디지털영상처리II (2021학년도 2학기)

6주: 영상특징과 서술자(3)

김 남 규 (ngkim@deu.ac.kr)

디지털영상처리II (2021학년도 2학기)

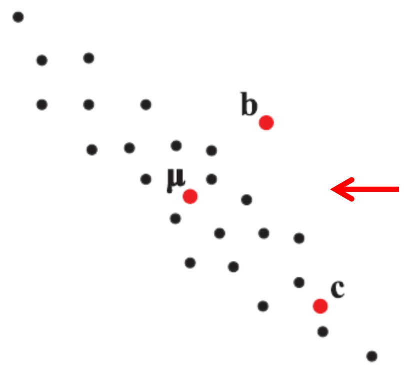
6주: 영상특징과 서술자(3)

1 매칭과 RANSAC

김 남 규 (ngkim@deu.ac.kr)

거리 척도

- Euclidean distance & Mahalanobis distance



← μ는 b와 c 중 누구와 더 가까운가?

$$\Sigma = \begin{bmatrix} var(x) & cov(x,y) \\ cov(x,y) & var(y) \end{bmatrix}$$

$$cov(x,y) = \frac{\sum_i^n (x_i - \mu) \cdot (y_i - \mu)}{N}$$

- Euclidean distance: $d_E(\mathbf{a},\mathbf{b}) = \|\mathbf{a}-\mathbf{b}\| = \sqrt{\sum_{i=1}^{d}(a_i - b_i)^2}$ 예) μ와 b가 가까움

- Mahalanobis distance: 공분산 행렬(covariance matrix)을 이용한 확률 분포

$$d_M(\mathbf{a},\mathbf{b}) = \sqrt{(\mathbf{a}-\mathbf{b})\boldsymbol{\Sigma}^{-1}(\mathbf{a}-\mathbf{b})^{\mathrm{T}}}$$

예) μ와 c가 가까움

아핀 변환 추정: 최소제곱법(LSM, Least Squared Method)

- $X=\{(\mathbf{a}_1,\mathbf{b}_1),(\mathbf{a}_2,\mathbf{b}_2),\cdots,(\mathbf{a}_n,\mathbf{b}_n)\}$

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{pmatrix} \qquad (b'_{i1}\ \ b'_{i2}\ \ 1) = (a_{i1}\ \ a_{i2}\ \ 1)\begin{pmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{pmatrix}$$

- Error 함수: $E(\mathbf{T}) = \sum_{i=1}^{n} \| \mathbf{b}_i - \mathbf{b}'_i \|^2$

$$= \sum_{i=1}^{n} \left((b_{i1} - (t_{11}a_{i1} + t_{21}a_{i2} + t_{31}))^2 + (b_{i2} - (t_{12}a_{i1} + t_{22}a_{i2} + t_{32}))^2 \right)$$

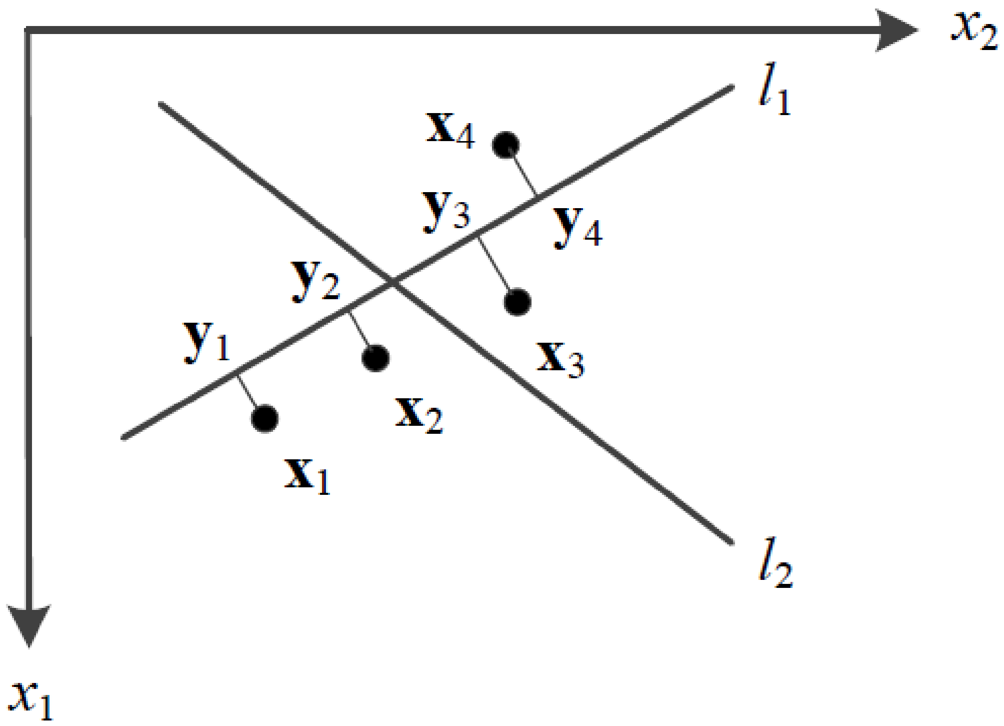
  - Error를 최소화하는 T? : $\dfrac{\partial E}{\partial t_{ij}} = 0$

