

08강

인공지능

컴퓨터 시각과 패턴인식(1)

컴퓨터과학과 이병래교수

학습목차

1 컴퓨터 시각의 개념

2 디지털 영상

3 전처리

4 영상분할



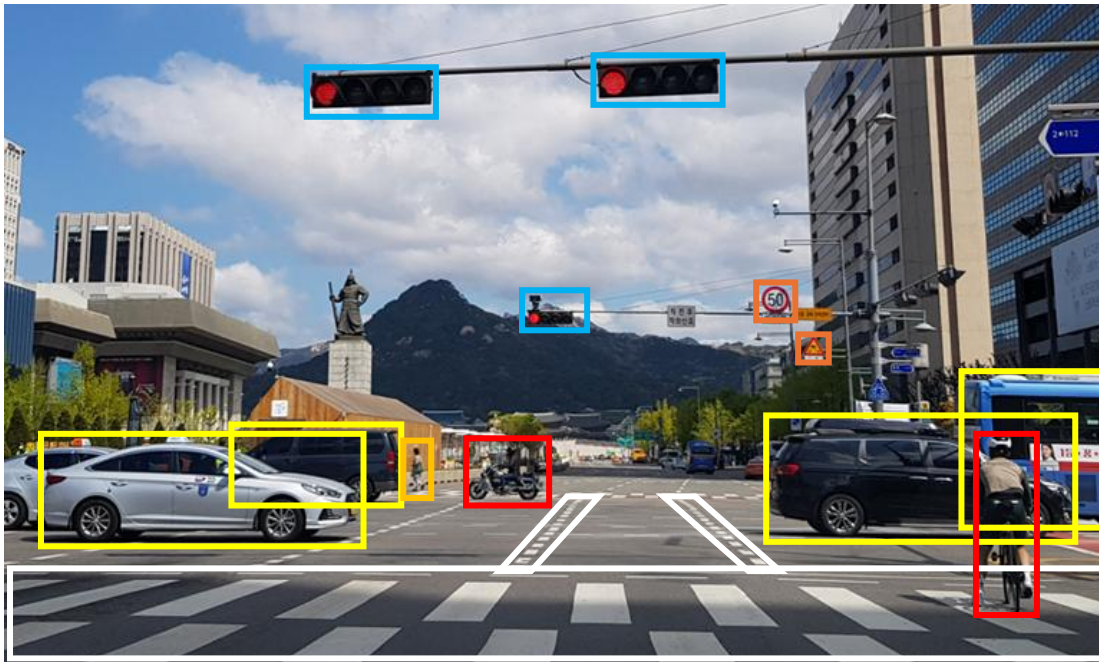


컴퓨터 시각의 개념

1. 컴퓨터 시각

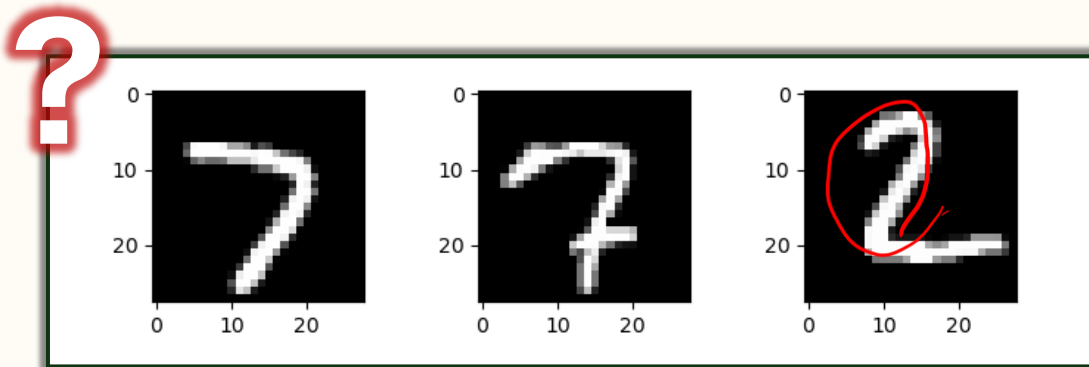
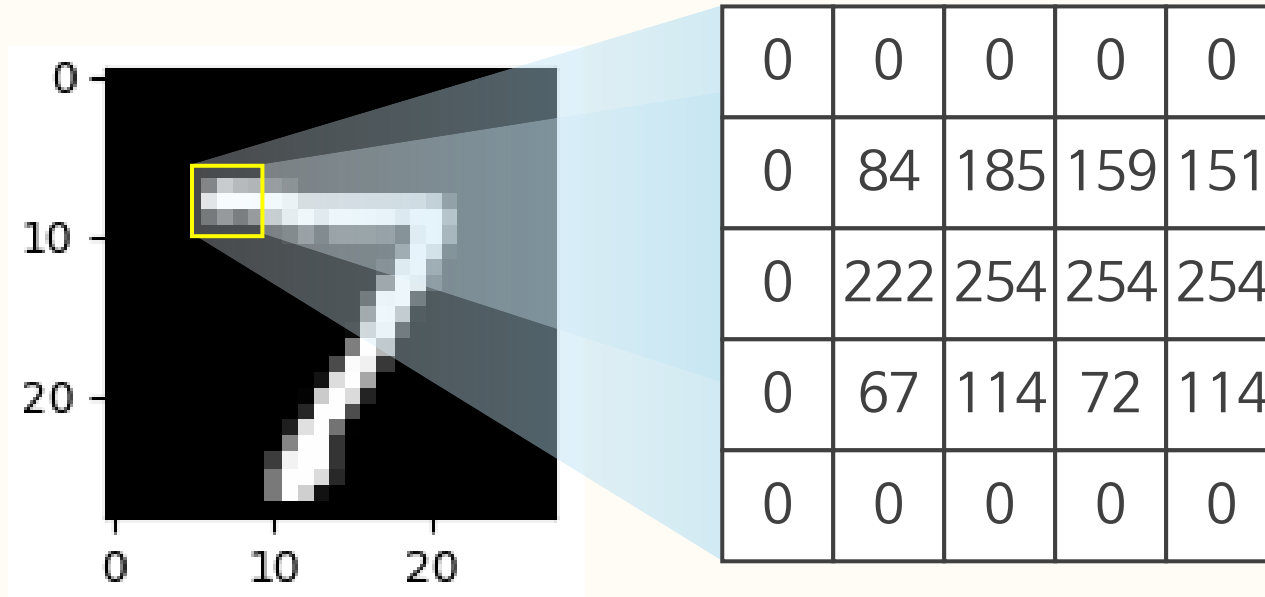
■ 컴퓨터 시각(computer vision)이란?

- 인간의 시각체계에 해당되는 능력을 컴퓨터에 심어 넣기 위해 연구하는 인공지능의 한 분야



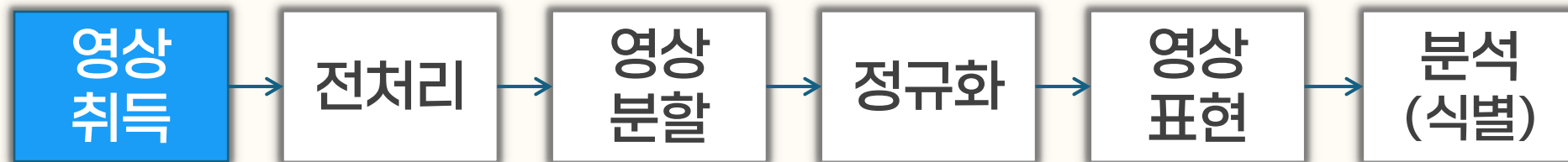
1. 컴퓨터 시각

■ 시각 데이터와 시각 정보의 인지



2. 컴퓨터 시각 시스템

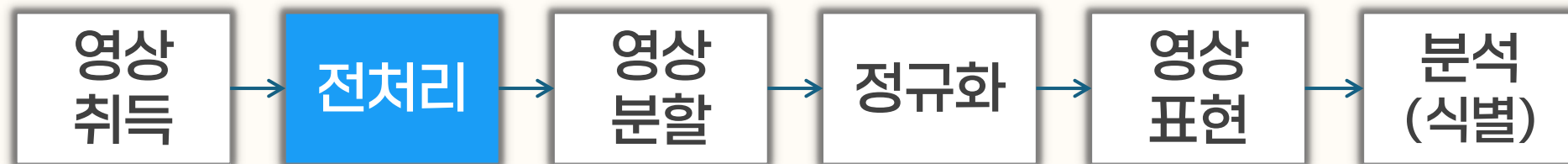
■ 컴퓨터 시각 시스템의 처리 단계



- 다양한 유형의 센서를 이용하여 디지털 영상 취득
 - 영상 센서, 적외선 센서를 이용한 시각 영상
 - 레인지 센서를 이용한 거리 영상
 - X-선 단층촬영 영상

2. 컴퓨터 시각 시스템

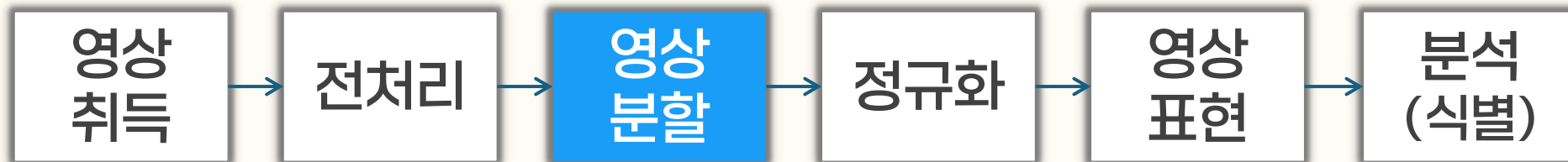
■ 컴퓨터 시각 시스템의 처리 단계



- 취득한 영상을 보다 처리하기 좋은 형태로 가공하는 처리
 - 영상 취득 과정에서 발생된 훼손이나 잡음의 제거
 - 영상의 개선: 명암, 대비 등을 개선하여 필요한 정보를 검출하기 좋게 함
 - 영상의 변환

