



자료구조 with Python

게임소프트웨어과 2학년

3. 컴퓨터의 정보 처리 방법

1

자료와 정보

1. 컴퓨터의 자료 처리

- ✓ 컴퓨터는 이진 코드(Binary code)를 이용하여 전기 신호의 상태(On/Off) 정보를 1(On) 또는 0(Off)으로 표현한다.
- ✓ 컴퓨터는 모든 데이터를 이진 코드로 변환하여 처리한 후 다시 이를 사람이 이해할 수 있는 형태로 변환하여 나타내 주어야 한다.

2. 인코딩과 디코딩

- ✓ 인코딩(Encoding)은 자료를 컴퓨터에 저장하거나 통신에 사용할 목적으로 부호화 (이진 코드 변환) 하는 것
- ✓ 디코딩(Decoding)은 부호화된 자료를 원래대로 되돌리는 것
- ✓ 우리가 컴퓨터에 내리는 명령들은 인코딩과 디코딩의 과정을 수행하여 컴퓨터가 명령을 처리하고, 그 결과를 우리가 확인할 수 있는 것이다.

1-2

효율적인 디지털 정보의 표현



교내에서 학급 홍보 자료 경연 대회가 열린다. 우수 작품으로 선정되면 표창을 받고 간식 파티를 할 수 있다. 혜원이네 반 학생들은 학급 홍보 자료를 제작하기 위해 필요한 자료와 표현 방법에 대해 알아보기로 했다.



- (1) 학급 홍보 동영상에 어떤 형태의 자료가 들어갈 수 있을까?
- (2) 여러 유형의 자료를 어떻게 만들고 편집할 수 있을까?
 - ① 자료 유형별로 적합한 파일 형식에 대해 조사해 보자.
 - ② 자료 유형별로 필요한 편집 도구를 조사해 보자.