LSM 아핀 변환 계산

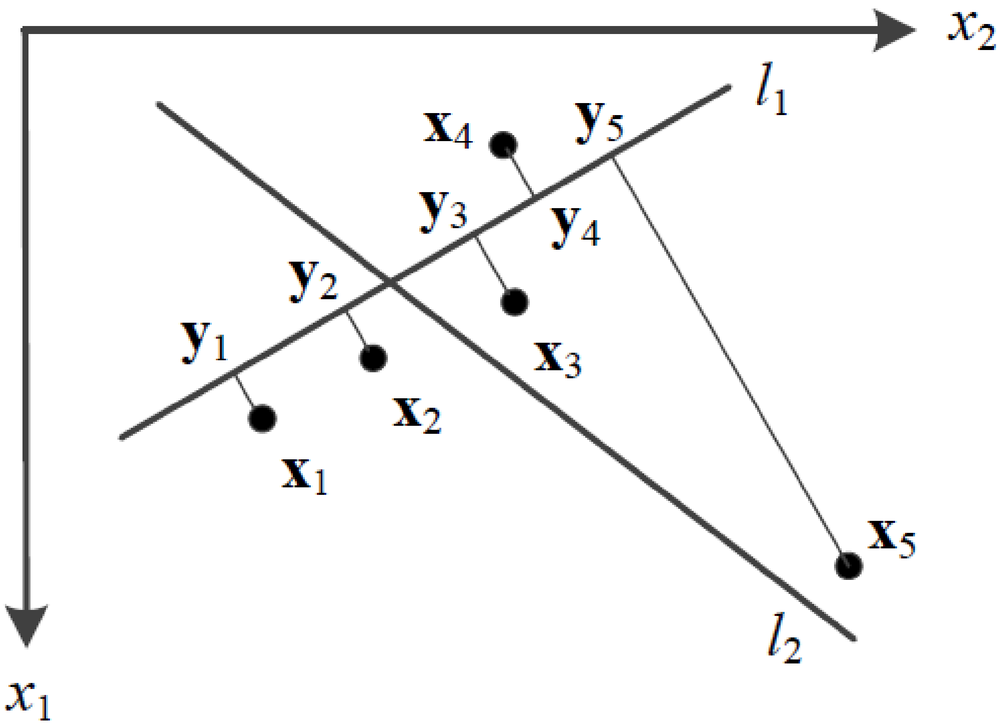
$$
\begin{pmatrix}
\sum_{i=1}^{n} a_{i1}^{\;2} & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1} & 0 & 0 & 0 \\
\sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i2}^{\;2} & \sum_{i=1}^{n} a_{i2} & 0 & 0 & 0 \\
\sum_{i=1}^{n} a_{i1} & \sum_{i=1}^{n} a_{i2} & \sum_{i=1}^{n} 1 & 0 & 0 & 0 \\
0 & 0 & 0 & \sum_{i=1}^{n} a_{i1}^{\;2} & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1} \\
0 & 0 & 0 & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1}^{\;2} & \sum_{i=1}^{n} a_{i2} \\
0 & 0 & 0 & \sum_{i=1}^{n} a_{i1} & \sum_{i=1}^{n} a_{i2} & \sum_{i=1}^{n} 1
\end{pmatrix}
\begin{pmatrix} t_{11} \\ t_{21} \\ t_{31} \\ t_{12} \\ t_{22} \\ t_{32} \end{pmatrix}
=
\begin{pmatrix}
\sum_{i=1}^{n} a_{i1} b_{i1} \\
\sum_{i=1}^{n} a_{i2} b_{i1} \\
\sum_{i=1}^{n} b_{i1} \\
\sum_{i=1}^{n} a_{i1} b_{i2} \\
\sum_{i=1}^{n} a_{i2} b_{i2} \\
\sum_{i=1}^{n} b_{i2}
\end{pmatrix}
$$

LSM의 한계

- 이상점(아웃라이어, outlier) 에 의해 잘못된 결과 도출



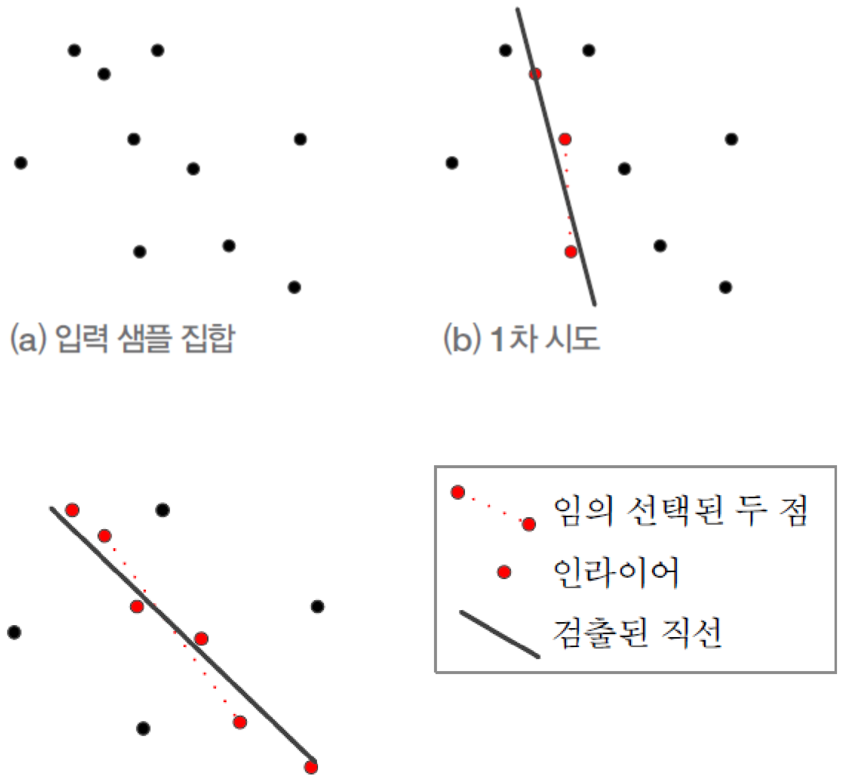
(a) 원리



(b) 한계

RANSAC

- RANdom SAmple Consensus(무작위 표본 합의)



(a) 입력 샘플 집합
(b) 1차 시도
(c) 2차 시도

$$\begin{pmatrix} \sum_{i=1}^{n} a_{i1}^{2} & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1} & 0 & 0 & 0 \\ \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i2}^{2} & \sum_{i=1}^{n} a_{i2} & 0 & 0 & 0 \\ \sum_{i=1}^{n} a_{i1} & \sum_{i=1}^{n} a_{i2} & \sum_{i=1}^{n} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sum_{i=1}^{n} a_{i1}^{2} & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1} \\ 0 & 0 & 0 & \sum_{i=1}^{n} a_{i1}a_{i2} & \sum_{i=1}^{n} a_{i1}^{2} & \sum_{i=1}^{n} a_{i2} \\ 0 & 0 & 0 & \sum_{i=1}^{n} a_{i1} & \sum_{i=1}^{n} a_{i2} & \sum_{i=1}^{n} 1 \end{pmatrix} \begin{pmatrix} t_{11} \\ t_{21} \\ t_{31} \\ t_{12} \\ t_{22} \\ t_{32} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{i1}b_{i1} \\ \sum_{i=1}^{n} a_{i2}b_{i1} \\ \sum_{i=1}^{n} b_{i1} \\ \sum_{i=1}^{n} a_{i1}b_{i2} \\ \sum_{i=1}^{n} a_{i2}b_{i2} \\ \sum_{i=1}^{n} b_{i2} \end{pmatrix}$$