2. 컴퓨터 시각 시스템

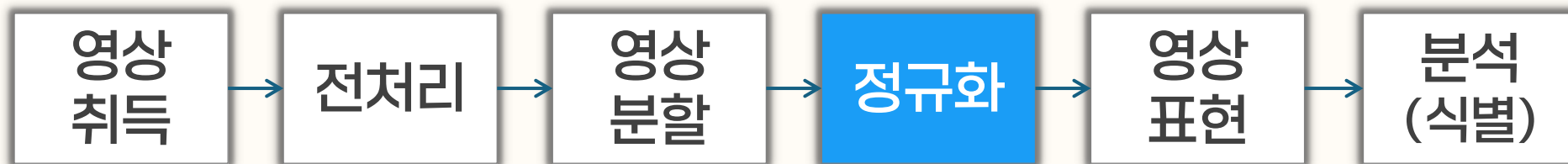
■ 컴퓨터 시각 시스템의 처리 단계



- 영상 해석의 단위가 될 수 있는 영역을 다른 부분과 구분하는 처리
- 유사한 성격을 갖는 영상의 부분은 하나의 영역으로 묶고, 상이한 특성을 갖는 부분은 서로 다른 영역으로 구분

2. 컴퓨터 시각 시스템

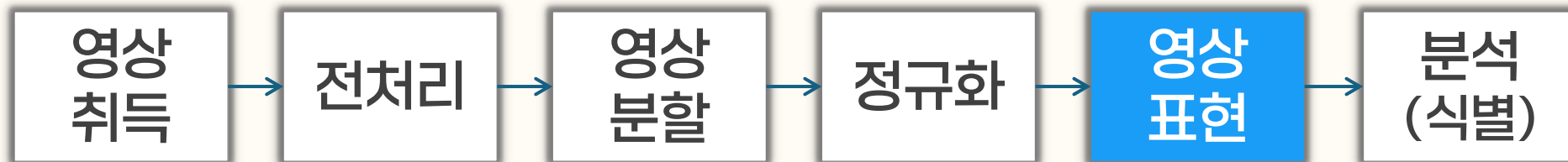
■ 컴퓨터 시각 시스템의 처리 단계



- 영상 내에서 발생할 수 있는 변형을 회복하여 기준이 되는 형태로 변환하는 것
 - 객체의 크기, 위치
 - 객체가 놓인 방향
 - 영상의 밝기, 대비 등

2. 컴퓨터 시각 시스템

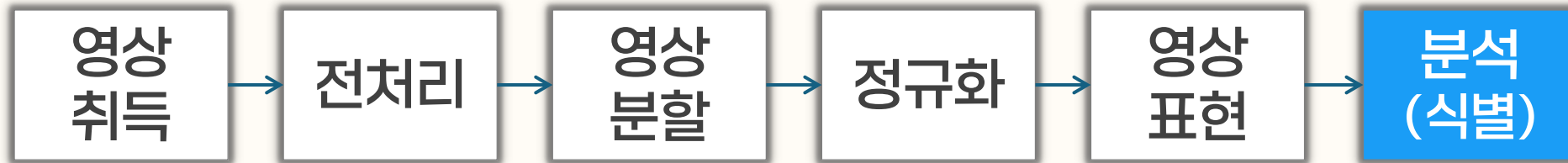
■ 컴퓨터 시각 시스템의 처리 단계



- 분할된 영역들로 구성되는 객체를 해석하기 위해 각각의 영역을 적절한 방법으로 컴퓨터 내에 묘사하는 것 - **특징**
 - 대상의 유용한 속성을 잘 표현할 수 있도록 함
 - 객체의 변형에도 고유한 속성을 유지할 수 있는 표현방법이 바람직함
 - 선(lines), 에지(edge), 모서리(corner), 덩어리(blob) 등

