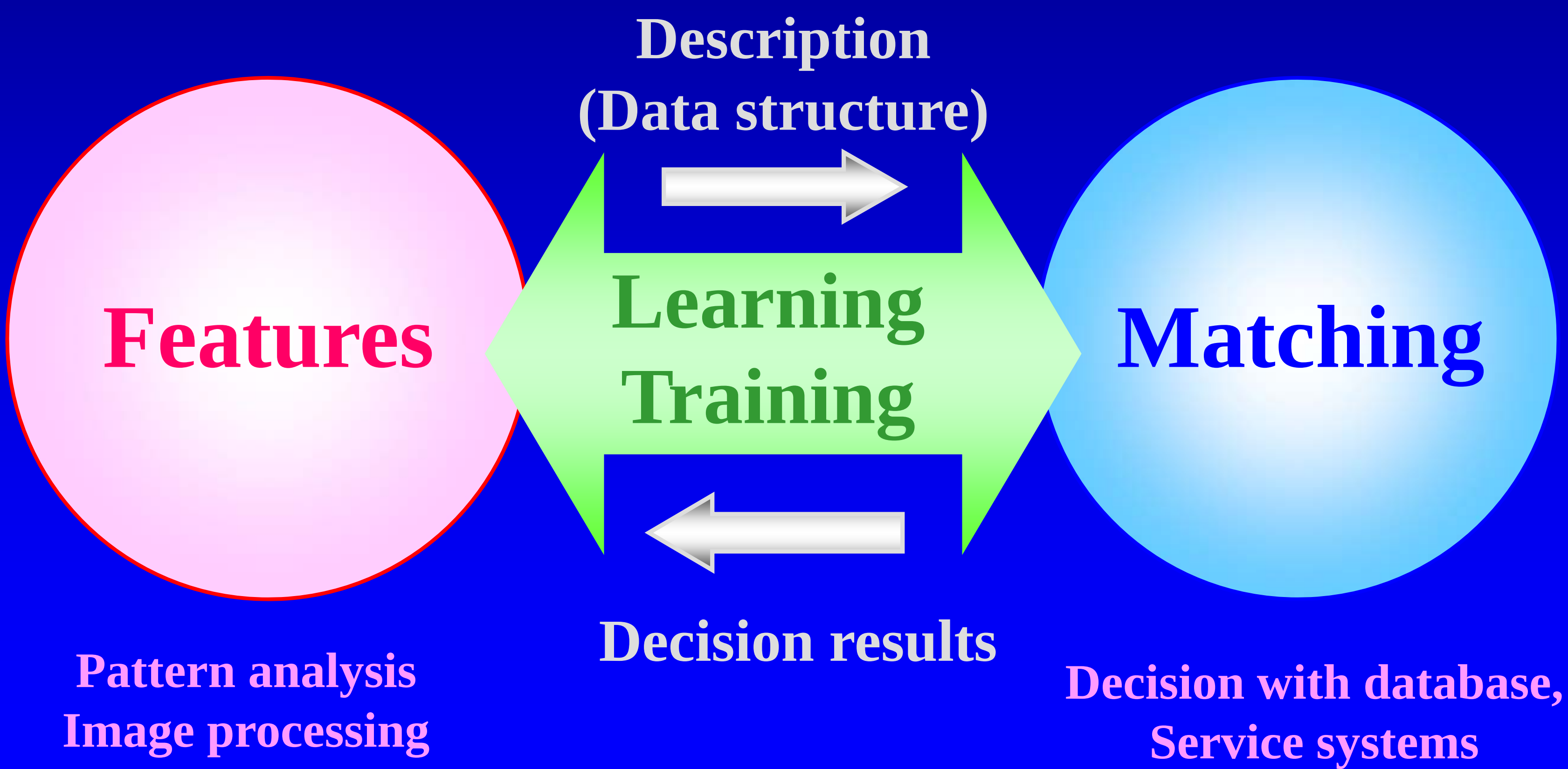
SIFT

Lecturer: Sang Hwa Lee

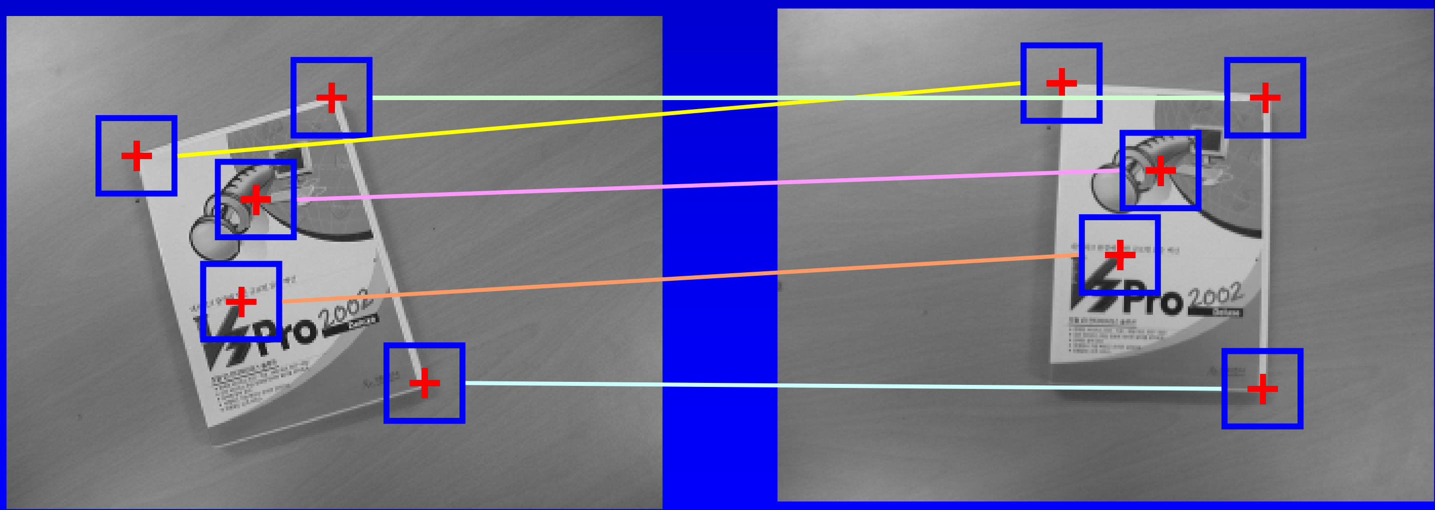
Introduction (I)

- **Elements of feature-based pattern recognition**



Introduction (II)

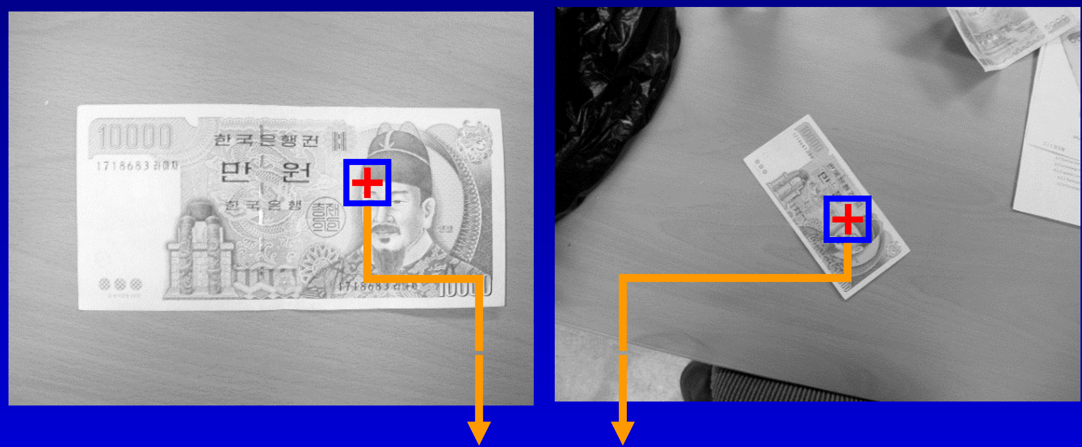
- **Feature-based pattern recognition**
  - Finding interest points
  - Analyzing features around the points
  - Comparing the features with those in database

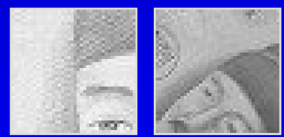


Introduction (III)

- **Variation of features**
  - Noise
  - Image scale and resolution
  - Rotation
  - 2D/3D projective transformation
  - Illumination change
- **We have to normalize the features with respect to the variations**
- **We need the invariant features for robust pattern matching**
  - SIFT (Scale Invariant Feature Transform)

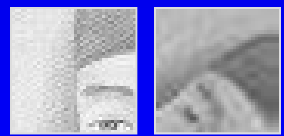
How to make invariant feature?