$Q = \emptyset$;
for(j=1 to k) {
  X에서 세 개 대응점 쌍을 임의로 선택한다.
  이들 세 쌍을 입력으로 $\boxed{}$를 풀어 $\mathbf{T}_j$를 추정한다.
  이들 세 쌍으로 집합 *inlier*를 초기화한다.
  for(이 세 쌍을 제외한 X의 요소 p 각각에 대해) {
    if(p가 허용 오차 t 이내로 $\mathbf{T}_j$에 적합) p를 *inlier*에 넣는다.
  }
  if($|inlier| \geq d$) // 집합 *inlier*가 d개 이상의 샘플을 가지면
    *inlier*에 있는 모든 샘플을 가지고 새로운 $\mathbf{T}_j$를 계산한다.
  if($\mathbf{T}_j$의 적합 오류$<e$) $\mathbf{T}_j$를 집합 Q에 넣는다.
}
Q에 있는 변환 행렬 중 가장 좋은 것을 $\mathbf{T}$로 취한다.

디지털영상처리II (2021학년도 2학기)

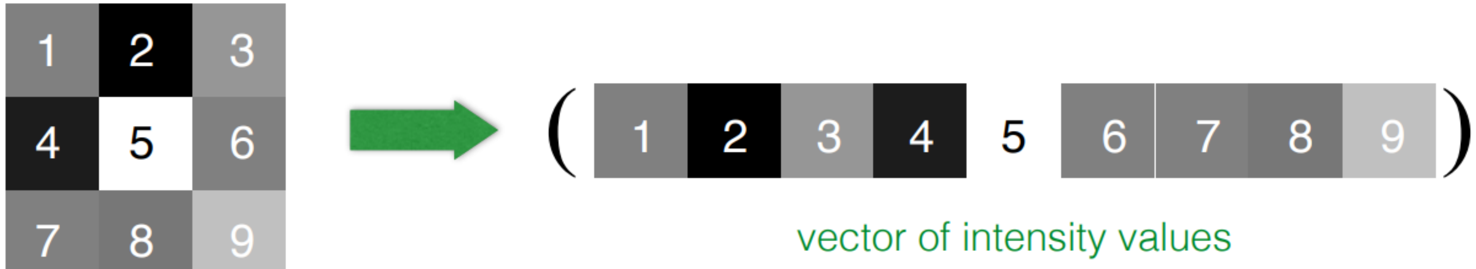
6주: 영상특징과 서술자(3)

2 HOG(Histogram of Gradient)

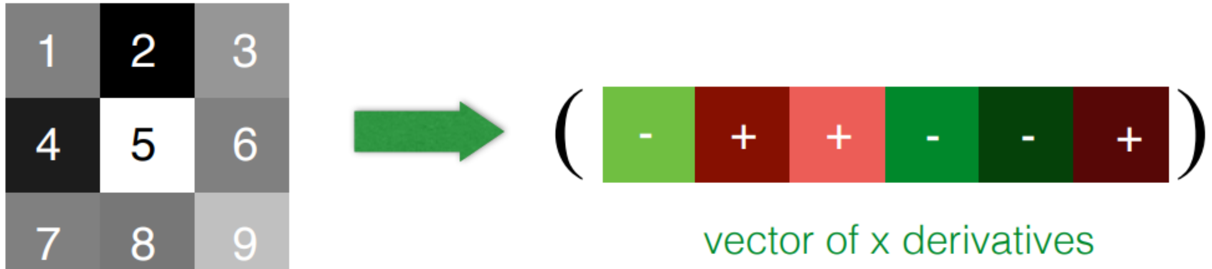
김 남 규 (ngkim@deu.ac.kr)