2. 컴퓨터 시각 시스템

■ 컴퓨터 시각 시스템의 처리 단계



- 분할 영역의 영상, 검출된 특징 등을 바탕으로 필요한 정보를 구함
 - 대상물체의 인식, 크기 측정, 위치 계산, 결함 검사 등

3. 컴퓨터 시각의 활용

■ 시각 인식

- 글자, 물체 등의 객체 인식
- 얼굴, 지문, 홍채 등의 생체인식

■ 의료 진단을 위한 영상 분석

■ 제조 공정에서 제품 결함의 자동 시각 검사

■ 보안 시스템 - 객체의 검출 및 추적

■ 로봇, 자율주행

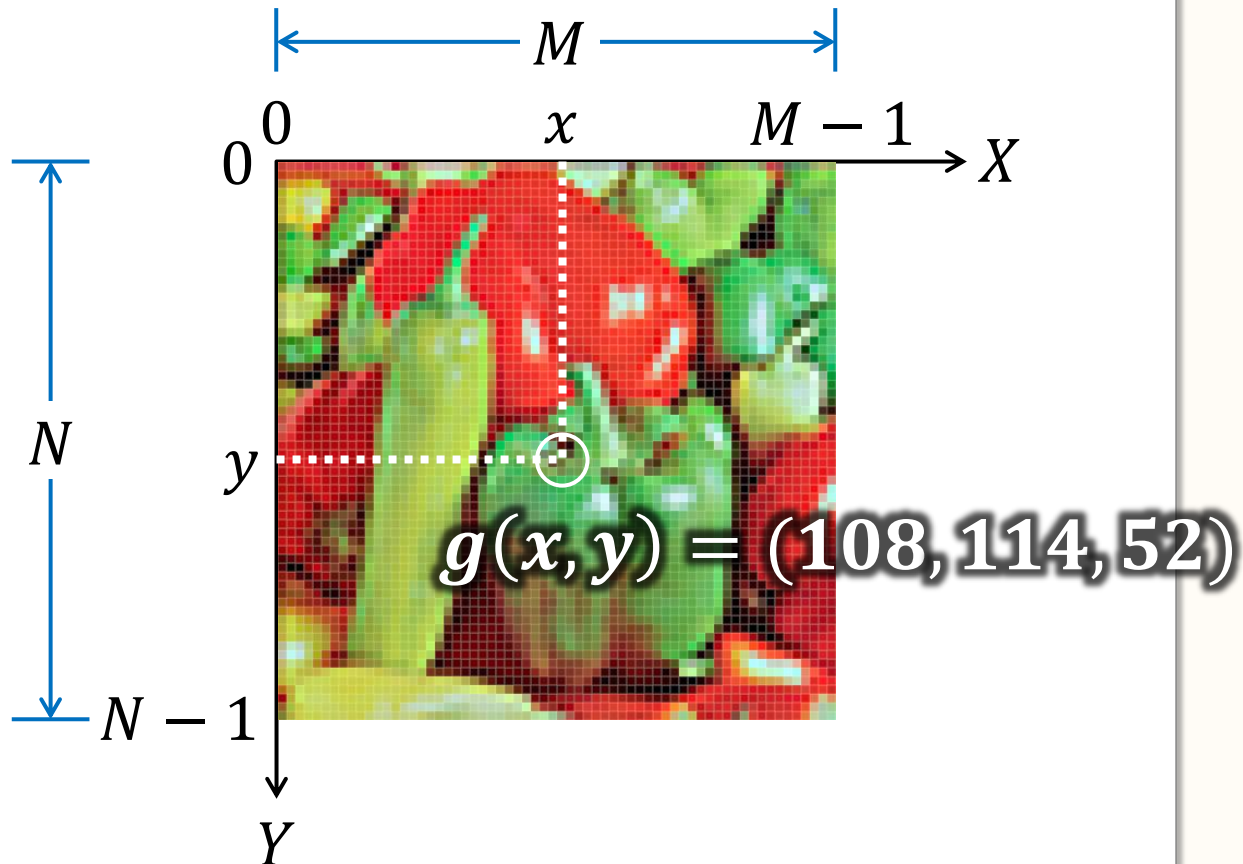
■ 사람과 컴퓨터의 상호작용을 위한 입력 장치



디지털 영상

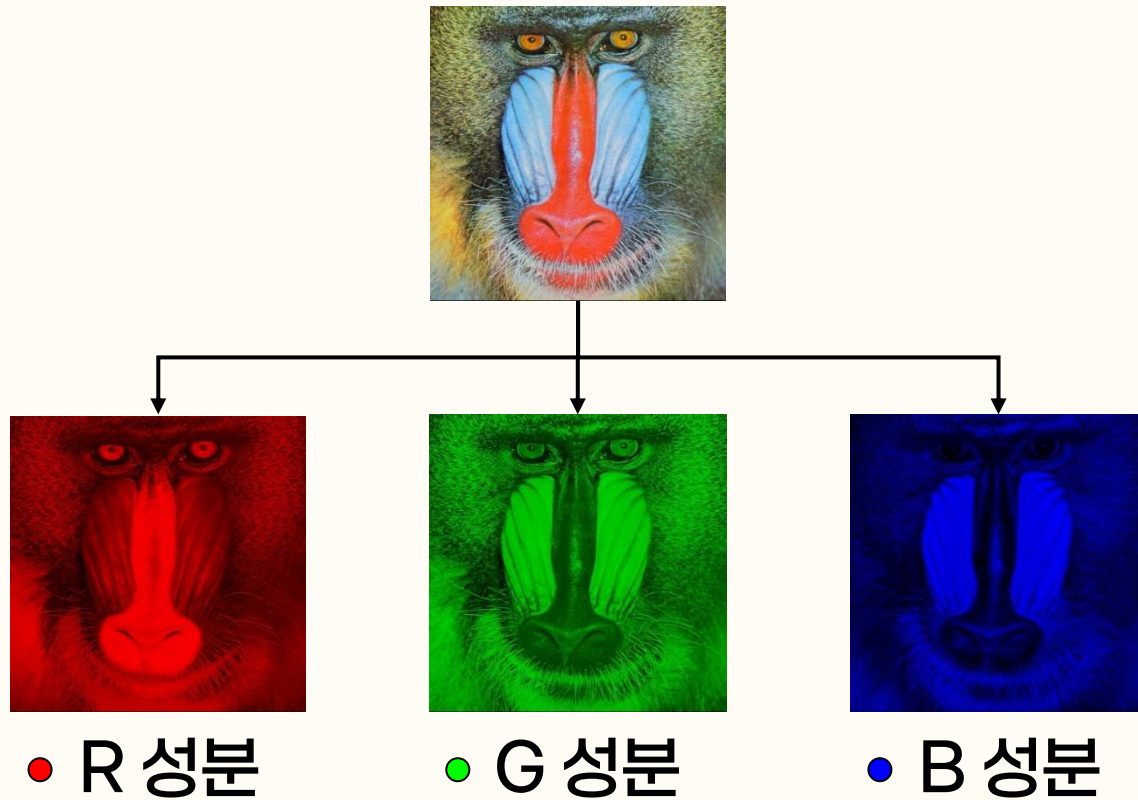
1. 디지털 영상

❑ 디지털 영상과 픽셀(pixel)



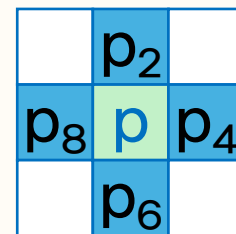
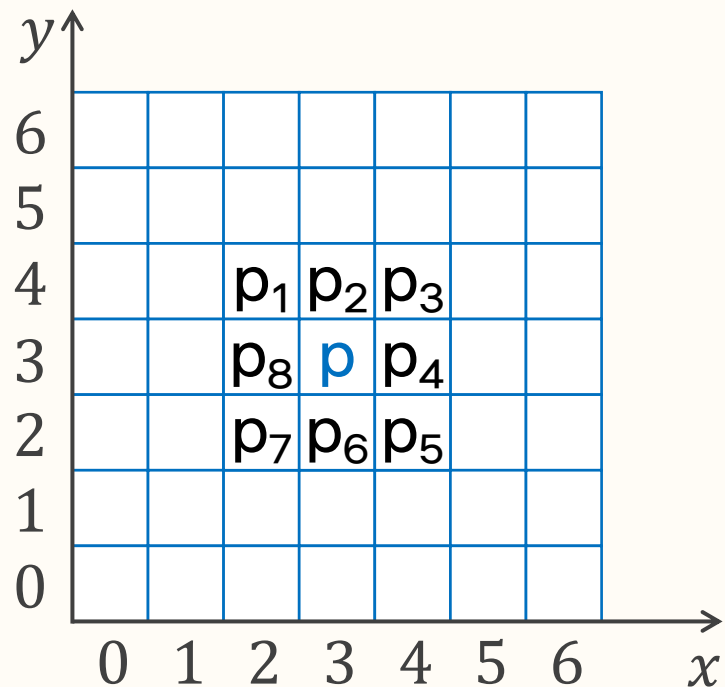
1. 디지털 영상

