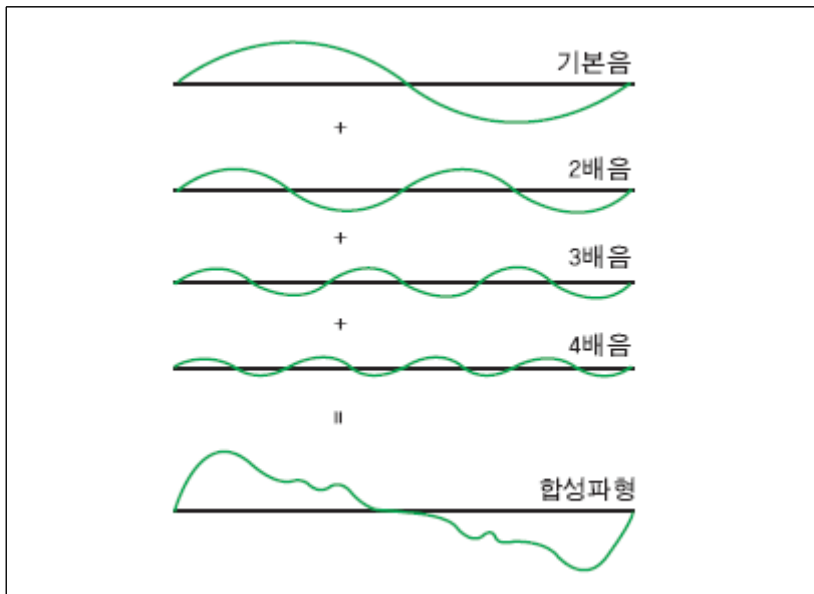
$$\text{음압 레벨(dB)} = 10 \times 2 \log \frac{P}{P_0}$$

- 기준 음압을 $P_0(0.0002\mu\text{bar})$, 현재 음압을 P 라 하면
- 사람이 가장 편하게 들을 수 있는 소리의 범위는 0dB~90dB 사이이다.

음의 크기(dB)	소리의 예
160	가까운 곳에서 들리는 제트기의 소리
140	청취하기에 고통스러운 소리
120	공항 활주로에서 들리는 소리
100	지하철에서의 소음
80	일반적인 공장에서의 소음, 번잡한 거리에서의 소음
60	일상적인 대화 소리
40	나직한 대화 소리
20	조용한 거실에서 들리는 소리
0	최저 가청음

다. 음색(Tone Color)

- 같은 음의 높이와 크기를 가져도 악기마다 고유한 소리의 특징(기본파와 고조파가 합성된것)이 있다. 이러한 특징을 음색이라고 한다.
- 악기에는 각 악기마다 고유의 주파수(기본파)가 있다.
- 이 기본파는 각 악기마다의 특성인 울림통에서 진동하며, 기본파 외에 그 악기의 기본파의 정수배에 해당하는 고조파라는 것을 생성하게 된다(아래 그림).
- 기본파와 고조파가 합성이 되어, 그 악기의 고유한 소리가 생성된다.



학습내용2 : 사운드의 개요

1. 사운드의 분류

- 사운드의 분류 : 음악, 음성, 음향효과의 세 가지로 나눌 수 있다
- 이 세 가지 중 음악과 음성(대화)은 단독으로 사용될 수 있지만, 멀티미디어 환경에서는 이 세 가지가 복합적으로 다른 미디어와 함께 사용되는 것이 일반적이다.
- 오감 : 사람이 가지는 감각은 크게 시각, 청각, 후각, 촉각, 미각의 다섯 가지가 있다.
- 오감 중에서 가장 많이 사용되는 정보 전달수단이 시각과 청각이다.
- 시각 정보가 중심이 되는 경우는 책, 그림 등이 있고,
- 청각 정보가 중심이 되는 경우는 대화, 연설, 음악, 뉴스 등이 있다.
- 그리고 비디오와 애니메이션에서는 이 두 가지 정보가 함께 사용되고 있다.
- 컴퓨터에서는 주로 이 두 가지 감각을 이용하고 있고,
- 사운드는 그 중 하나이므로 컴퓨터를 이용한 정보 전달에서도 매우 중요한 역할을 차지한다.
- 다른 측면에서 볼 때, 멀티미디어 정보에서 사운드를 사용하는 이유는 정보전달 시 미디어의 상승효과와 동기유발에 있다.
- 정보를 전달할 때 시각적 방식과 청각적 방식을 동시에 사용하여 정보를 전달하면 정보전달 효과가 커지는 것으로 알려져 있다. 또한 사운드를 사용하게 되면 정보를 전달받는 사람의 동기나 흥미를 유발시켜 멀티미디어 정보를 효과적으로 전달할 수 있다.