지역 특징 서술자

- 회전 + 크기 불변 기술자 찾기
- Image patch: 형태 불변
  값의 크기에 의존적



- Image gradient: 크기 불변
  형태 변화에 의존적

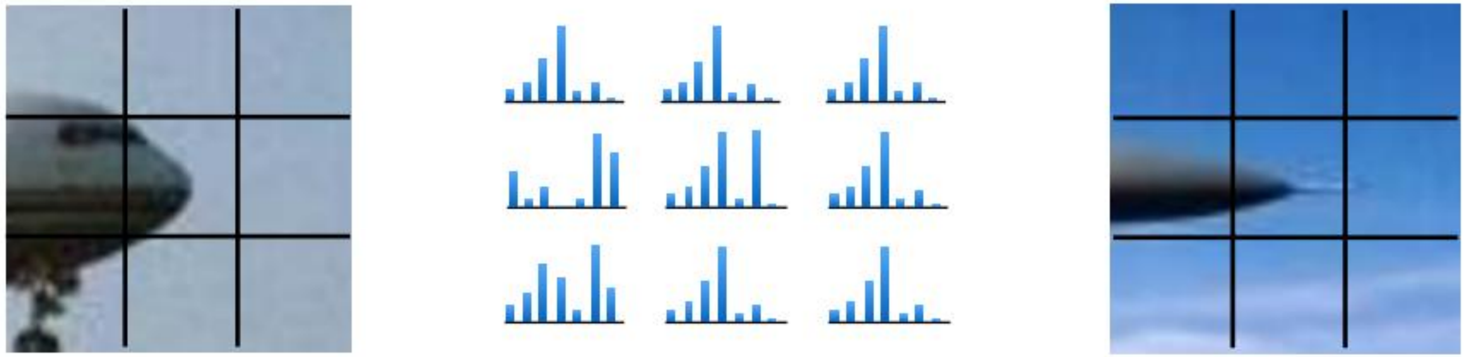


- Color histogram: 크기와
  회전에 불변

Histogram

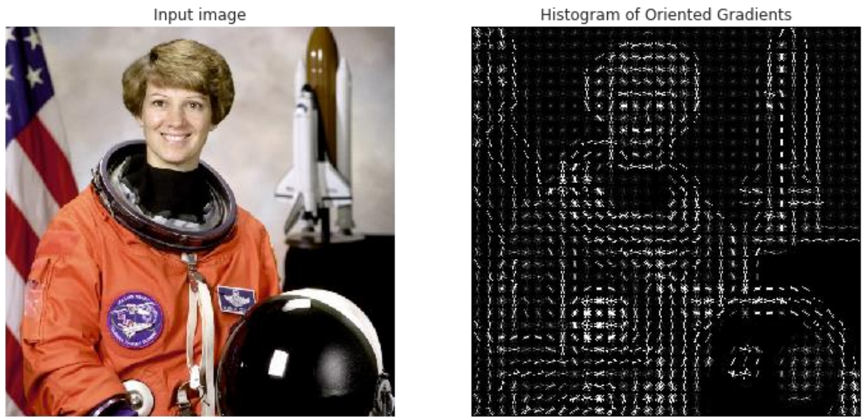
- Color histogram:
  유사한 레이아웃에서 문제

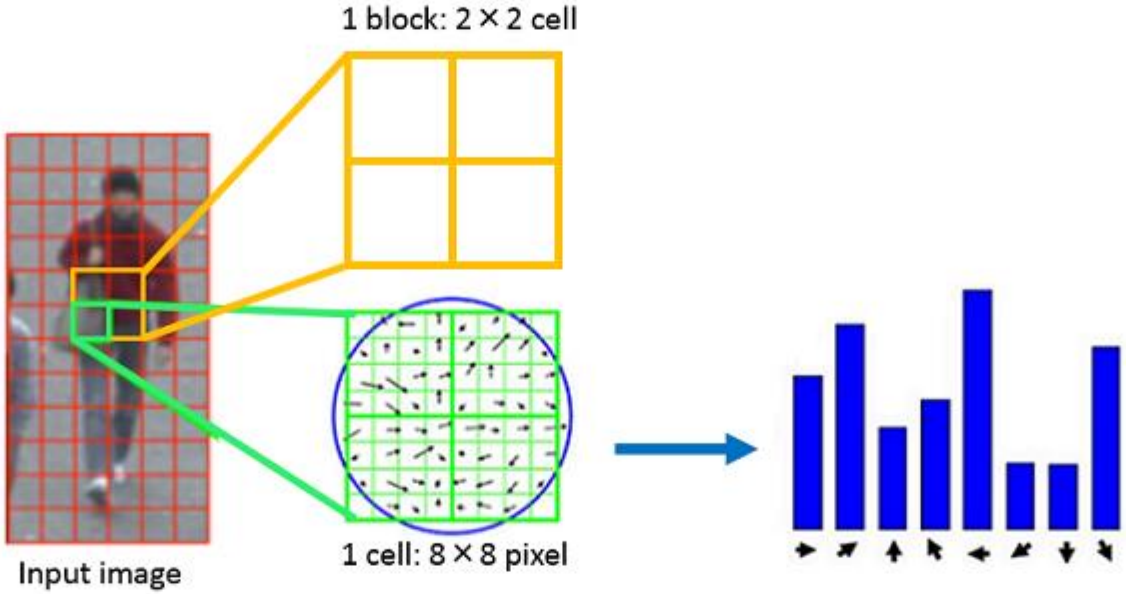
- Spatial histogram:
  레이아웃 구별 가능하나,
  회전 불변하지 않음



Histogram of oriented gradient (HoG)

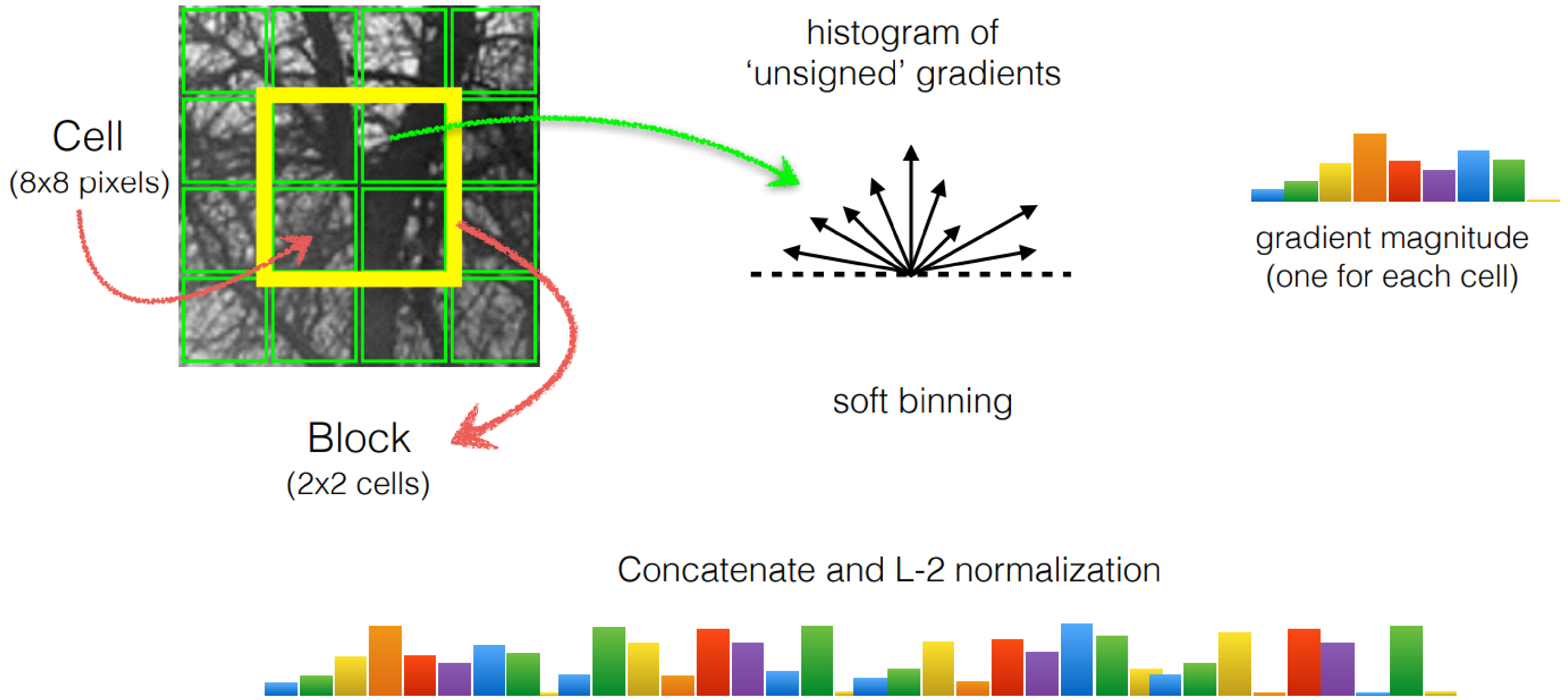
- Gradient 방향 히스토그램 정의





- Orientation 방향 개수 (bins) : 예) 8개
- Cell: 방향 계산을 위해 주변 픽셀 크기 : 예) 8x8
- Block: 주변 특성을 정하는 크기 : 예) 2x2