02 자료의 종류와 표현

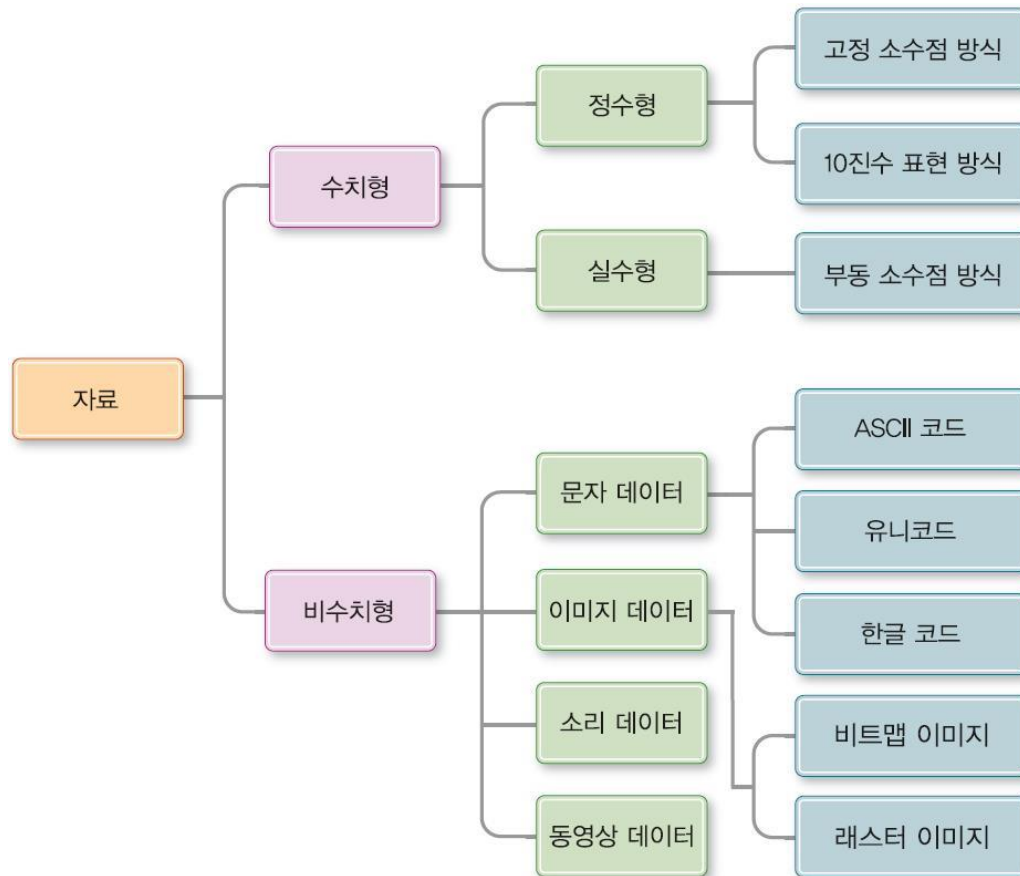
학습 목표

- ❖ 자료의 종류를 알 수 있다.
- ❖ 컴퓨터에서 자료를 표현하는 방법을 알 수 있다.

02 자료의 종류와 표현

1 수치 데이터의 표현

1. 자료의 분류



[그림 1 - 06] 자료의 분류

02 자료의 종류와 표현

1 수치 데이터의 표현

1. 자료의 분류

- **무형성:** 정보는 일정한 형태를 지니지 않으며 유동적이다.
- **독점성:** 정보는 공개되는 순간 그 가치가 급격히 떨어진다.
- **적시성:** 정보는 수요자가 원하는 시점에 바로 전달되어야 한다.
- **비소모성:** 정보는 상품과 달리 소유의 개념이 아니라 사용의 개념을 가진다.
- **상대성:** 같은 정보도 사용자가 어떤 목적으로 이용하느냐에 따라 상대적으로 다른 가치를 지닌다.

1. 숫자와 문자의 표현

1

자료와 정보

1. **숫자와 문자**는 우리가 가장 많이 사용하는 자료 형태이다.
2. **수의 표현**은 **이진수**를 이용한다.
 - ✓ 컴퓨터가 표현할 수 있는 수의 범위를 초과하는 연산을 수행하는 경우 잘못된 값을 출력할 수 있으므로 주의해야 한다.
3. **문자의 표현**은 아스키 코드, 유니 코드와 같이 각 문자에 해당하는 값을 미리 정의해 놓은 **표준 코드**를 사용한다.
4. 표현하고자 하는 문자에 따라 적절한 인코딩 방식을 사용해야 한다.
 - ✓ 7비트 크기로 문자를 표현하는 아스키 코드로는 다양한 언어의 문자를 표현할 수 없기 때문에 유니코드를 사용하게 되었다.
 - ✓ 유니코드를 문자로 표현하기 위한 인코딩 방식은 UTF-8, UTF-16, UTF-32 등이 있다.
 - ✓ 웹 브라우저 등에서 인코딩 방식을 잘못 선택하면 한글 등의 문자들이 깨져서 나타나는 경우가 있다.

02 자료의 종류와 표현

1 수치 데이터의 표현

2. 자료의 표현 단위

[표 1 - 01] 자료의 표현 단위

비트(bit)	정보 표현의 최소 단위로, 2진수 0 또는 1을 나타낸다.
니블(nibble)	1/2바이트를 의미하며, 1니블은 4비트이다.
바이트(byte)	컴퓨터에서 사용되는 자료의 기본 저장 단위로, 1바이트는 2니블을 의미하며, 1바이트는 8 비트이다.
워드(word)	컴퓨터에서 CPU가 한 번에 처리하는 명령어의 기본 단위이다. 1워드는 4바이트 또는 full word라고 하고, 1/2워드는 2바이트로 half word라고 한다.(단, 1워드가 항상 4바이트는 아니다. 2바이트 또는 8바이트를 1워드로 하는 시스템도 있다.) 1워드 = 4바이트 = 32비트
필드(field)	파일 구성의 최소 단위로 항목(item)을 의미한다.
레코드(record)	하나 이상의 필드들이 모여 구성된 자료 처리 단위이다.
파일(file)	여러 개의 레코드가 모여 구성되며, 디스크의 저장 단위이다.
데이터베이스	자료를 편리하고 효율적으로 사용할 수 있도록 저장한 파일 형태의 집합이다.