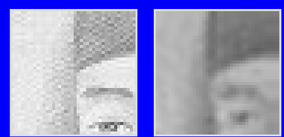
Scale & Rotation Variation

Scale normalization



~~Scale~~ & Rotation Variation

Scale & rotation normalization



~~Scale & Rotation~~ Variation

What is the Invariant Feature ?

Invariant to

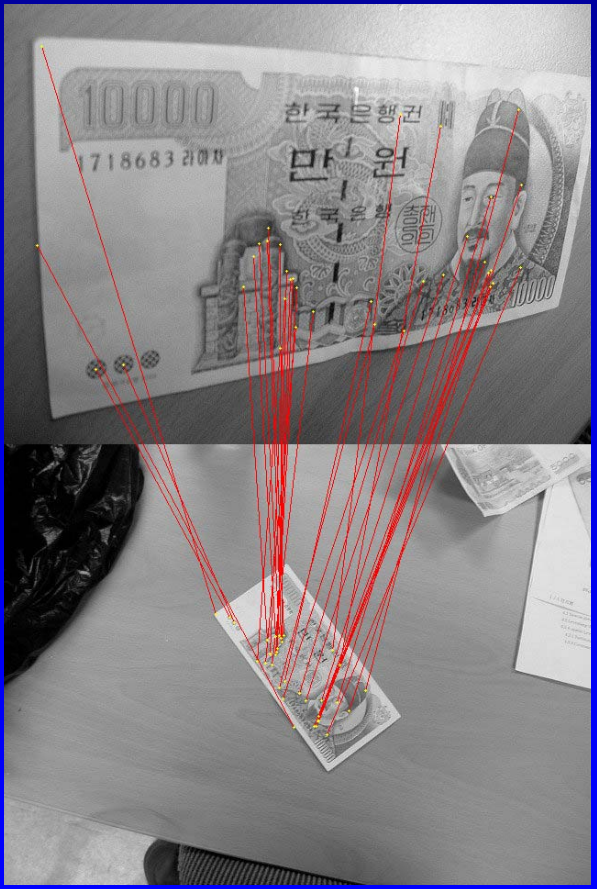
scale, resolution, rotation, illumination, perspective distortion, …

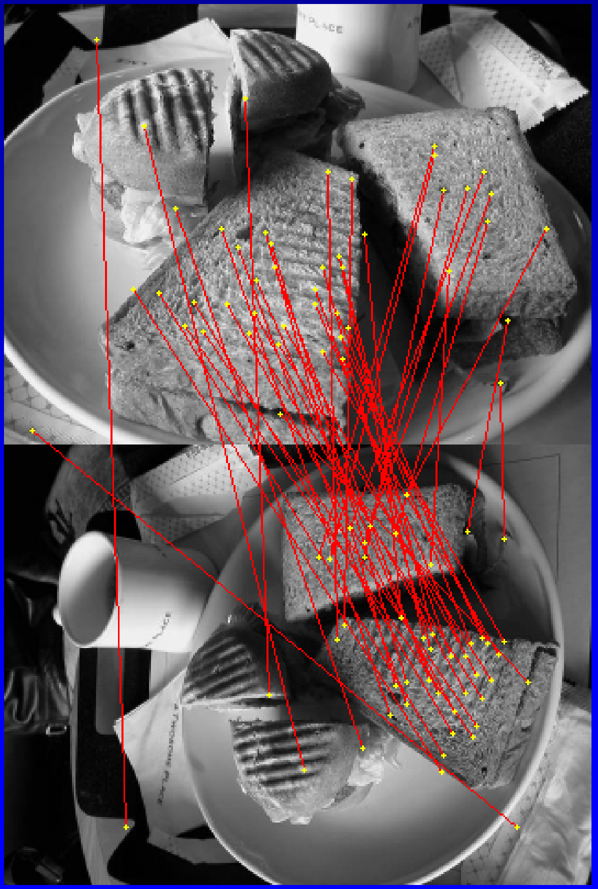
SIFT-Overview (I)