HOG 특징 서술자

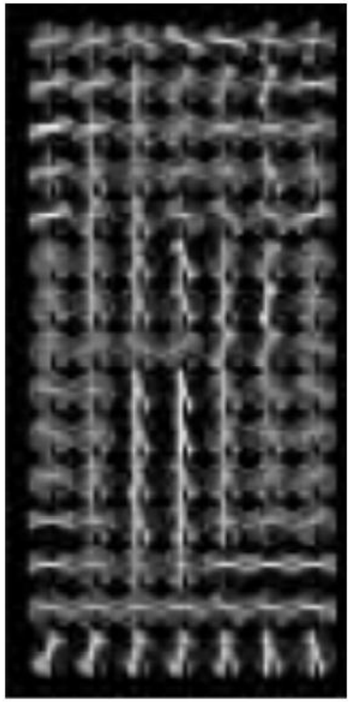


Cell
(8x8 pixels)
histogram of
'unsigned' gradients
gradient magnitude
(one for each cell)
soft binning
Block
(2x2 cells)
Concatenate and L-2 normalization

HOG Detector

1 cell step size

visualization



128 pixels
16 cells
15 blocks

15 x 7 x 4 x 36 =
3780

64 pixels
8 cells
7 blocks

Integral Image(적분 영상)

1	5	2	3
2	4	1	1
2	1	1	2
1	2	3	1

입력영상

0	0	0	0	0
0	1	6	8	11
0	3	12	15	19
0	5	15	19	25
0	6	18	25	32

적분영상

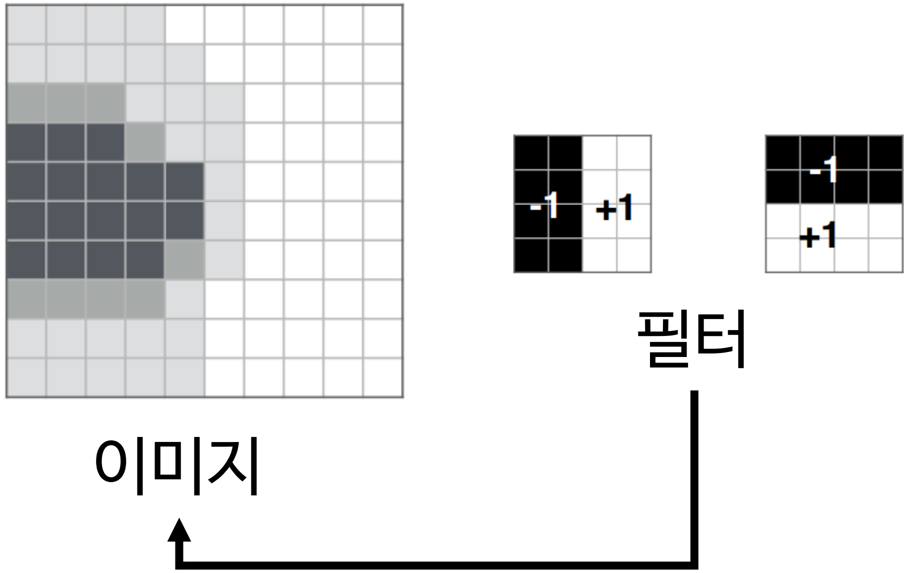
1	5	2	3
2	4	1	1
2	1	1	2
1	2	3	1

$$A(x,y) = \sum_{x' \le x, y' \le y} I(x', y')$$

S=RightBottom + LeftTop – RightTop – LeftBottom

= 32 + 6 – 11 – 18 = 9

적분 영상 활용



S = 15 + 0 – 0 – 3 = 12

1	5	2	3
2	4	1	1
2	1	1	2
1	2	3	1

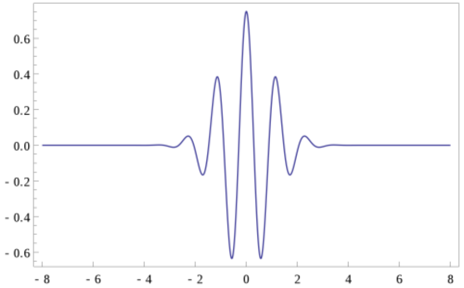
0	0	0	0	0
0	1	6	8	11
0	3	12	15	19
0	5	15	19	25
0	6	18	25	32

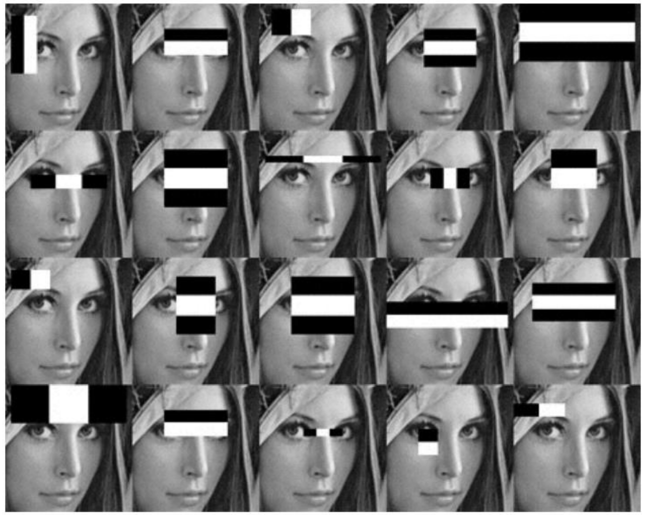
입력 영상에서 필터 연산: 7(+), 7(+), +/- → 15 번

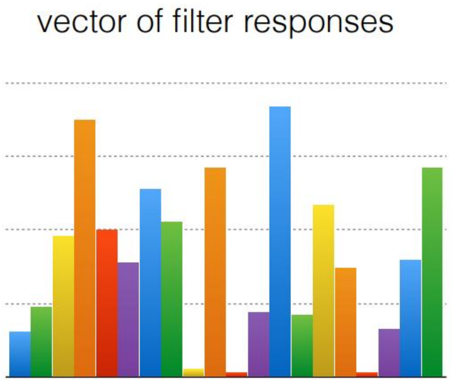
적분 영상에서 필터 연산: 3(+/-), 3(+/-), +/- → 7 번