02 자료의 종류와 표현

1 수치 데이터의 표현

3. 정수형 데이터의 표현

- 고정 소수점 표현 방식

- 부호와 절대치 표현법
- 부호와 1의 보수 표현법
- 부호와 2의 보수 표현법

- 10진수 표현 방식

- 팩 10진 형식
- 언팩 10진 형식

○○○ 컴퓨터는 왜 2진수를 사용할까? ○○○

- ▶ 전기적 신호로 의미를 전달 할 수 있는 가장 간단한 방법이기 때문이다.
- ▶ 컴퓨터가 의미를 구분할수 있는 최소 단위가 ON/OFF 이기 때문에 컴퓨터는 2진수를 사용한다.
- ▶ 그렇게 표현된 0 또는 1을 한개의 비트라고 한다.

10진수	16진수	2진수
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
10	A	1010
11	B	1011
12	C	1100
13	D	1101
14	E	1110
15	F	1111
16	10	10000



진법 변환



- 1) $(892)_{10} \rightarrow 2\text{진수}, 8\text{진수}, 16\text{진수로}$
- 2) $(385)_{10} \rightarrow 2\text{진수}, 8\text{진수}, 16\text{진수로}$
- 3) $(783.8125)_{10} \rightarrow 2\text{진수로}$
- 4) $(10101010)_2 \rightarrow 10\text{진수로}$
- 5) $(11110001)_2 \rightarrow 10\text{진수로}$
- 6) $(1010.101)_2 \rightarrow 8\text{진수}, 16\text{진수로}$

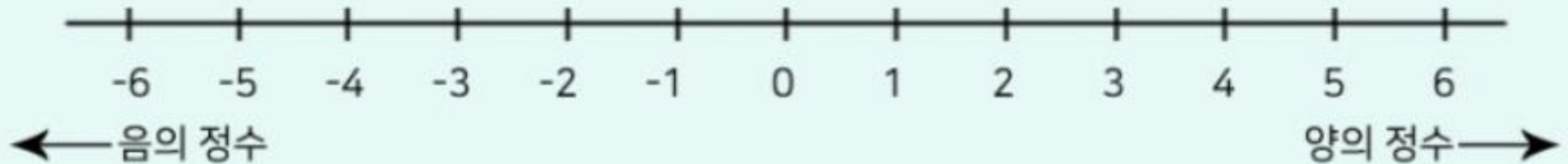


보수를 쓰는 이유



보수

- ① 양수에 대한 음의 값
- ② 컴퓨터가 사칙연산을 수행할 때 뺄셈을 덧셈방식으로 변환하여 산출
- ③ 예) ' $5-3=2$ ' 를 ' $5+(-3)=2$ '로 계산
→ 이때 ' -3 '이 ' 3 '의 보수, 즉 '음의 정수'가 됨



○○ 보수를 구하는 방법 3가지 ○○

음의 정수 표현방법

- ① 부호화 절대치(또는 부호화 절댓값) : 가장 왼쪽 비트가 부호를 의미하며, 그것을 '부호비트'라고 함
- ② 1의 보수 : 양수의 반전
- ③ 2의 보수 : 1의 보수에 1을 더한 값



10진수	부호화 절댓값	1의 보수	2의 보수
-4	-	-	100
-3	111	100	101
-2	110	101	110
-1	101	110	111
-0	100	111	-
+0	000	000	000
+1	001	001	001
+2	010	010	010
+3	011	011	011