2. 사운드의 종류

- 분류1 : 사운드는 일반적으로 음성(Voice), 음악(Music), 음향효과(Sound Effect)의 세 가지로 분류할 수 있고,
- 분류2 : 컴퓨터에서 처리하는 방법에 따라서는 디지털 오디오(Digital Audio)와 미디(MIDI)로 분류할 수 있다.

가. 일반적인 분류---음성(Voice), 음악(Music), 음향효과(Sound Effect)

① 음성(Voice)

- 음성은 정보를 전달하는 주요한 수단 중 하나이다. 컴퓨터에서 음성을 사용하면 텍스트(Text)를 사용하는 경우보다 정보를 더 빠르고, 이해하기 쉽고, 설득력 있게 전달할 수 있다.
- 컴퓨터에서 사용하는 음성은 디지털화된 음성과 합성된(Synthesized)음성으로 나눌 수 있는데,
- 디지털화된 음성은 사람 음성에 가깝지만 많은 용량을 차지하고,
- 합성된 음성은 용량은 작지만 자연스럽게 들린다. 우리가 SF영화 같은데서 들을 수 있는 로봇 음성은 합성된 음성의 좋은 예이다.

② 음악(Music)

- 음악은 감상용으로 사용하는 경우도 있지만, 멀티미디어 정보를 전달하는데 부수적인 효과를 제공한다.
- 예를 들어, 분위기나 장면의 전환, 감정의 고조 등을 나타내는데 사용된다.
- 특히 텍스트나 이미지와 같은 시각적인 정보를 전달할 때, 음성과 음향효과를 같이 사용하면 효과가 더욱 커지게 된다.

③ 음향 효과(Sound Effect)

- 음향 효과는 단독으로 사용되기보다는 정보를 전달할 때 강조하거나 보조하는 수단으로 사용된다.
- 음향 효과는 자연적인(Natural) 것과 합성한(Synthetic) 것으로 나눌 수 있다.
- 자연적인 것은 일상적인 주변에서 일어나는 소리를 의미
- 합성한 음향효과는 인공적으로 만들어진 소리를 의미한다.
- 음향 효과는 배경 효과로 사용되어 특정 장소나 상황을 좀 더 실감 있게 전달할 수 있고(비오는 상황이나 공사장 등), 어떤 특정한 정보를 나타내기 위해 사용될 수 있다(영화에서 전화 오는 소리나, 보이지 않는 상황에서의 그릇 깨지는 소리 등).

나. 컴퓨터에서 처리하는 방법에 따른 분류===디지털 오디오(Digital Audio)/미디(MIDI)

① 디지털 오디오(Digital Audio)

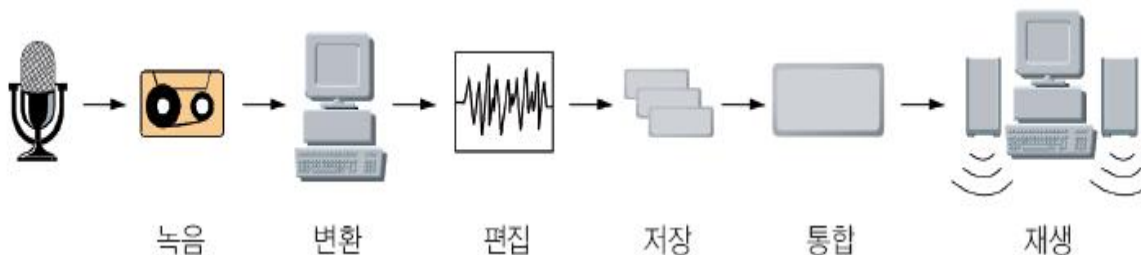
- 디지털 오디오는 아날로그 파형(Analog Wave)을 디지털 형태로 변환한 것
- 원음에 가깝게 재생하기 위해서는 많은 정보가 필요하다.
- 우리가 듣는 음악 CD의 수준으로 디지털 오디오를 만들려면 3분 정도의 음악을 위해 약 30MB 정도가 필요
- 디지털 오디오는 사람의 음성이 존재하거나, 디지털화된 소리를 처리할 수 있는 컴퓨터 환경이 지원될 때, 또는 연주하는 각 악기에 대한 제어를 할 수 없는 경우에 사용된다.
- 디지털화된 소리를 처리하기 위해서는 메모리나 하드디스크의 용량이 크고 CPU의 성능도 좋아야 한다.

