비정형 데이터
Feat. 이미지

비정형 데이터의 문제 해결 사례

입력 문장: 아보카도 모양의 의자

출처 https://openai.com/blog/dall-e/

비정형데이터

- 이메일, 문자메시지, 사진, 음성, 영상 등 정해진 규칙이 없고 구조화되지 않은 데이터
- 정보 통신의 발달로 수많은 비정형 데이터가 만들어지고 있음.

이미지 처리 응용 사례

의학 영상 처리: 악성 흑색종과 기저 세포암 진단

이미지 처리 응용 사례

보안 시스템 개발: 홍채인식, 안면인식, 지문인식

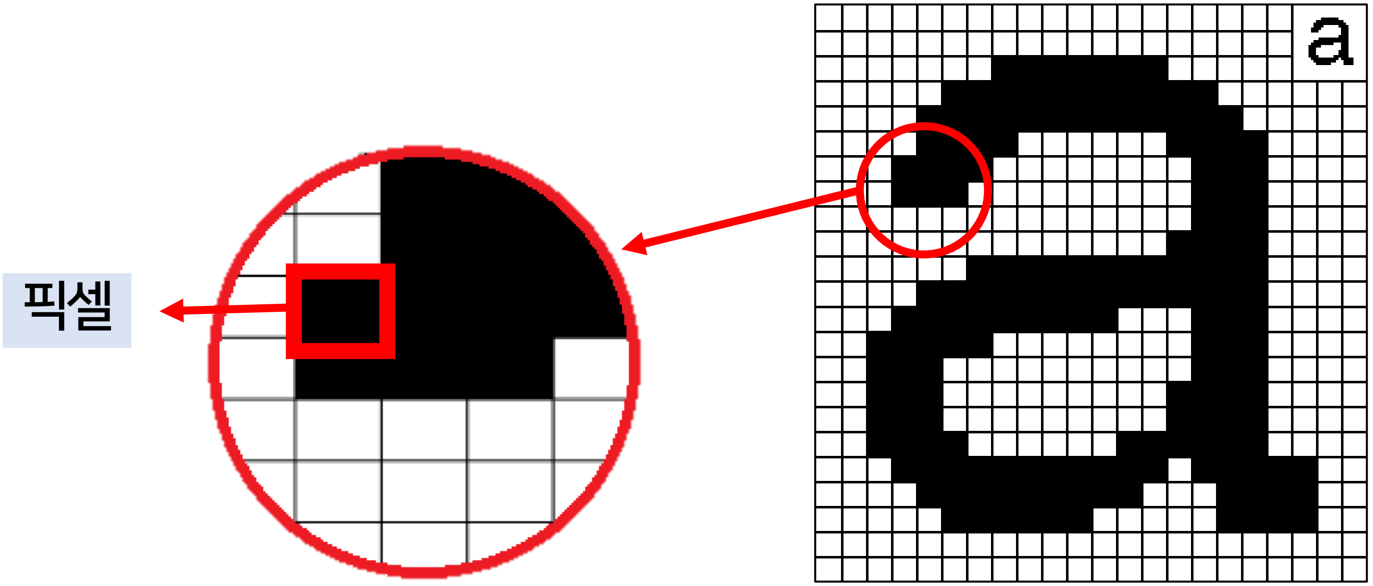
이미지 처리 응용 사례

이미지 복원: 흑백 사진을 컬러사진으로 복원

▲ 딥러닝으로 흑백 사진을 컬러 사진으로 재복원

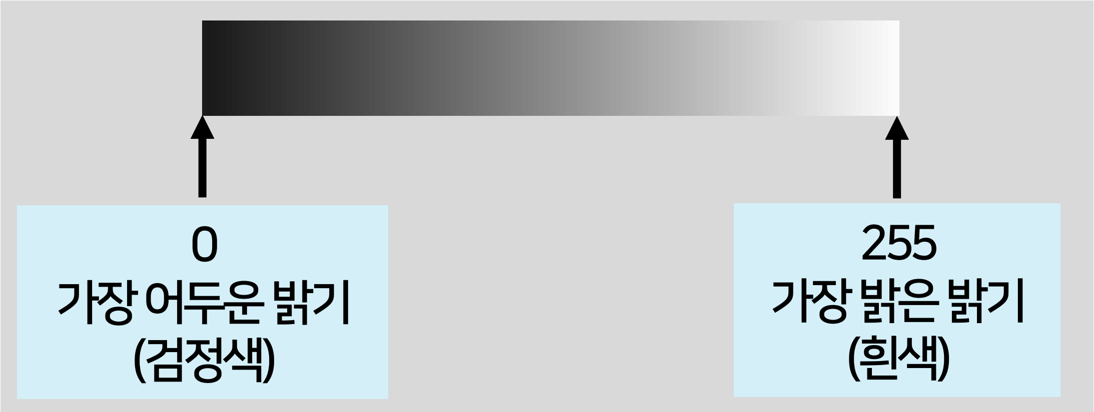
컴퓨터가 이미지를 표시하고 저장하는 방법

- 픽셀(pixel) - 디지털 이미지를 구성하는 기본 단위, 화소
- 비트맵 이미지 - 픽셀들이 모여 하나의 그림을 표현함.
- 파일 형식 - JPG, PNG, BMP...



흑백 사진의 저장 방법

- 색상 정보가 없이 오직 밝기 정보만으로 구성.
- 밝기 정보를 0~255의 정수 범위로 256단계 표현