○○ 보수를 구하는 방법 3가지 ○○

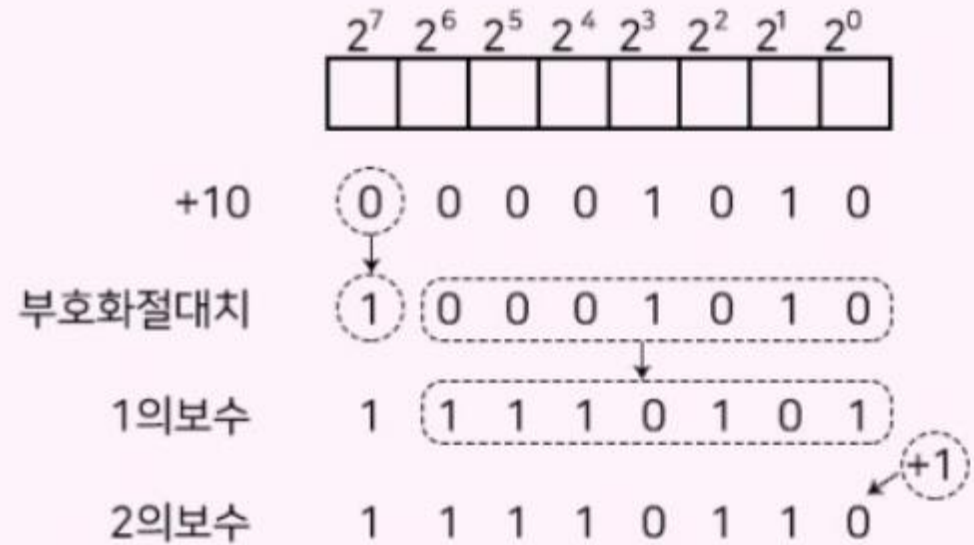
'음의 정수' 구하기

① 양수값을 8비트로 나타낸 후

② 부호화 절대치 방식으로 '-10' 표현

③ 1의 보수 : '양수' 전환하기

④ 2의 보수 : 1의 보수에 1더하기



-13을 부호화 절대치 방식으로 표현하면?

- ① 13을 2진수로 바꾸어 8비트로 표현 -----13을 2진수로 전환하면 00001101
- ② 양수를 음수로 바꾸기 위해 부호비트 전환-----가장 왼쪽의 부호비트 전환하면 10001101



표현 방식	+13	-13
부호화 절대값	0 0001101	1 0001101
1의 보수	0 0001101	1 1110010
2의 보수	0 0001101	1 1110011

-14를 부호화된 2의 보수로 표현하면?

- ① 14를 2진수로 바꾸어 8비트로 표현 -----14를 2진수로 전환하면 00001110
- ② 양수를 음수로 바꾸기 위해 부호비트 전환-----가장 왼쪽의 부호비트 전환하면 10001110
- ③ 양수와 음수를 전환하여 1의 보수 산출-----부호비트는 유지하고 나머지값 전환 11110001
- ④ 1의 보수에 +1을 하여 2의 보수 산출-----1의 보수에 +1을 하면 11110010

02 자료의 종류와 표현

1 수치 데이터의 표현

4. 실수형 데이터의 표현

● 부동 소수점 데이터 표현 방식

- 실수를 표현하기 위해 사용
- 4바이트 또는 8바이트로 표현
- 고정 소수점 데이터 표현 방식에 비해 기억 용량을 더 많이 차지하고 구조도 복잡하기 때문에 연산 속도가 느림
- 유효 숫자의 범위가 크고 작은 수의 표현이 가능
- 부호 비트에는 양수이면 0, 음수이면 1을 표시
- 지수부에는 지수 값에 127을 더한 값을 저장
- 가수부는 표현할 수를 2진수로 변환한 후 정규화 과정을 통하여 소수점 이하 부분만 저장

02 자료의 종류와 표현

2 문자 데이터의 표현

1. 유니코드

- 국제 표준 코드
- 하나의 문자를 표현하는 데 2바이트를 사용
- 사용 운영 체제와 프로그램에 관계 없이 문자마다 고유한 코드 값 제공
- 모든 문자를 16비트로 표현
- 최대 65,536자 표현 가능