필터 반응과 특징 서술자

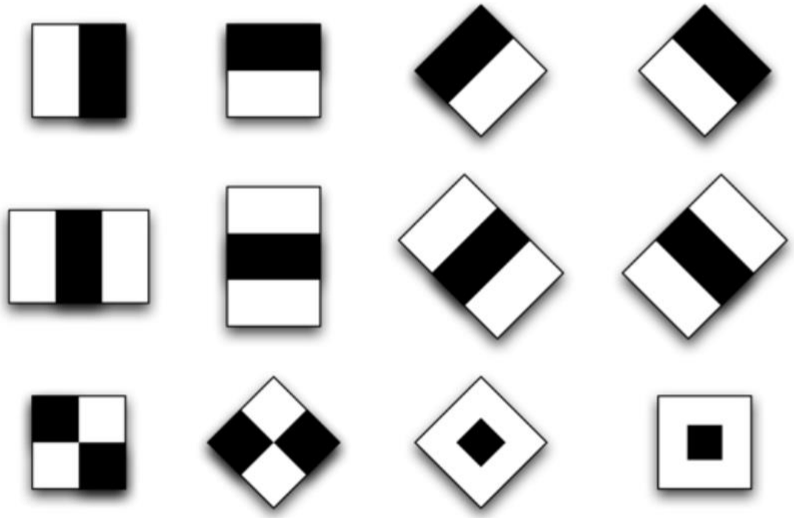
- Haar Wavelets

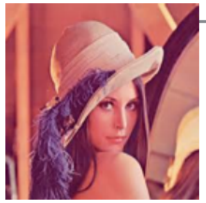


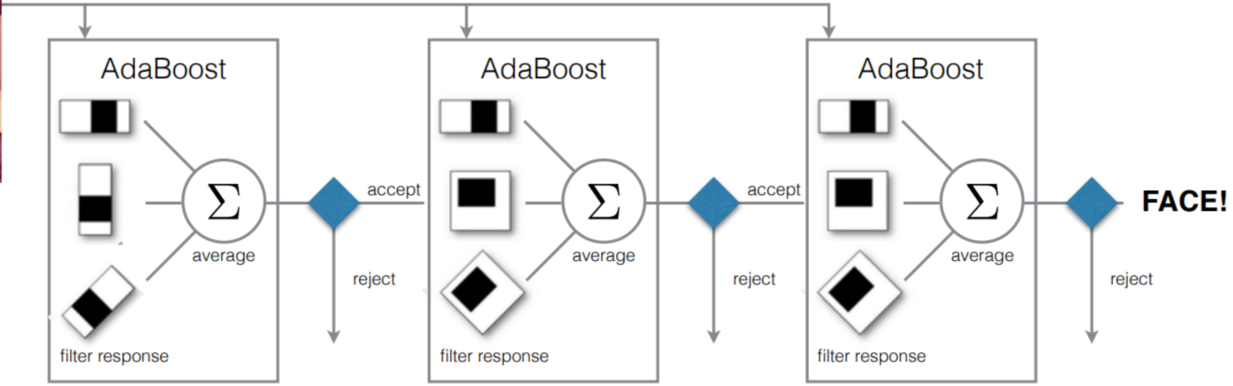




- Bank of filters





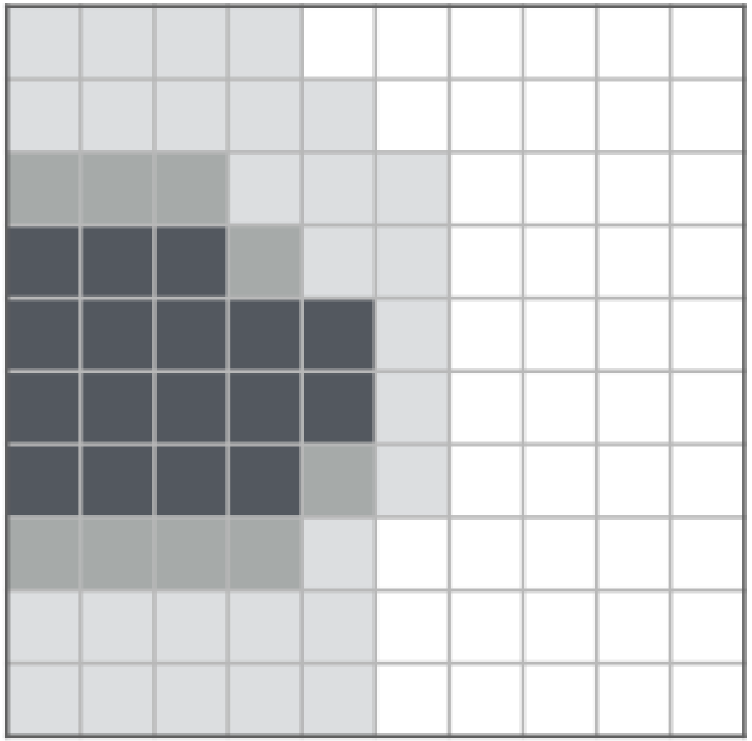


6주: 영상특징과 서술자(3)

3 BRIEF, ORB, SURF, SIFT

김 남 규 (ngkim@deu.ac.kr)