■ 색 정보의 표현 - RGB 색 좌표계

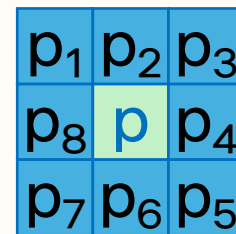


1. 디지털 영상

■ 사각형 픽셀 구조 및 연결성



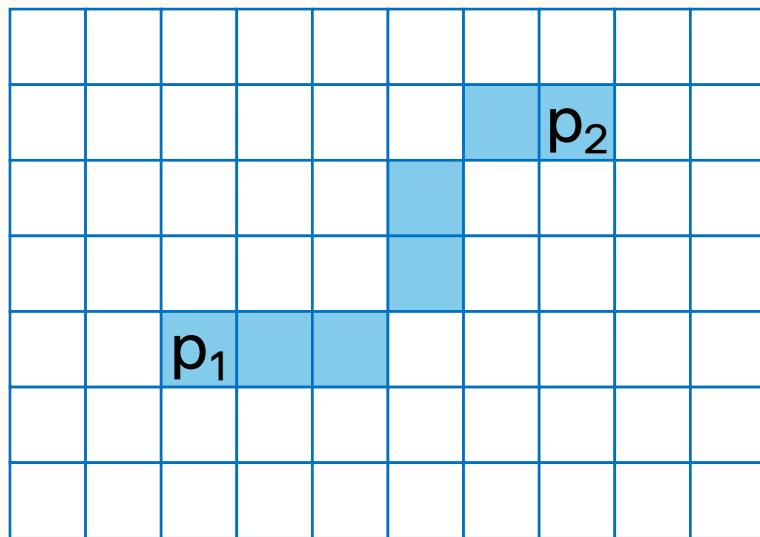
4-이웃 연결성



8-이웃 연결성

1. 디지털 영상

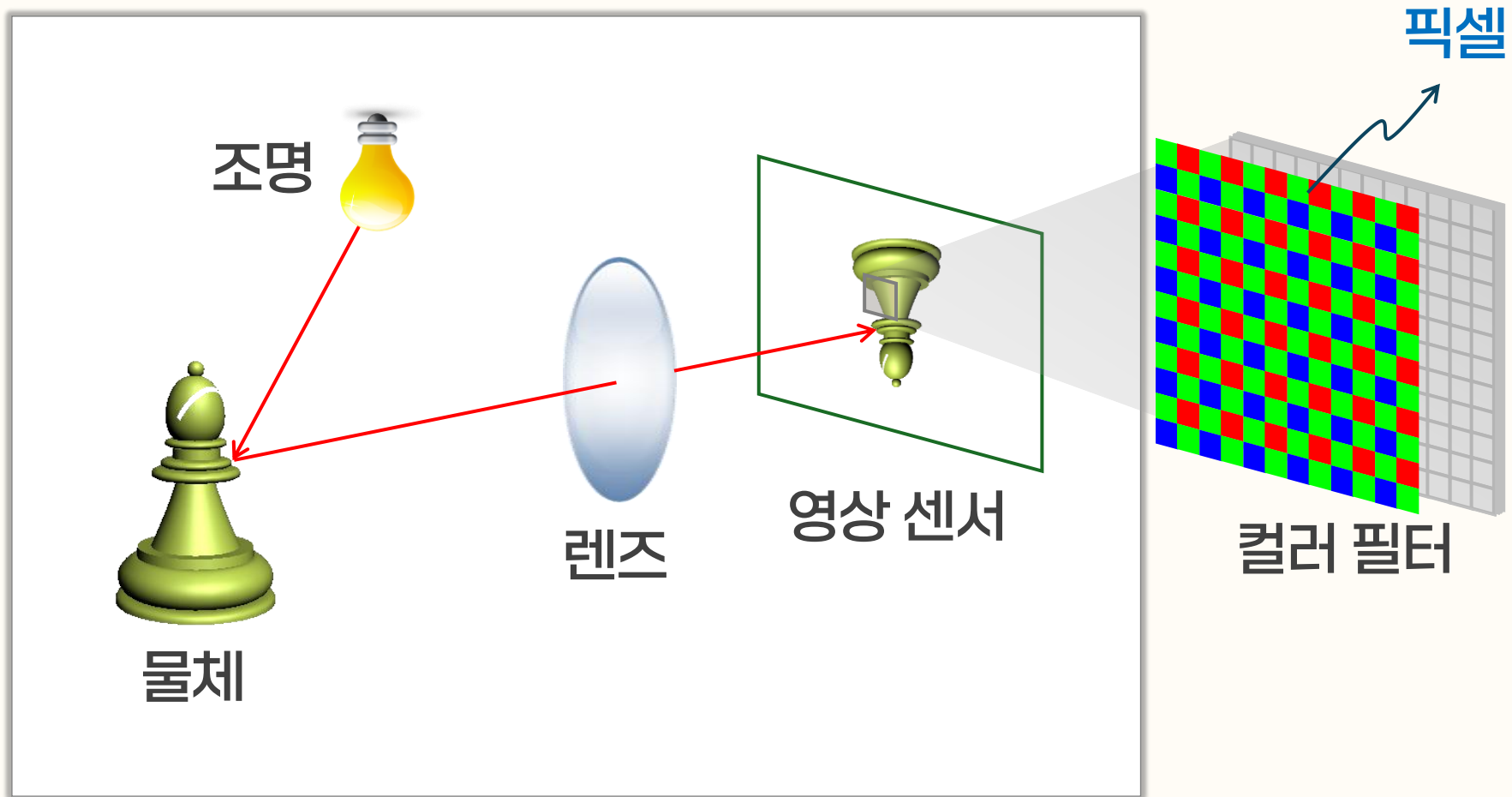
■ 사각형 픽셀 구조 및 연결성



- 4-이웃 연결성 적용 ➡ p_1 과 p_2 를 연결하는 경로 없음
- 8-이웃 연결성 적용 ➡ p_1 으로부터 p_2 까지 연결됨

2. 디지털 영상의 입력

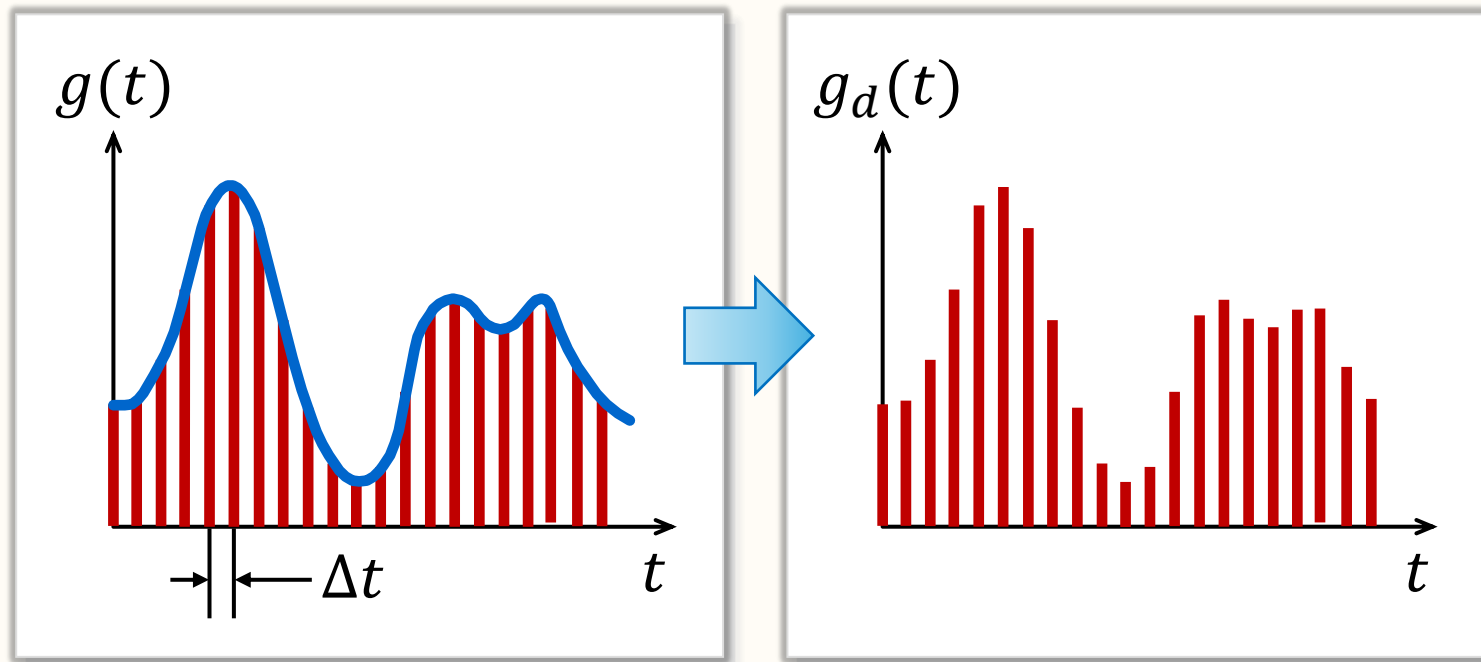
❏ 디지털 영상의 형성 과정



2. 디지털 영상의 입력

■ 표본화(sampling)

- 연속적으로 변화하는 아날로그 신호를 이산(discrete) 신호로 변환하는 것



2. 디지털 영상의 입력

■ 표본화(sampling)

Nyquist-Shannon 표본화 이론

어떠한 신호 $g(t)$ 가 주파수 B 에서 대역제한되어 있고, 이를 표본화 주파수 $f_s = 1/\Delta t$ 로 표본화하여 신호 $g_d(t)$ 를 얻었다. 이 경우 만약 $f_s \geq 2B$ 라면 표본화된 신호 $g_d(t)$ 로부터 원래의 신호 $g(t)$ 를 복원할 수 있다.

 $f_s < 2B$ 이면 에일리어싱(aliasing) 발생

2. 디지털 영상의 입력

▣ 표본화(sampling)

256×256



64×64



32×32



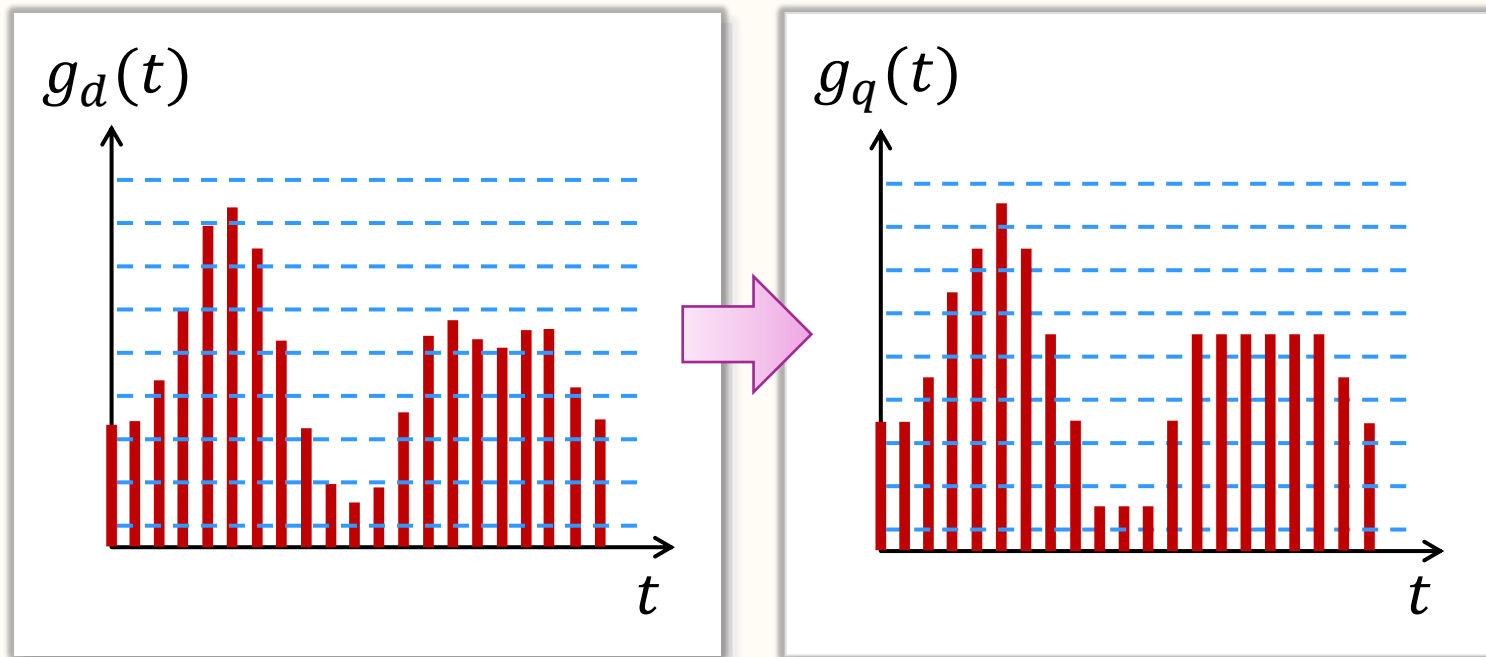
16×16



2. 디지털 영상의 입력

■ 양자화(quantization)

- 이산신호의 값에 정해진 정밀도로 분할된 구간의 대푯값을 부여하는 과정



2. 디지털 영상의 입력

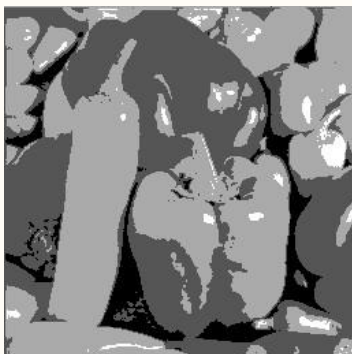
양자화(quantization)



256단계



16단계



4단계



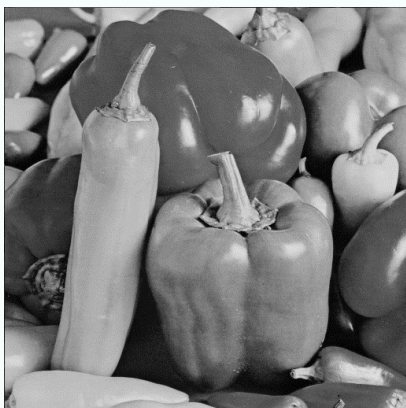
2단계

2. 디지털 영상의 입력

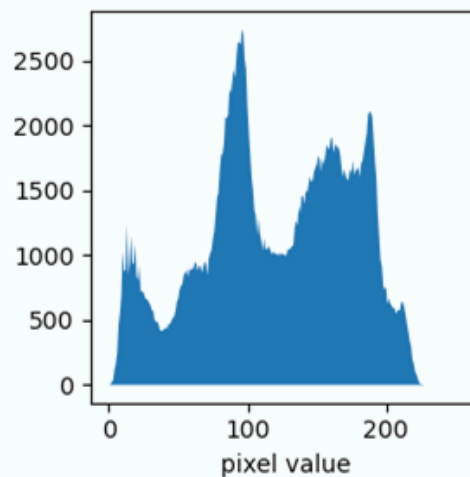
■ $M \times N$ 크기의 256단계 흑백 영상 I

$$I = \{g(x, y) \mid g(x, y) \in \{0, 1, \dots, 255\}, \\ x \in \{0, 1, \dots, M - 1\}, y \in \{0, 1, \dots, N - 1\}\}$$