02 자료의 종류와 표현

2 문자 데이터의 표현

2. ASCII 코드



[그림 1 - 07] ASCII 코드 형식



[그림 1 - 08] BCD 코드 형식



[그림 1 - 09] EBCDIC 코드 형식

02 자료의 종류와 표현

2 문자 데이터의 표현

3. 한글 코드

- 완성형 코드와 조합형 코드로 2종류
- **완성형 코드:** 한글 코드 표준, 하나의 음절에 코드를 부여하여 2바이트에 표현, 총 2,350자만 표현 가능하여 모든 글자를 나타낼 수 없음.
- **조합형 코드:** 한글의 초성, 중성, 종성 각각에 5비트를 조합하여 표현, 최상위 1비트는 영문자와 구별하기 위해 사용, 모든 한글을 표현할 수 있지만 국제 표준화 '기구의 정보 교환용 확장법에 위배되어 표준으로 쓰기에는 부적절

02 자료의 종류와 표현

3 이미지 데이터의 표현

1. 비트맵 이미지
2. 벡터 이미지



[그림 1 - 10] 비트맵 이미지와 벡터 이미지의 비교

[표 1 - 02] 비트맵 이미지와 벡터 이미지의 장단점

구분	비트맵	벡터
장점	<ul style="list-style-type: none">• 화면에 보여 주는 속도가 빠르다.• 미세한 색의 표현도 가능하다.	<ul style="list-style-type: none">• 평균적으로 기억 공간을 적게 차지한다.• 확대 시 화질 저하가 거의 없다.• 이동과 변형이 용이하다.
단점	<ul style="list-style-type: none">• 기억 공간을 많이 차지한다.• 이동과 변형이 어렵다.• 확대 시 화질이 저하된다.	<ul style="list-style-type: none">• 화면에 보여 주는 속도가 느리다.• 색의 변화를 나타내기 어렵다.

2. 이미지의 표현

1

자료와 정보

1. 이미지의 표현 방법은 대표적으로 **비트맵 방식**과 **벡터 방식**이 있다.
2. **비트맵 방식**은 **픽셀**을 이용하며, 일정 배율 이상 확대하면 이미지가 깨져 보이게 된다.
 - ✓ 1인치 크기에 사용되는 픽셀의 수 또는 이미지 전체에 포함된 픽셀의 수를 해상도라고 하며, 해상도에 따라 이미지의 선명도 및 파일 크기 등이 결정된다.
3. 많이 사용되는 비트맵 이미지 파일 형식: **BMP, JPEG, GIF, PNG** 등
4. **벡터 방식**은 x와 y좌표 값, 선 길이, 선 굵기, 색상 등에 대한 정보를 **수식의 형태로 표현**하기 때문에 확대해도 이미지가 깨지는 현상이 발생하지 않는다.
5. 이미지의 해상도 표현 단위
 - ✓ 이미지의 해상도는 일반적으로 1인치에 표현할 수 있는 픽셀의 수로 나타낸다. 보통 인쇄물의 경우 **DPI(Dots Per Inch)**, 모니터와 같은 디스플레이 장치의 경우 **PPI(Pixels Per Inch)**를 사용한다.
 - ✓ **DPI**는 출력물 1인치당 표현되는 픽셀의 수를, **PPI**는 보이는 화면 1인치당 표현되는 픽셀의 수를 의미하는 것이다.