② 미디(MIDI: Musical Instrument Digital Interface)

- 미디의 정의 : 전자 악기와 컴퓨터 간에 정보를 전송하기 위해 만든 통신 프로토콜
- 음악을 표현하기 위해 사용됨
- 미디는 음 자체에 대한 파형정보를 가지고 있는 것이 아니라 음을 연주하는 방법과 연주시기 등에 대한 정보를 가짐
- 음 자체에 대한 정보를 가지고 있지 않기 때문에 파일 크기는 작고,
- 음의 질은 실제 음에 대한 정보를 가지고 있는 음원의 영향을 받는다.
- 3분 정도의 미디 음악을 듣기 위해서는 약 8 KB가 필요하다.
- 미디는 사람의 음성이 존재하지 않고,
- 연주할 각 악기에 대한 제어가 가능하고,
- 좋은 음원 모듈을 가지고 있을 경우에 이용된다.
- 특히 음악에 대해 어느 정도의 지식을 가진 사람은 미디 프로그램을 통해 손쉽게 음악을 작곡, 편곡할 수 있다.

3. 사운드의 제작 과정

그림은 디지털 사운드의 제작과정을 간략하게 표현한 것이고, 실제적으로 각 단계에서 일어나는 작업은 다음과 같다.



가. 계획(Plan)

어떤 작업을 할 것인지에 대한 계획을 세우고, 필요한 조건들을 검토한다. 그리고 최종 결과에서 사용될 사운드에 대한

설계를 한다.

나. 녹음(Record)

사람의 음성이나 현실 세계의 소리를 사운드 도구를 이용하여 녹음한다. 컴퓨터를 이용하는 경우에는 녹음과 함께 디지털로의 변환이 동시에 이루어진다. 이렇게 직접 녹음하는 것이 사용하고자 하는 목적에 가장 잘 부합되지만, 인터넷에서 음향 효과를 모아 놓은 사이트나 사운드 클립(Sound Clip) CD를 이용할 수도 있다. 이 경우에는 사운드가 일반적으로 디지털 형태로 존재하기 때문에 변환 과정이 필요 없다. 그러나 이러한 경우에도 목적에 완전히 맞는 경우는 드물기 때문에 편집과정을 거쳐 목적에 맞게 바꾸어 주어야 한다.

다. 변환(Capture)

기존의 카세트테이프나 CD 등에 녹음되어 있는 음성이나 음악 등을 컴퓨터에서 사용하기 위하여 디지털 사운드로 변환한다. 일반적으로 아날로그 파형은 디지털로 변환되나 이미 디지털로 되어 있는 경우는 전송만 하면 된다.

라. 편집(Edit)

원하는 목적에 맞게 사운드 편집 프로그램을 이용하여 편집(Edit) 또는 믹싱(Mixing)한다. 예를 들어, 녹음한 것 중에서 제일 좋은 것을 선택하여 원하는 부분만을 뽑아내어 에코(Echo) 효과 등을줄 수 있다.

마. 저장(Store)

편집이 끝난 중간 결과를 원하는 포맷으로 저장한다. 다운샘플링(Down Sampling: 음의 원래 표본화율이나 해상도를 낮추는 것)하여 저장할 수 있고, 다른 프로그램과 호환이 되게 하기 위한 포맷으로 저장할 수도 있다.

바. 통합(Integration)

최종 결과를 만들기 위해 저장된 사운드들을 다른 프로그램(예를 들어, 동영상이나 프리젠테이션 프로그램 등)과 통합한다.

사. 재생(Playback)

최종 결과를 재생하여 보완될 점을 고친다.