흑백 영상



그레이 레벨 히스토그램





전처리

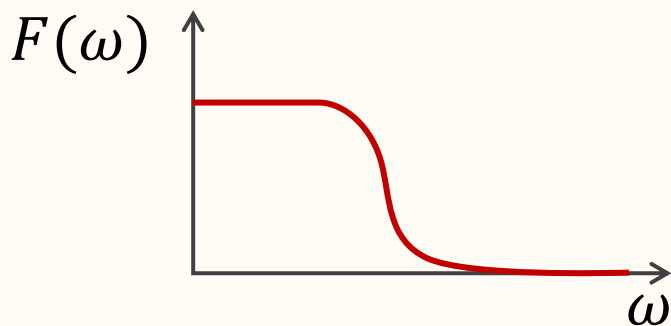
1. 영상 필터링

필터(filter)

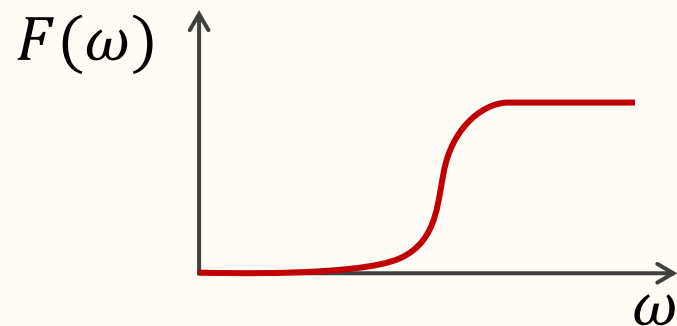
- 입력신호에서 불필요한 요소를 제거하여 시스템에 필요한 요소만으로 구성된 신호를 출력하는 장치



저역 통과 필터(LPF)



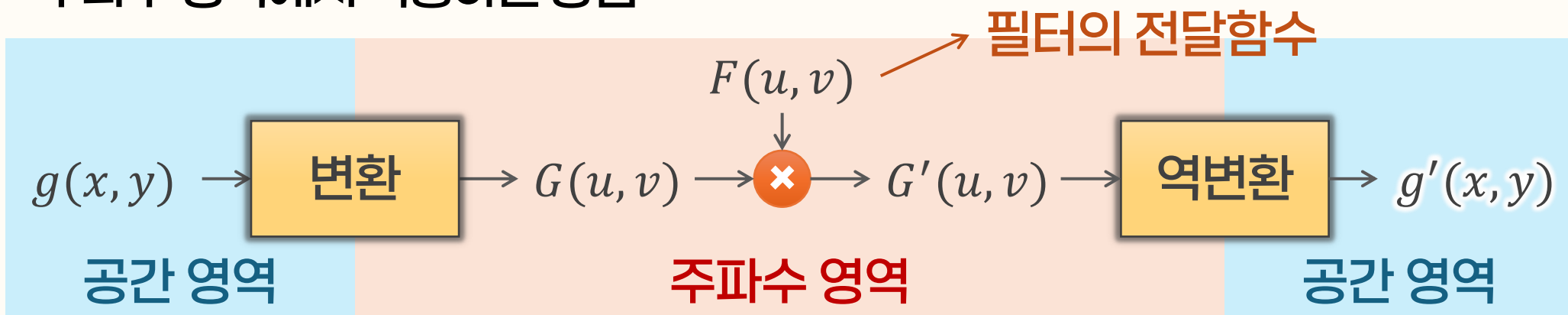
고역 통과 필터(HPF)