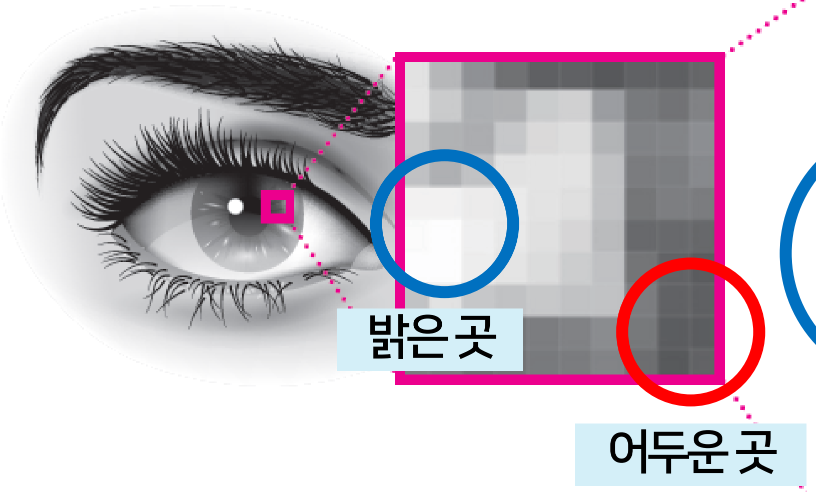


256단계 = 2^8
8개의 bit 필요

Python에서는
uint8 dataType
사용

흑백 사진의 저장 방법

- 색상 정보가 없이 오직 밝기 정보만으로 구성.
- 밝기 정보를 0~255의 정수 범위로 256단계 표현



230	194	147	108	90	98	84	96	91	101
237	206	188	195	207	213	163	123	116	128
210	183	180	205	224	234	188	122	134	147
195	189	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
249	241	237	244	232	226	202	116	125	126
251	254	241	[illegible]	230	217	196	102	103	99
243	255	240	231	227	214	203	[illegible]	95	91
204	231	208	200	207	201	[illegible]	121	95	95
144	140	120	115	125	127	[illegible]	118	92	91
121	121	108	109	122	121	134	106	86	97