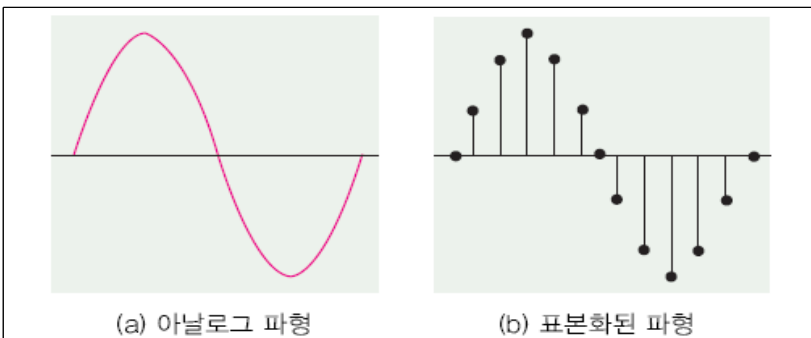
학습내용3 : 디지털 사운드로의 변환

- 일상적인 사운드는 아날로그 형태인데, 컴퓨터에서 처리하기 위해서는 디지털의 형태로 변환되어야 한다.
- 이러한 작업은 그림에서 보듯이 ADC(Analog-to-Digital Converter)장치에서 이루어지는데,
- 표본화(Sampling) 및 양자화(Quantizing) 과정을 거쳐서 디지털 형태로 출력이 된다.
- 실세계에서 듣기 위해서는 반대의 과정인 DAC (Digital-to-Analog Converter)를 거쳐 아날로그 형태로 바뀌어야 한다.

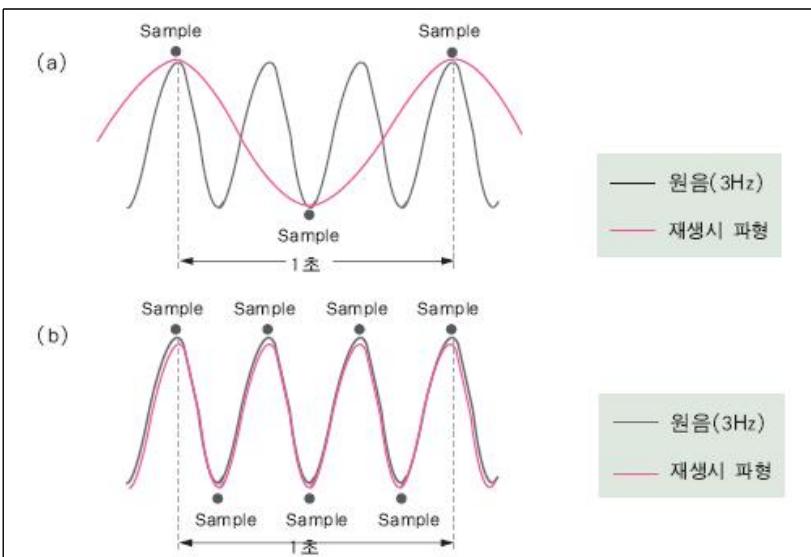


1. 표본화(Sampling)

- 표본화란 아날로그 파형을 디지털 형태로 변환하기 위해 표본을 취하는 것을 말한다(그림).
- 사운드의 표본화율(Sampling Rate)은 1초 동안에 취한 표본수(디지털화하는 횟수)를 의미한다.
- 단위로는 주파수 단위와 같은 Hz를 사용한다.



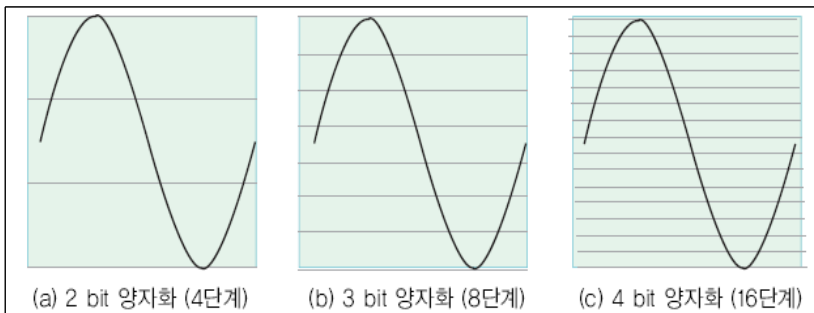
- 표본화율이 높을수록 원음에 가까운 음으로 디지털화 되지만 데이터 양이 증가하게 된다.
- 실제로 표본화하는 예를 보면 다음과 같다.