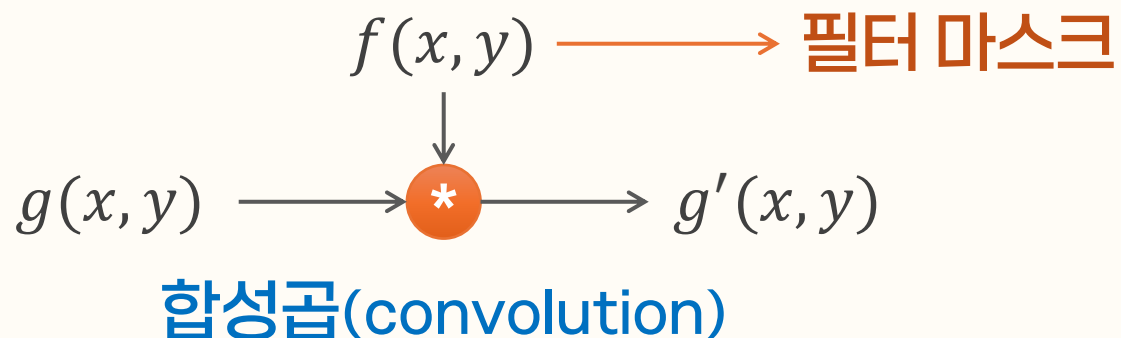
1. 영상 필터링

필터의 적용 방법

- 주파수 영역에서 적용하는 방법



- 공간 영역에서 적용하는 방법



2. 합성곱을 이용한 영상 필터링

합성곱과 필터 마스크

그레이스케일 영상



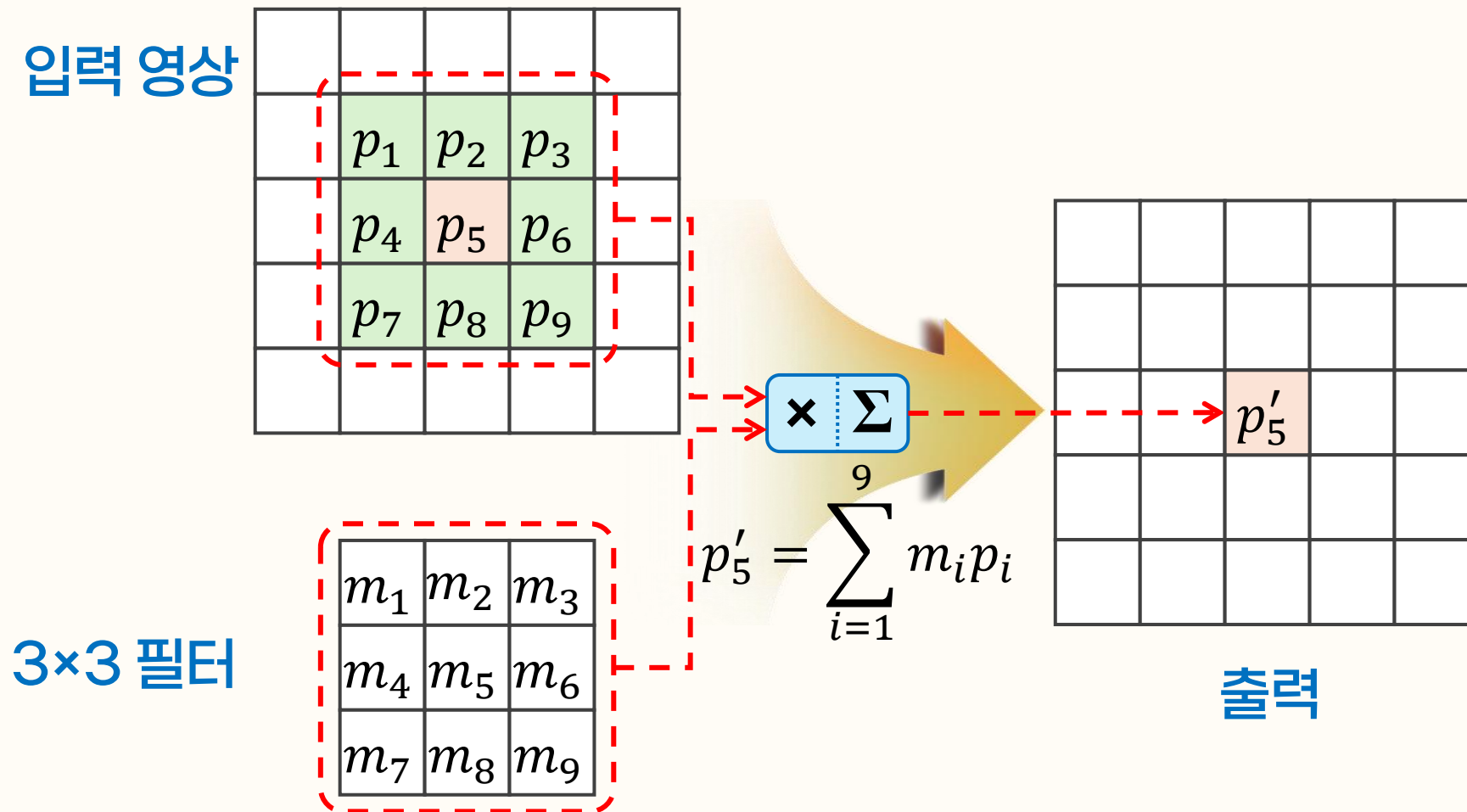
필터 마스크

$1/9$	$1/9$	$1/9$
$1/9$	$1/9$	$1/9$
$1/9$	$1/9$	$1/9$



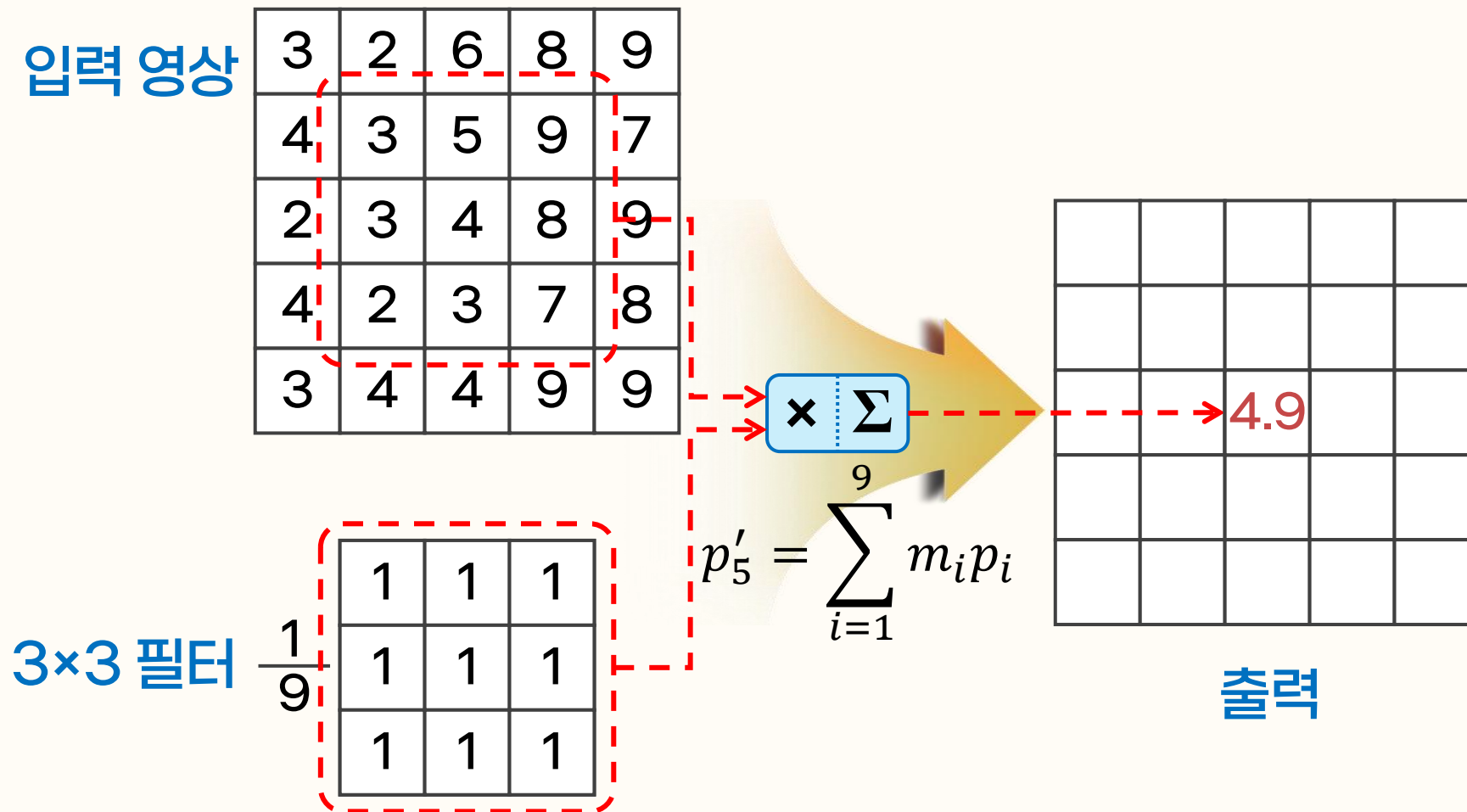
2. 합성곱을 이용한 영상 필터링

합성곱과 필터 마스크



2. 합성곱을 이용한 영상 필터링

합성곱과 필터 마스크



2. 합성곱을 이용한 영상 필터링

합성곱과 필터 마스크

입력 영상

3	2	6	8	9
4	3	5	9	7
2	3	4	8	9
4	2	3	7	8
3	4	4	9	9

3×3 필터 $\frac{1}{9}$

1	1	1
1	1	1
1	1	1

	3.6	5.3	7.2	
	3.3	4.9	6.7	
	3.2	4.9	6.8	

출력

3. 공간 영역 필터링의 예 - 가우시안 잡음의 평활화

▣ 영상의 잡음(noise, 雑音, 잡영, 雜映)

- 관측 과정에서 발생하는 불규칙한 영상 신호 성분
- 가우시안 잡음(Gaussian noise), 점 잡음(salt-and-pepper noise) 등

▣ 잡음 제거 필터

- 저역통과 필터(low-pass filter) 활용 : 고주파 성분을 억제하여 불규칙하게 변화하는 잡음을 억제
 - ➡ **평활화(smoothing)**: 잡음으로 인한 영상의 급격한 변화를 완화하여 영상 품질을 개선하는 처리

3. 공간 영역 필터링의 예 - 가우시안 잡음의 평활화

▣ 평활화를 위한 필터

- 평균 필터(상자 필터, box filter)
 - $n \times n$ 크기의 동일한 값으로 구성된 필터 마스크
 - 예: $n = 3$ 인 평균 필터

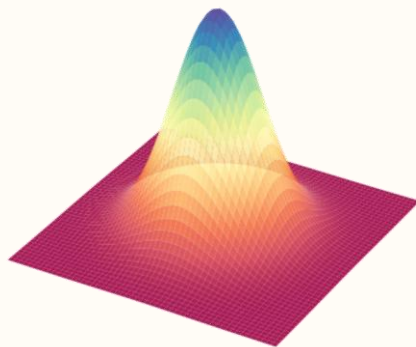
$$M_{avr} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

3. 공간 영역 필터링의 예 - 가우시안 잡음의 평활화

▣ 평활화를 위한 필터

- 가우시안 저역통과 필터(Gaussian LPF)

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



- 3×3 크기로 근사화한 가우시안 저역통과 필터 마스크

$$M_G = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

3. 공간 영역 필터링의 예 - 가우시안 잡음의 평활화

▣ 평활화 필터의 적용 예

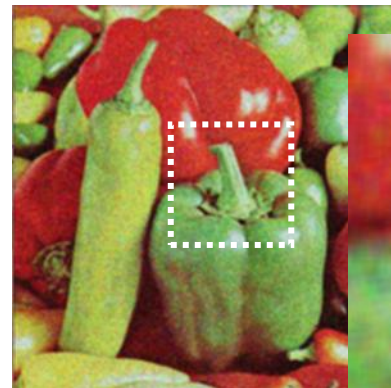
원 영상



가우시안 잡음



평균 필터



가우시안 LPF



4. 공간 영역 필터링의 예 - 점 잡음의 평활화

■ 점 잡음(salt-and-pepper noise, impulse noise)이란?

- 흰 점과 검은 점이 무작위로 나타나는 잡음
- 가우시안 필터로 효과적인 잡음 제거를 할 수 없음

원 영상



점 잡음 영상



가우시안 LPF




4. 공간 영역 필터링의 예 - 점 잡음의 평활화

■ 중간값 필터(median filter)

- 각각의 픽셀 위치에서 정해진 크기의 영역내의 픽셀 값들 중에서 중간값을 선택함
- 점잡음에 효과적임

84	81	90
85	210	87
82	89	91



84	81	90
85	87	87
82	89	91

(81 82 84 85 87 89 90 91 210)

4. 공간 영역 필터링의 예 - 점 잡음의 평활화

■ 중간값 필터(median filter)

- 중간값 필터를 이용한 점잡음 제거



가는 선이나 점 등도 함께 제거될 수 있음

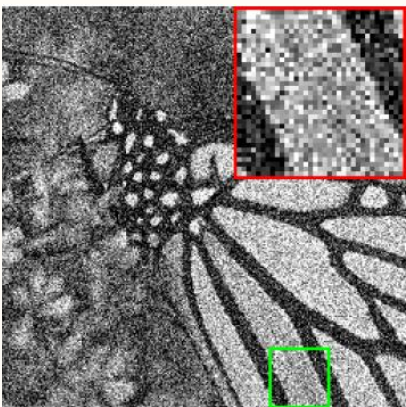
5. 딥러닝을 이용한 영상 잡음 제거

❏ 예: DnCNN(Denoising Convolutional Neural Network)

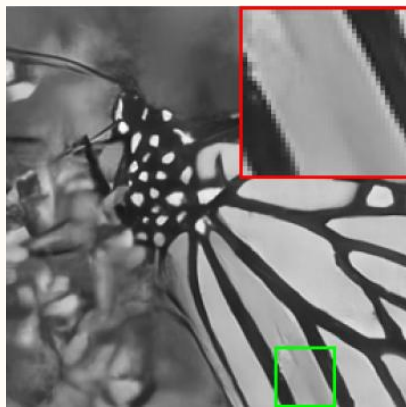
- 잔차 학습(Residual learning) 기반 모델

❏ 예: SwinIR(Swin Transformer for Image Restoration)

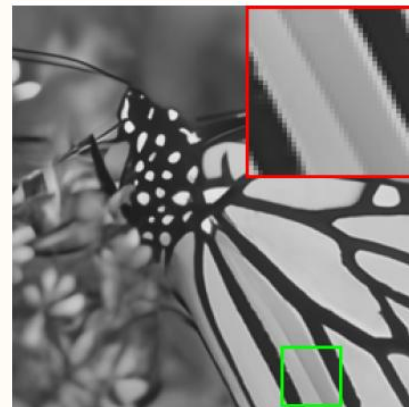
- 트랜스포머(transformer) 기반 모델



잡음 영상



DnCNN 결과



SwinIR 결과

출처: LIANG, Jingyun, et al., "Swinir: Image restoration using swin transformer," Proc. IEEE/CVF ICCV, pp. 1833–1844, 2021.



영상 분할

1. 영상분할의 개념

■ 영상분할(image segmentation)이란?