특징 서술자 형식



Keypoints: 관심이 있을 위치 예) 코너, 에지 등

Patch: keypoint를 중심으로 주변 픽셀 공간
patch_size로 정의

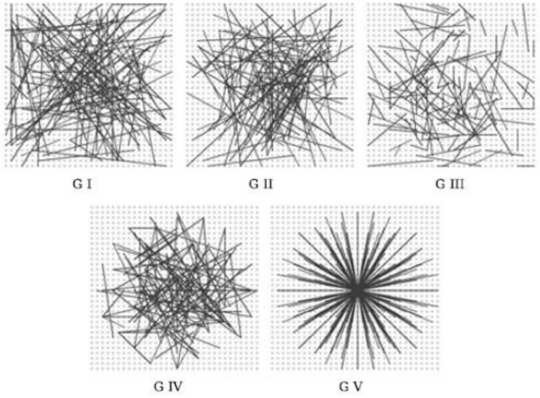
Feature: Patch 에서의 정보

Descriptor : feature 의 모임 정보

BRIEF

- Binary Robust Independent Elementary Features

$$Where\ \tau(p;x,y)\ is\ defined\ as:$$
$$\tau(p;x,y)=\begin{cases}1 & : p(x)<p(y)\\ 0 & : p(x)\geq p(y)\end{cases}$$
$$p(x)\ is\ the\ intensity\ value\ at\ pixel\ x.$$



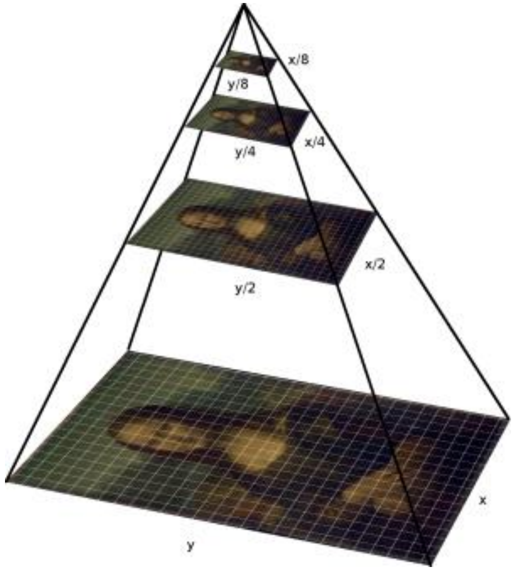
  - 패치(Patch size 는 일반적으로 49) 에서 무작위 쌍으로 계산
  - 보통 256개 쌍으로 부터 2진화 벡터 정의

$$f(n)=\sum_{1<i<n} 2^{i-1}\tau(p;x_i,y_i)$$

- 회전과 크기 변환에 모두 불변하지 못함

ORB(Oriented FAST and rotated BRIEF)

- 크기 불변: image pyramid

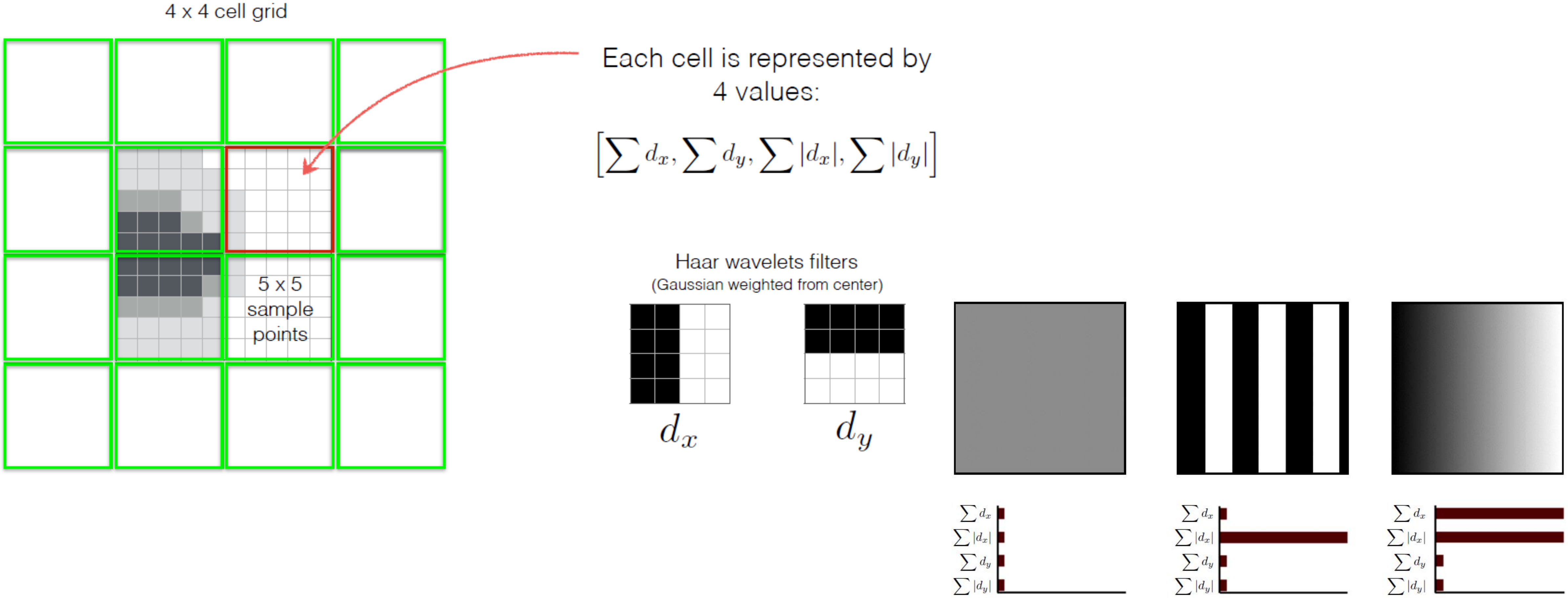


- 회전 불변: rBRIEF
  - 12도 씩 30개의 벡터 정의

- 특징 벡터의 매칭 거리 측정: 해밍 거리
  - '1011101'과 '1001001'사이의 해밍 거리는 2이다.
    (1011101, 1001001)
  - '2143896'과 '2233796'사이의 해밍 거리는 3이다.
    (2143896, 2233796)
  - "toned"와 "roses"사이의 해밍 거리는 3이다.
    (toned, roses)