그레이스케일 값이 크다

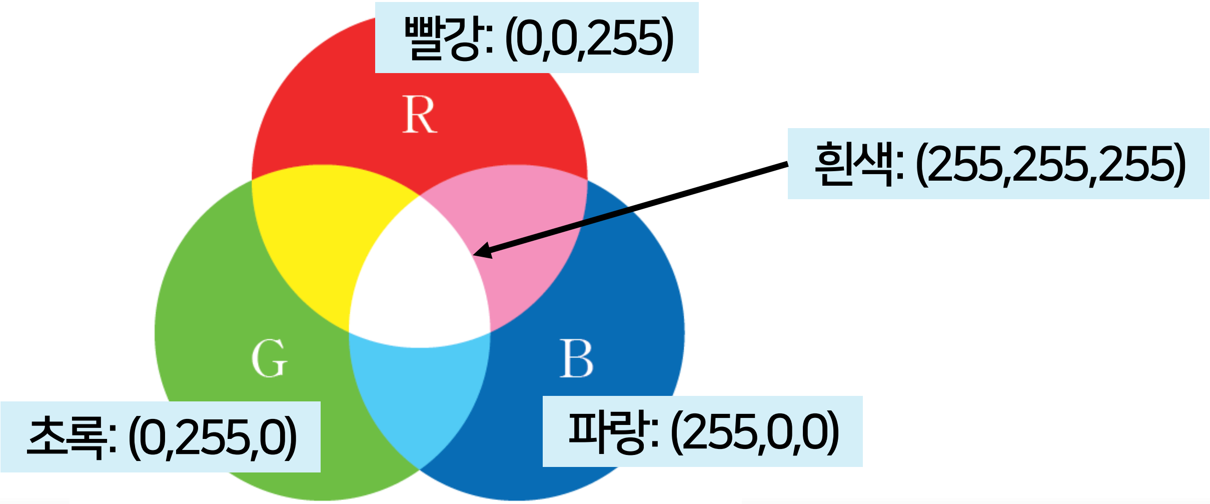
그레이스케일 값이 작다

컬러 사진의 저장 방법

- 빨강(R), 초록(G), 파랑(B)이 혼합되어 색을 표현한다.
- OpenCV는 색상을 B, G, R 순서로 표현한다.
- 3가지의 각 색상은 0~255 사이의 정수 256단계로 표현된다.

→ 256^3 = 16,777,216 색상 표현 가능

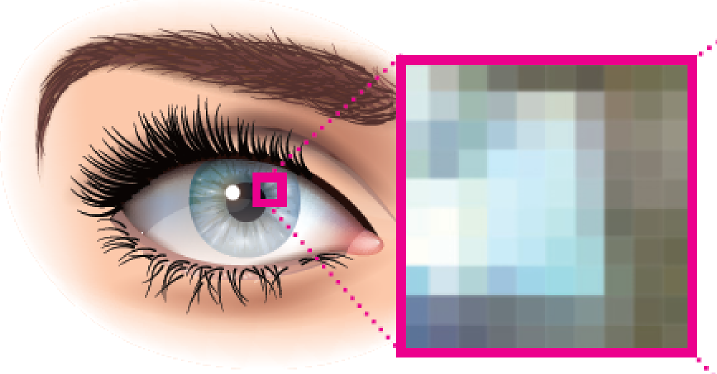
■ 빛의 3원색



컬러 사진의 저장 방법

- 하나의 픽셀에 3개의 색상이 표현된다.
- 파이썬의 경우 (세로X가로X색) 3차원 행렬로 컬러사진을 표현한다.

세로 픽셀수*가로 픽셀수*색상(RGB)



		233	[illegible]	137	96	90	95	63	73	73	82
	237	202	159	120	105	110	88	107	112	121	109
226	191	147	110	101	112	98	123	110	119	142	131
221	191	176	182	203	214	169	144	133	145	155	122
185	160	161	184	205	223	186	137	147	161	140	115
181	174	189	207	206	215	194	136	142	151	133	87
245	237	237	231	208	206	192	122	142	144	111	74
254	254	241	224	199	192	181	99	122	117	117	74
239	248	232	207	187	182	184	110	114	110	113	74
293	215	293	167	158	164	181	114	112	111	105	82
113	119	110	111	113	123	135	120	188	186	113	
93	97	91	103	107	111	122	112	104	114		

(10, 10, 3)

행렬이란?

1 행렬의 뜻

(1) 여러 개의 수 또는 문자를 직사각형 형태로 배열하여 괄호로 묶어 나타낸 것을 **행렬**이라고 한다. 이때, 행렬을 이루는 각각의 수나 문자를 그 행렬의 **성분**이라고 한다.