- 영상을 구성하는 물체의 영역을 구분하여 분할하는 것

➡ 유사한 속성을 갖는 부분으로 분할



영상의 밝기

색 성분

텍스처 등

2. 임계치에 의한 이진화

■ 이진화(binartization)

- 영상의 밝기에 따라 두 영역으로 구분하는 것
- 배경으로부터 전경 객체를 구분하기 위한 적절한 임계치를 선택함(thresholding)

$$b(x, y) = \begin{cases} 1 & \text{if } g(x, y) > T \\ 0 & \text{if } g(x, y) \leq T \end{cases}$$

- $g(x, y)$: 흑백 영상
- $b(x, y)$: 이진 영상

2. 임계치에 의한 이진화

■ 이진화(binartization)

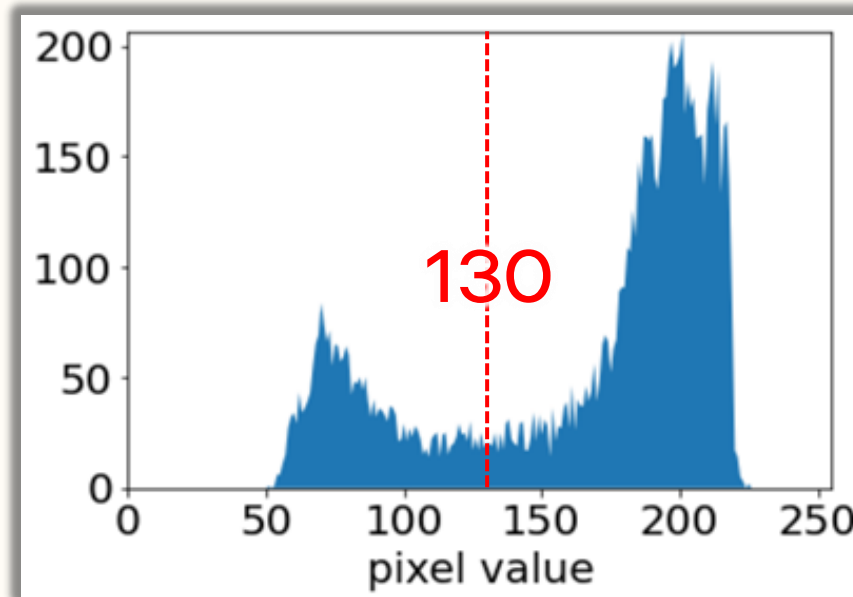
- 영상의 밝기에 따라 두 영역으로 구분하는 것
- 배경으로부터 전경 객체를 구분하기 위한 적절한 임계치를 선택함(thresholding)



흑백 영상



이진 영상



3. 영역분할 방법

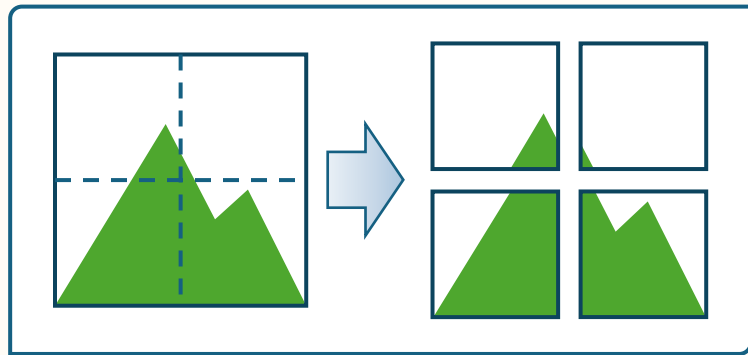
■ 영상의 공간적 특성을 이용한 방법

- 어떠한 영역이 인접 영역과 유사한 속성을 갖는다면 이들을 하나의 영역으로 결합
- 영역의 내부가 균일하지 않다면 이를 분할
- 알고리즘
 - 영역 성장(region growing)
 - 분할과 합병(split and merge)

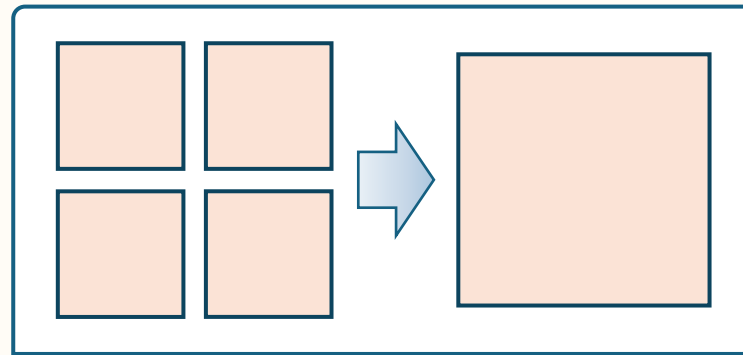
3. 영역분할 방법

■ 분할과 합병

- 4분 트리(quad tree) 분할을 이용하는 알고리즘
- 분할 및 합병 과정을 반복함



분할



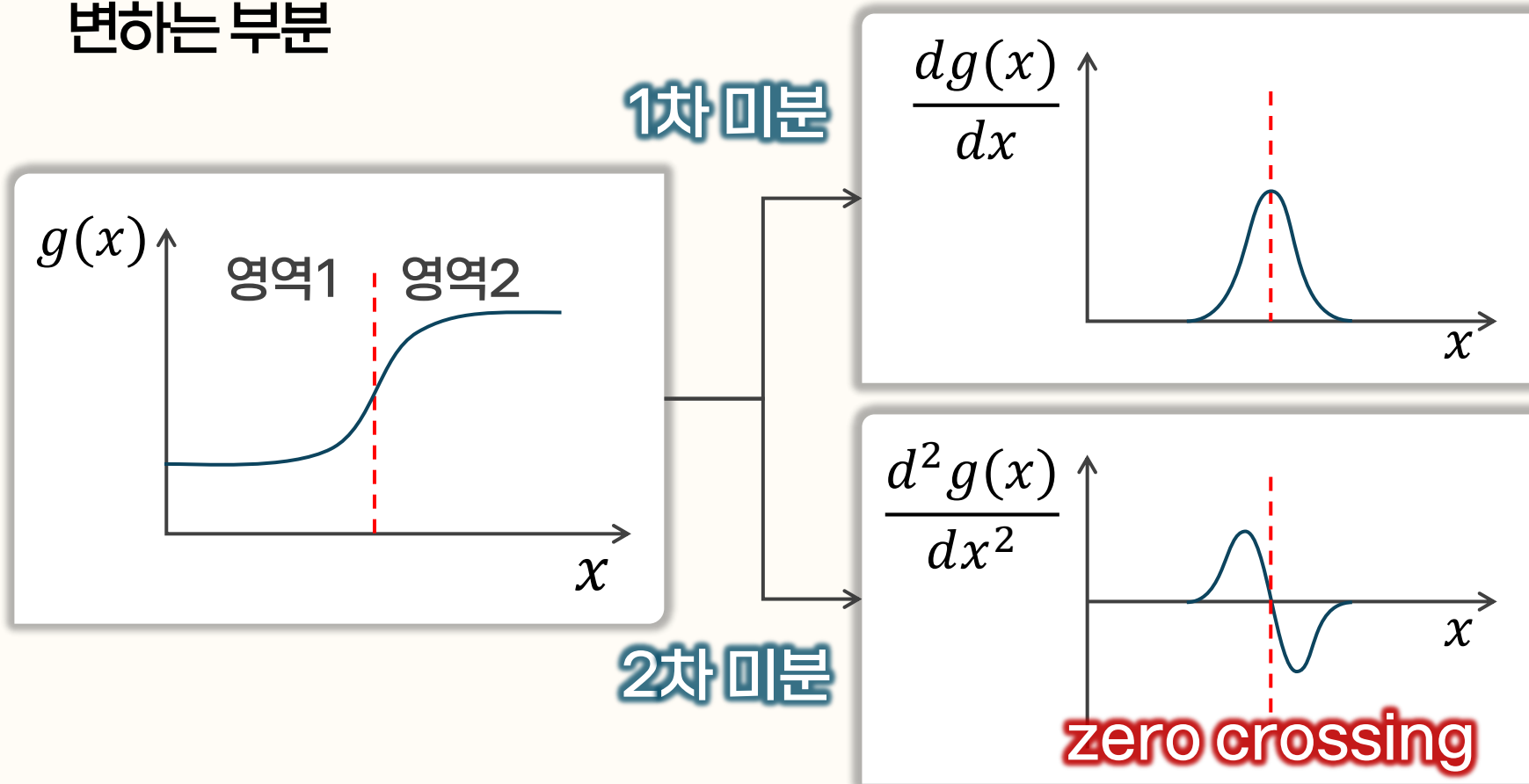
합병

- 분할 및 합병 반복 후 모든 영역들을 인접 영역과 균일성 검사를 하여 균일하다면 이들을 하나의 영역으로 합병

4. 경계검출 방법

■ 에지(edge)

- 밝기가 다른 두 영역 사이에서 픽셀의 그레이레벨이 크게 변하는 부분



4. 경계검출 방법

■ 소벨(Sobel) 연산자

- **1차 미분**으로 에지를 검출하는 연산자
- 2개의 마스크로 정의됨

$$M_{SobelH} = \frac{1}{4} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$M_{SobelV} = \frac{1}{4} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