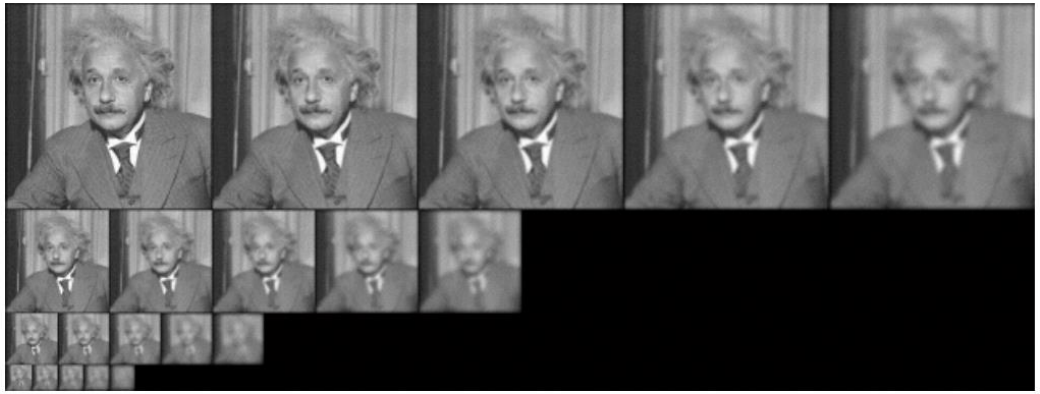
SURF(Speed Up Robust Feature)

- 차원: 4 x 4 x 4 = 64



SIFT (Scale Invariant Feature Transform) (1/2)

- Keypoint : multi-scale extreme detection



Gaussian

Laplacian

$$f(\mathbf{x}) = f + \frac{\partial f}{\partial \mathbf{x}}^T \mathbf{x} + \frac{1}{2}\mathbf{x}^T \frac{\partial^2 f}{\partial \mathbf{x}^2}\mathbf{x}$$

$$\mathbf{x} = \{x, y, \sigma\}$$

$$m(x,y) = \sqrt{\underbrace{(L(x+1,y) - L(x-1,y))^2}_{\text{x-derivative}} + \underbrace{(L(x,y+1) - L(x,y-1))^2}_{\text{y-derivative}}}$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y)))$$

$$\{x, y, \sigma, \theta\}$$

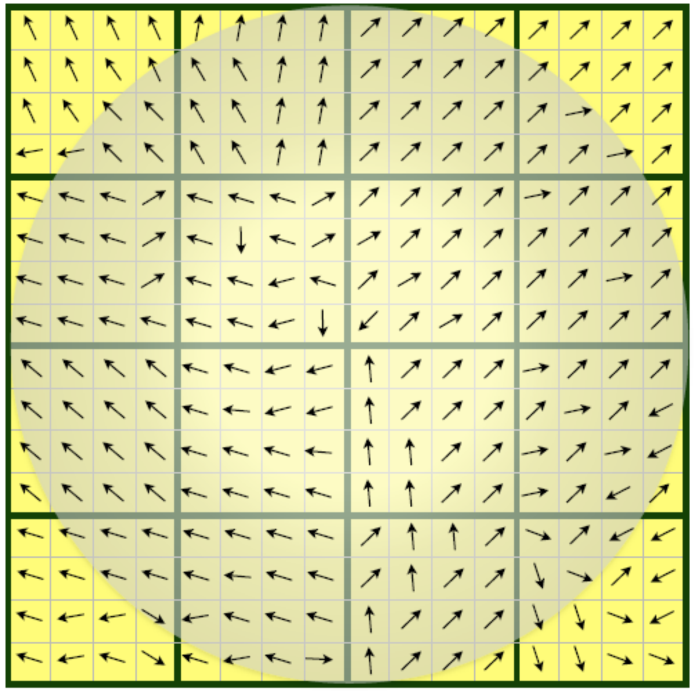
location scale orientation

SIFT (Scale Invariant Feature Transform) (2/2)

- descriptor

Image Gradients

(4 x 4 pixel per cell, 4 x 4 cells)

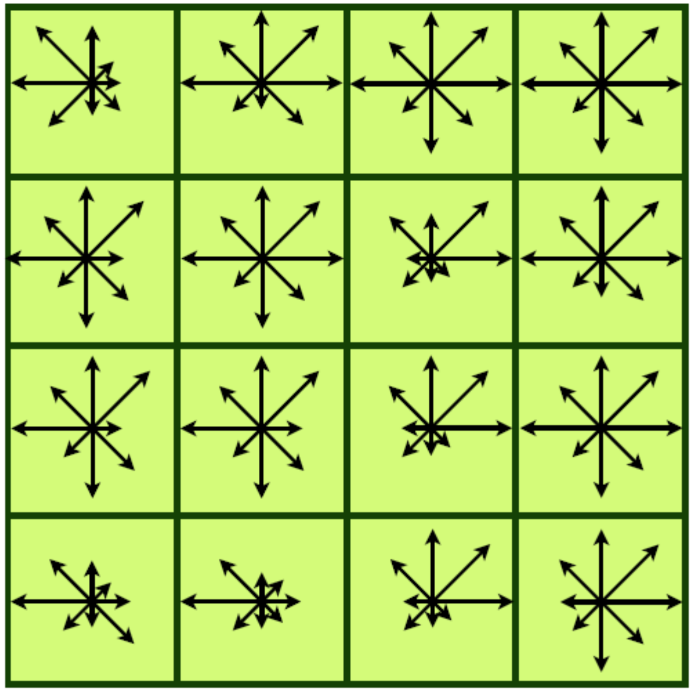


Gaussian weighting

(sigma = half width)

SIFT descriptor

(16 cells x 8 directions = 128 dims)



6주차 : 영상특징과 서술자

본 강의 자료의 내용 및 그림은 아래 책으로부터 발췌 되었음

- 파이썬으로 배우는 영상처리, Sandipan Dey 지음, 정성환 , 조보호 , 배종욱 옮김, 도서출판 홍릉, 2020년
- Digital Image Processing, 4th Ed., Rafael C . Gonzalez, Richard E. Woods 지음, Pearson, 2018년
- 컴퓨터 비전(Computer Vision) 기본 개념부터 최신 모바일 응용 예까지 IT CookBook, 오일석 지음, 한빛아카데미, 2014년

김 남 규 (ngkim@deu.ac.kr)