(2) 행렬의 성분을 가로로 배열한 줄을 **행**이라 하고, 세로로 배열한 줄을 **열**이라고 한다. 일반적으로 m개의 행과 n개의 열로 이루어진 행렬을 **$m\times n$ 행렬** 또는 **m행 n열의 행렬**이라고 한다. 특히, $n\times n$ 행렬을 **n차 정사각행렬**이라고 한다.

3 행렬의 실수배

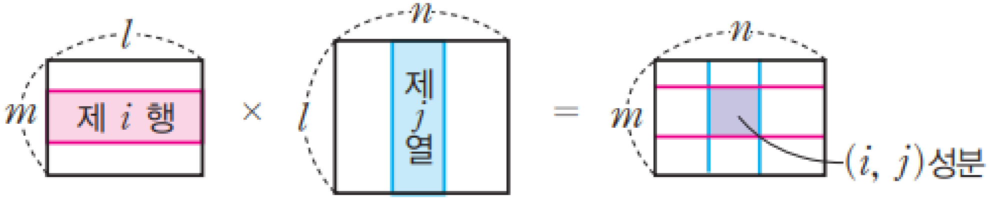
일반적으로 임의의 실수 k에 대하여 행렬 A의 각 성분을 k배 한 것을 성분으로 하는 행렬을 행렬 A의 k배라고 하며, 이것을 기호로 kA와 같이 나타낸다.

이를 테면 2×2 행렬의 실수배는 다음과 같다.

행렬 $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$와 실수 k에 대하여 $kA=\begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$

4 행렬의 곱셈

일반적으로 두 행렬 A, B에 대하여 행렬 A의 열의 개수와 행렬 B의 행의 개수가 같을 때, 행렬 A의 제i행의 성분과 행렬 B의 제j열의 성분을 각각 차례로 곱하여 더한 값을 $(i,\ j)$성분으로 하는 행렬을 두 행렬 A, B의 곱이라고 하며, 이것을 기호로 AB와 같이 나타낸다.



▲ 행렬 A　　▲ 행렬 B　　▲ 행렬 AB

다음을 계산하여라.

(1) $\begin{pmatrix} 3 & 4 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix}$

(2) $\begin{pmatrix} 5 & 2 \\ -7 & 4 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$

$A+B=\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $B+C=\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$을 만족시키는 세 행렬 A, B, C에 대하여 다음 중 행렬 $C-A$와 같은 것은?

① $\begin{pmatrix} 3 & 4 \\ -1 & 2 \end{pmatrix}$ ② $\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}$ ③ $\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$ ④ $\begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}$ ⑤ $\begin{pmatrix} -3 & -4 \\ 1 & -2 \end{pmatrix}$

2 일차변환

좌표평면 위의 변환 $f:(x, y) \to (x', y')$에서

$$\begin{cases} x'=ax+by \\ y'=cx+dy \end{cases} (a, b, c, d\text{는 상수}) \quad \cdots\cdots ㉠$$

와 같이 x', y'이 상수항이 없는 x, y의 일차식으로 나타내어질 때, 이 변환 f를 **일차변환**이라고 하며, ㉠을 일차변환 f를 나타내는 식이라고 한다.

㉠을 행렬로 나타내면

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

이므로 $X'=\begin{pmatrix} x' \\ y' \end{pmatrix}$, $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $X=\begin{pmatrix} x \\ y \end{pmatrix}$로 놓으면 다음과 같다.

$$X'=AX \quad \cdots\cdots ㉡$$

따라서 일차변환 f를 나타내는 식 ㉠이 주어지면 ㉡와 같이 행렬 A가 결정되고, 역으로 행렬 A가 주어지면 ㉡에 의하여 ㉠과 같이 일차변환 f가 정해짐을 알 수 있다. 이때, 행렬 A를 일차변환 f를 나타내는 행렬 또는 일차변환 f의 행렬이라고 한다.