02 자료의 종류와 표현

4 소리 데이터의 표현



[표 1 - 03] 오디오 파일의 종류

[그림 1 - 11] 오디오 파일 아이콘

파일 확장자	내용
MP3	현재 가장 널리 쓰이는 오디오 파일 포맷으로, 손실 압축 포맷을 이용한다. 사람이 들을 수 있는 소리만 압축하여 용량을 대폭 줄이며, 대부분의 웹 및 모바일 웹에서 지원한다.
WAV	윈도 PC 계열에 주로 쓰이는 표준 오디오 파일로, 비압축 방식의 CD급 품질의 오디오 파일을 저장하기 위하여 사용한다.
AU	선 마이크로 시스템즈, 유닉스, 자바 등에서 사용되는 표준 오디오 파일 포맷이다.
WMA	마이크로소프트사가 개발한 윈도 미디어 오디오 포맷이다.
RA	인터넷에서 스트리밍 서비스를 제공하여 실시간으로 음악을 감상할 수 있게 개발된 리얼 오디오 포맷이다.
AAC	MP3와 동일한 손실 압축 포맷을 사용하지만, MP3에 비하여 압축 파일의 크기를 반으로 줄일 수 있으며 우수한 복원 품질을 가진다. 애플사가 주로 사용하는 오디오 파일 포맷이다.

3. 소리의 표현

1

자료와 정보

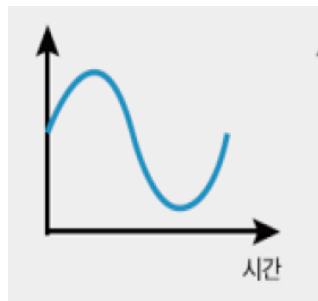
1. **WAV** 파일은 윈도 PC에서 주로 쓰이는 표준 오디오 파일 형식이다.
2. **MP3** 파일은 손실 압축 방법을 적용하여 파일의 크기를 줄일 수 있다.
3. **MIDI** 파일은 전자 악기끼리 디지털 신호를 주고받기 위한 파일로, 실제 오디오의 녹음 파일이 아니라 악기 명령의 모임이다.
4. 소리의 디지털 표현
 - ✓ 컴퓨터에서 소리를 표현하기 위해서는 전기 신호(아날로그 신호)를 이진 코드로 변환해야 하는데, 이를 위해 **PCM(Pulse-code modulation)**을 이용한다. PCM은 아날로그 신호를 일정한 간격으로 전압을 측정하여 디지털 값으로 변환하는 방법을 사용한다.
 - ✓ PCM은 다음 단계로 수행된다. (※ 뒷 장에 계속)

3. 소리의 표현

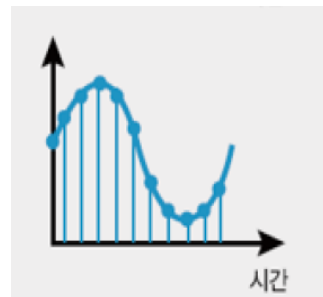
1

자료와 정보

- ✓ **표본화(Sampling)**: 연속된 파형에서 동일한 주기별로 신호 등급을 채취하는 과정이다.
- ✓ **양자화(Quantization)**: 채취한 PAM 신호(아날로그 데이터)를 이산적인 값(디지털 데이터)으로 바꾸어 표시하는 과정이다.
- ✓ **부호화(Encoding)**: 양자화 과정을 통해 얻은 값을 이진 코드(이진수)로 변환하는 과정이다. 부호화 과정을 통해 얻어진 이진 코드를 컴퓨터에서 처리할 수 있다.