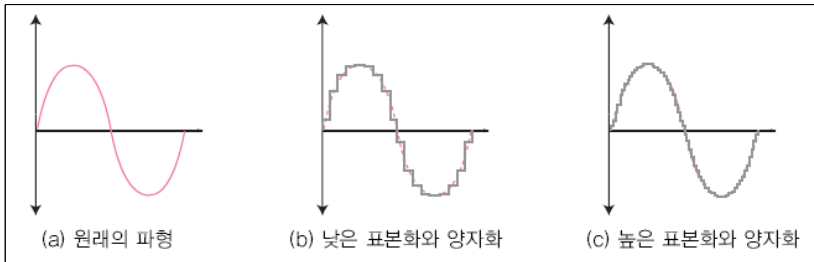
- 그림 (a)는 3Hz인 원음을 2Hz로 표본화한 것으로 표본화 되는 샘플의 중간 중간의 음이 반영되지 않는다.
- 그러나 그림(b)에서는 원음의 두 배인 6Hz로 표본화한 것으로 중간의 음도 표현이 되어 원음에 더 가깝게 표현할 수 있다.
- 표본화를 많이 할수록 원음을 잘 표현할 수 있으나 데이터 저장을 위한 메모리 용량을 많이 차지
- 표본화를 작게 하면 메모리 용량은 작아지지만 원음을 그대로 반영하지 못하게 된다.
- 원음을 재생하기 위한 최소한의 표본화율을 알아야 되는데, 이는 1920년대 후반 나이퀴스트(Harry Nyquist)에 의해 밝혀졌다.
- 나이퀴스트 정리(Nyquist Theorem)
- 표본화를 할 때 원음을 그대로 반영하기 위해서는 원음이 가지는 최고 주파수의 2배 이상으로 표본화해야 한다.
- 나이퀴스트 정리가 실제로 적용된 예로 우리가 많이 듣고 있는 음악 CD의 경우 표본화율은 44.1KHz이다.
- 그러므로 여기서 재생할 수 있는 최고 주파수는 22.05KHz가 된다.
- 음악CD에서 44.1KHz 주파수로 표본화하는 이유는 사람의 귀가 20KHz 이상의 사운드는 감지할 수 없기 때문이다.

2. 양자화(Quantizing)

- 양자화는 그 값을 얼마나 자세하게 표현할 것인가와 관련이 있다.
- 표본화된 각 점에서의 값을 표현하기 위해서는 그 값을 어느 정도의 정밀도로 표현할 지를 결정해야 한다.
- 이러한 정밀도는 바로 그 값을 표현하는 비트 수와 연관된다.
- 예를 들어, 8bit로 양자화를 하면 256단계로 값을 표현할 수 있지만, 16bit로 양자화를 하면 더 세밀한 65,536단계로 값을 표현할 수 있다.



- 위의 그림은 값을 표현하는데 각각 2bit, 3bit, 4bit로 표현한 예이다.
- 예를 들어, 우리가 듣는 음악 CD의 경우는 16bit로 값을 표현하고 있다.
- 이것은 음을 65,536단계로 표현하고 있다는 것을 의미한다.
- 이때 표본화하는 정밀도를 음의 해상도(Sampling Resolution 또는 Sampling Size)라고한다.
- 양자화를 할 때 원음(아날로그 음)에 포함되는 잡음(Noise)의 성분보다 높은 정밀도를 사용하는 것은 의미가 없다.
- 그 이유는 잡음이 원음에 포함되어 있는 것으로 간주되어 잡음까지 양자화되어 표현되기 때문이다.
- 낮은 표본화와 양자화를 거친 것과 높은 표본화와 양자화를 거친 디지털 파형을 비교해 보면 아래 그림과 같다.



- 위의 그림을 비교해 보면 높은 표본화와 양자화를 거친 것이 원음에 가깝게 표시가 되는 것을 알 수 있다.

3. 부호화(Coding)

- 표본화와 양자화를 거치고 나서 그것을 실제 디지털 정보(Bit Stream)로 표현을 해야 한다. 이러한 과정을 부호화라고 한다. 사운드 파일은 파일의 크기가 크기 때문에 부호화하는 과정에서 일반적으로 압축하여 저장하게 된다.

【학습정리】

1. 사운드는 일반적으로 음성(Voice), 음악(Music), 음향효과(Sound Effect)의 세 가지로 분류한다.
2. 컴퓨터에서 처리하는 방법에 따른 사운드의 분류는 디지털 오디오(Digital Audio), 미디(MIDI: Musical Instrument Digital Interface)로 나뉜다.
3. 사운드는 음의 높낮이와 관련이 있는 주파수(Frequency), 음의 크기와 관련이 있는 진폭(Amplitude), 그리고 각 음의 특성인 음색(Tone Color)의 세 가지 요소로 구성된다.
4. 디지털 사운드로의 변환은 표본화와 양자화를 한 이후에 그것을 실제 디지털 정보(Bit Stream)로 표현하는 과정을 부호화 과정을 거친다.