➡ $e_h(x, y)$: M_{SobelH} 의 결과

➡ $e_v(x, y)$: M_{SobelV} 의 결과

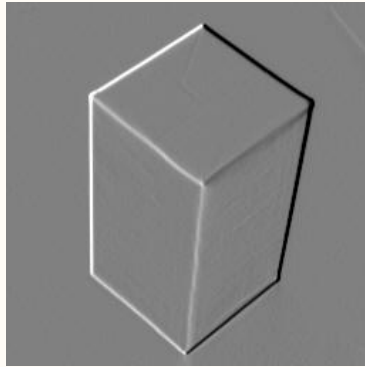
- 에지의 강도: $e_{Sobel}(x, y) = \sqrt{e_h(x, y)^2 + e_v(x, y)^2}$
- 에지의 방향: $\theta_{Sobel}(x, y) = \tan^{-1} \frac{e_v(x, y)}{e_h(x, y)}$

4. 경계검출 방법

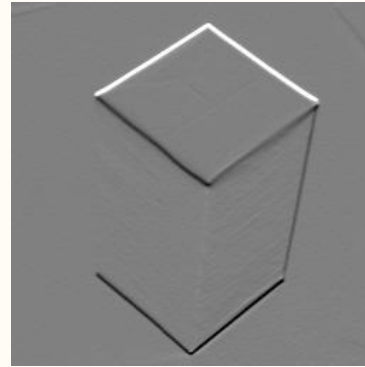
■ 소벨(Sobel) 연산자



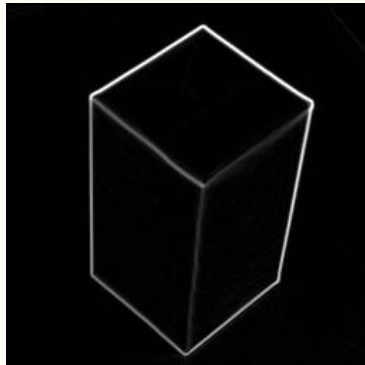
원 영상



M_{SobelH} 의 결과



M_{SobelV} 의 결과



에지 강도

4. 경계검출 방법

■ 라플라스(Laplace) 연산자

- **2차 미분**으로 에지를 검출하는 연산자

$$\begin{aligned} f(x, y) &= \nabla^2 g(x, y) \\ &= \frac{\partial^2}{\partial x^2} g(x, y) + \frac{\partial^2}{\partial y^2} g(x, y) \end{aligned}$$

- 근사화한 3×3 필터 마스크

$$M_{Lap} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

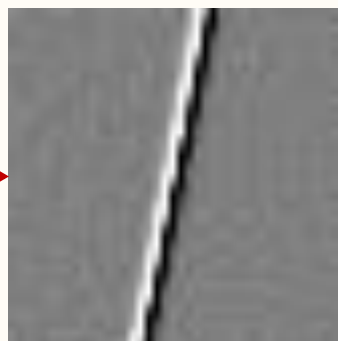
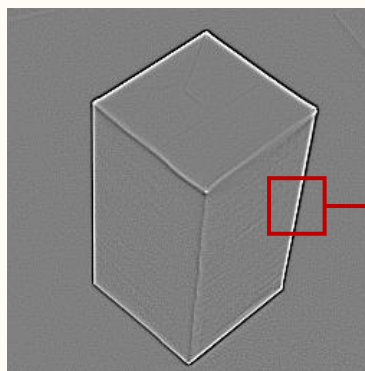
4. 경계검출 방법

라플라스(Laplace) 연산자

원 영상



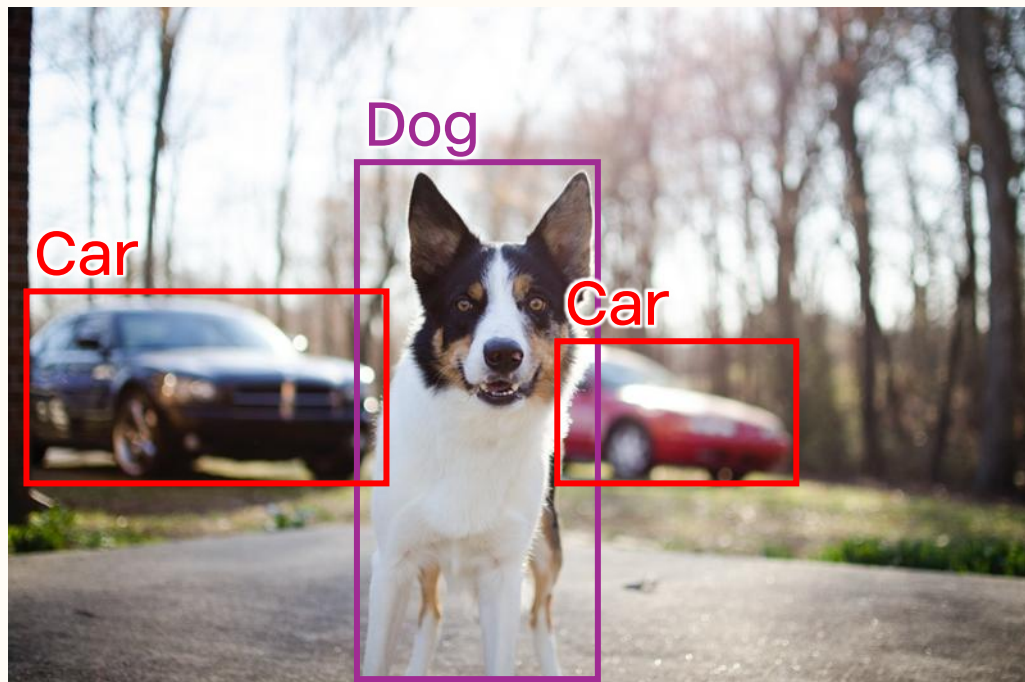
라플라스
연산 결과



5. 심층 신경망 모델을 이용한 영상 분할

■ 객체 검출

- 영상 안에 포함된 객체의 유형 및 위치를 찾아내는 처리

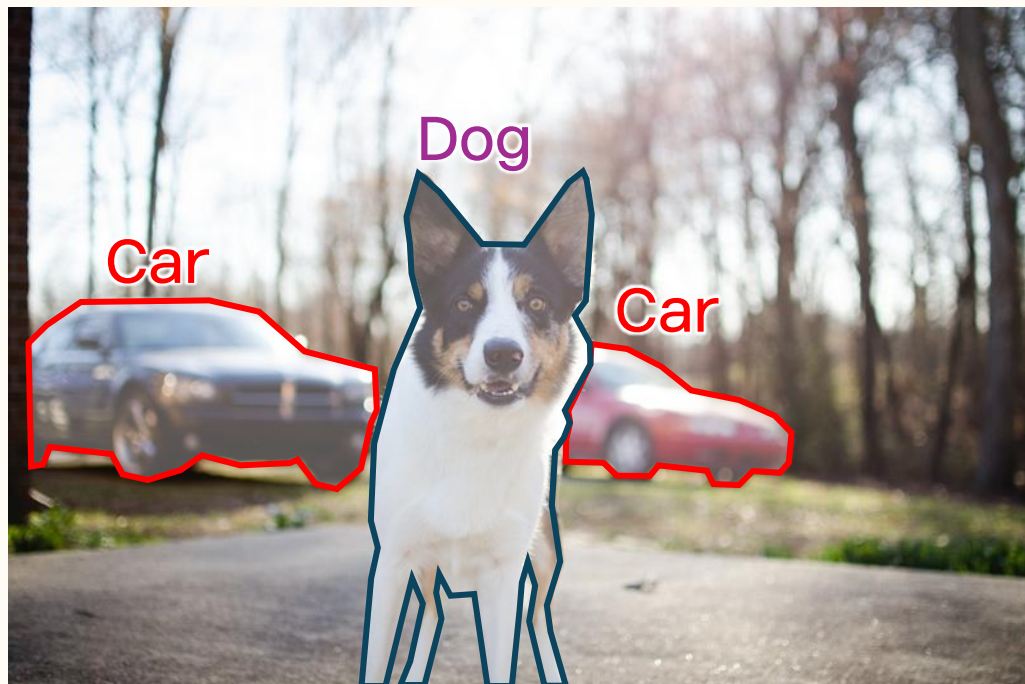


- 대표적 모델: R-CNN, YOLO 등

5. 심층 신경망 모델을 이용한 영상 분할

■ 인스턴스 분할(instance segmentation)

- 영상 안에 포함된 개별 인스턴스들의 영역을 구분하는 처리



- 대표적 모델: Mask R-CNN, SegFormer 등

정리하기

- ✓ 컴퓨터 시각(computer vision)은 인간의 시각체계에 해당되는 능력을 컴퓨터에 심어 넣기 위해 연구하는 인공지능의 한 분야이다.
- ✓ 아날로그 영상 신호를 표본화와 양자화를 통해 사각형 격자 구조로 나열된 픽셀들로 구성함으로써 디지털 영상을 입력한다.
- ✓ 공간 영역에서 필터를 적용하는 것은 $n \times n$ 크기의 필터 마스크를 각각의 픽셀 위치에 겹쳐 놓은 후 각각의 값을 곱하여 더한 값을 구하는 것이다.

정리하기

- ✓ 가우시안 저역통과 필터를 사용하면 영상에 포함된 고주파 잡음을 완화할 수 있다.
- ✓ 점잡음은 중간값 필터를 사용하면 효과적으로 제거할 수 있다.
- ✓ 영상분할이란 영상 내의 유사한 속성을 갖는 영역들을 분리함으로써 물체의 영역을 구분하는 처리이다.
- ✓ 영상의 잡음 제거나 객체 검출, 다양한 시각 처리에 심층 신경망 모델을 이용되고 있다.

09강

다음시간안내 ▶▶▶

컴퓨터 시각과 패턴인식(2)