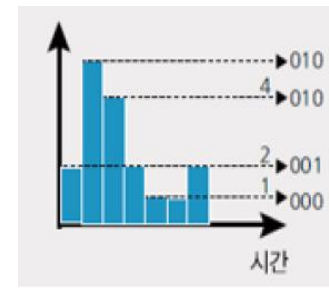
▲ 아날로그 신호



▲ 표본화



▲ 양자화



▲ 부호화

02 자료의 종류와 표현

5 동영상 데이터의 표현



[그림 1 - 12] 동영상 파일 아이콘

[표 1 - 04] 동영상 파일 형식의 종류

파일 확장자	내용
AVI	마이크로소프트사에서 개발한 동영상 파일 형식으로 전 세계적으로 널리 쓰인다. 윈도우에서 기본적으로 제공되는 파일 포맷이지만 압축률이 높지 않다는 단점이 있다. 압축된 파일이기 때문에 압축을 해제할 코덱이 필요하다.
MP4	MPEG(동영상 전문 그룹)에서 내놓은 동영상 파일 형식으로 MPEG-4의 일부로 규정된 파일 포맷이다. 인터넷을 통한 스트리밍 기술을 지원한다.
MOV	애플사의 표준 동영상 파일 형식으로 퀵 타임 프로그램을 통하여 지원한다.
ASF	인텔이 만든 멀티미디어 콘텐츠를 송수신하기 위하여 사용하는 데이터 포맷으로 스트리밍 기술을 지원한다.
RM	리얼 플레이어에서 사용되는 멀티미디어 스트리밍 데이터 파일이다. 인터넷에서 실시간으로 고품질의 동영상, TV, 라디오 등을 시청할 수 있다.

4. 동영상의 표현

1

자료와 정보

1. **동영상의 원리**는 여러 정지 화면을 빠르게 보여주어 우리가 화면을 볼 때는 연속적으로 움직이는 것처럼 느끼도록 하는 것이다.
2. 정지 화면 하나를 **프레임(Frame)**이라 한다. 단위 시간에 보이는 프레임의 수에 따라 자연스러운 움직임의 정도와 파일의 크기를 조절할 수 있다.
3. 많이 사용되는 동영상 파일 형식: **AVI, MPEG, WMV, MOV** 등
4. 동영상 콘텐츠를 제작할 때 일반적으로 **HDTV**의 경우 60fps, **UHD TV**의 경우 120fps, **비디오 게임**의 경우 30~60fps로 제작된다.
5. 동영상 파일을 컴퓨터에서 재생할 때 코덱이 설치되지 않아서 재생이 불가능한 경우가 있다.
 - ✓ **코덱(Codec)**이란 영상 또는 음성 등의 아날로그 신호를 디지털 방식으로 변환하는 코더(Coder)와 디지털 신호를 영상이나 음성으로 바꿔주는 디코더(Decoder)의 합성어이다.
 - ✓ 코덱을 통해 동영상 정보를 압축하여 저장하고, 압축을 해제하여 재생할 수 있다.



생각열기 활동

교내에서 학급 홍보 자료 경연 대회가 열린다. 우수 작품으로 선정되면 표창을 받고 간식 파티를 할 수 있다. 혜원이네 반 학생들은 학급 홍보 자료를 제작하기 위해 필요한 자료와 표현 방법에 대해 알아보기로 했다.

(1) 학급 홍보 동영상에 어떤 형태의 자료가 들어갈 수 있을까?

(2) 여러 유형의 자료를 어떻게 만들고 편집할 수 있을까?

① 자료 유형별로 파일 형식

② 자료 유형별로 편집 도구

뛰어넘기 활동

1. 수영이는 학교 축제 기간 중 촬영한 동영상에 배경 음악을 삽입한 후 학교 홈페이지에서 바로 재생할 수 있도록 편집 중 이다. 여러 가지 동영상 파일 형식 중 어떤 방식으로 제작하는 것이 적절할까? 왜 그렇게 생각하는지 함께 설명해 보자.

뛰어넘기 활동

2. 윤석이는 친구와 함께 하늘을 향해 점프하는 과정을 여러 컷의 사진으로 찍었다. 이 사진을 편집하여 움직이는 영상처럼 편집하려고 한다. 여러 가지 이미지 파일 형식 중 어떤 방식으로 제작하는 것이 적절할까? 왜 그렇게 생각하는지 함께 설명해 보자.



3. 윤영이는 부모님의 생신을 맞이하여 부모님이 좋아하시는 팝송 3,000곡을 1GB 용량의 USB 메모리에 담아 드리려고 한다. 최대한 많은 음악을 넣기 위하여 파일을 변환하려고 할 때, 어떤 파일 형식으로 변환하는 것이 적절할까? 왜 그렇게 생각하는지 함께 설명해 보자.