확인 문제 2

행렬 $A=\begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$로 나타내어지는 일차변환 f에 의하여 점 $\mathrm{P}(1, 1)$, $\mathrm{Q}(2, -3)$이 각각 옮겨지는 점 P', Q'의 좌표를 구하시오.

수학으로 풀어보기

점 $\mathrm{P}(1, 1)$, $\mathrm{Q}(2, -3)$이 일차변환 f에 의하여 각각 $\mathrm{P}'(x_1, y_1)$, $\mathrm{Q}'(x_2, y_2)$로 옮겨진다고 하면

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}=\begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}=\begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}=\begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix}=\begin{pmatrix} 7 \\ 0 \end{pmatrix}$$

따라서 $\mathrm{P}'(1, 5)$, $\mathrm{Q}'(7, 0)$이다.

답 $\mathrm{P}'(1, 5)$, $\mathrm{Q}'(7, 0)$

대칭변환

❶ x축에 대한 대칭변환 $\begin{cases} x'=x \\ y'=-y \end{cases}$ 의 행렬은 $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

❷ y축에 대한 대칭변환 $\begin{cases} x'=-x \\ y'=y \end{cases}$ 의 행렬은 $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

❸ 원점에 대한 대칭변환 $\begin{cases} x'=-x \\ y'=-y \end{cases}$ 의 행렬은 $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

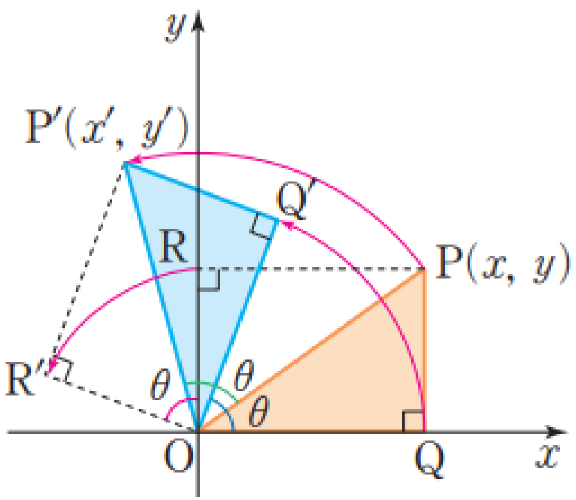
❹ 직선 $y=x$에 대한 대칭변환 $\begin{cases} x'=y \\ y'=x \end{cases}$ 의 행렬은 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

회전변환

원점을 중심으로 점 $\mathrm{P}(x,\ y)$를 각 θ만큼 회전하여 점 $\mathrm{P}'(x',\ y')$으로 옮기는 회전변환의 행렬은

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

좌표평면 위의 점 $\mathrm{P}(x, y)$를 원점을 중심으로 각 θ만큼 회전하여 점 $\mathrm{P}'(x', y')$으로 옮기는 변환은 다음과 같다.



오른쪽 그림과 같이 두 점 $\mathrm{Q}(x, 0)$, $\mathrm{R}(0, y)$를 원점 O를 중심으로 각 θ만큼 회전한 점을 각각 Q', R'이라고 하면

$$\mathrm{Q}'(x\cos\theta, x\sin\theta)$$

$$\mathrm{R}'(y\cos(90^\circ+\theta), y\sin(90^\circ+\theta))$$

즉, $\mathrm{R}'(-y\sin\theta, y\cos\theta)$이다.

이때, 직사각형 $\mathrm{OQ'P'R'}$에서 두 대각선 $\mathrm{OP'}$, $\mathrm{Q'R'}$의 중점이 일치하므로

$$\frac{x'}{2}=\frac{x\cos\theta-y\sin\theta}{2},\ \frac{y'}{2}=\frac{x\sin\theta+y\cos\theta}{2}$$

이다. 따라서 다음이 성립한다.

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases}$$

이것을 행렬을 이용하여 나타내면 다음과 같다.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

이와 같이 좌표평면 위의 점을 원점을 중심으로 각 θ만큼 회전하여 옮기는 일차변환을 **회전변환**이라고 한다.