- **References**
  - D. Lowe, "Distinctive image features from scale-invariant keypoints", *IJCV*, vol. 60, no. 2, pp. 91-110, 2004.
- **Scale Invariant Feature Transform (SIFT)**
  - Invariant to scales, rotations, perspective distortion, illumination change
- **Additional advantages using SIFT**
  - A few features are enough for pattern matching
    - Occluded patterns are also recognized.
  - Learning method is usually simple
    - Nearest neighbor matching
  - Real time implementation

SIFT-Overview (II)

- **SIFT examples**



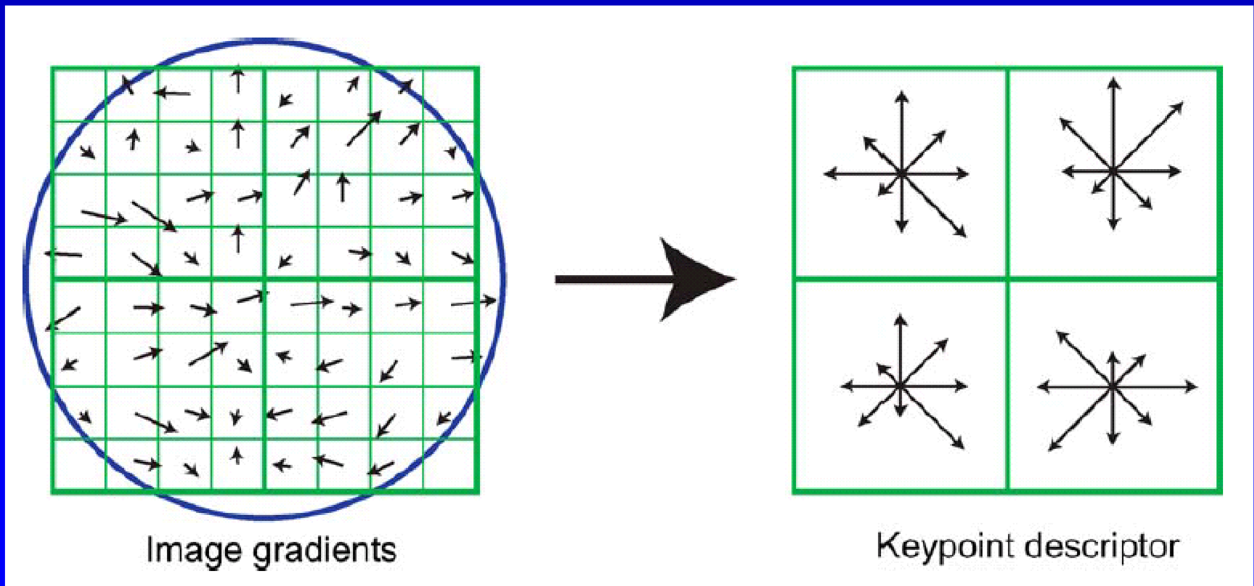


SIFT-Overview (III)

- **Procedure of SIFT**
  - Scale-space extreme point detection
    - Difference of Gaussian (DOG)
  - Keypoint localization
    - Ratio of principal curvature
  - Orientation assignment
    - Direction of local gradient at a keypoint
  - Keypoint descriptor
    - Orientation representation in a region around the keypoint

SIFT-Overview (IV)

- **Keypoint descriptor (4x4x8)**
  - Orientation histogram in a 16x16 block
  - 128 dimensional vector: SIFT feature



SIFT feature vector in an 8x8 block

SIFT: Scale-Space Extrema Detection (I)

- **Identification of candidate locations for keypoints**
- **Gaussian convolution with variable variances**
  - Pyramid structure
  - Scale space with variable σ

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

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

- **Difference of Gaussian (DOG)**
  - Similar properties of Laplacian of Gaussian (LOG)
    - LOG produced the most stable scale-invariant features

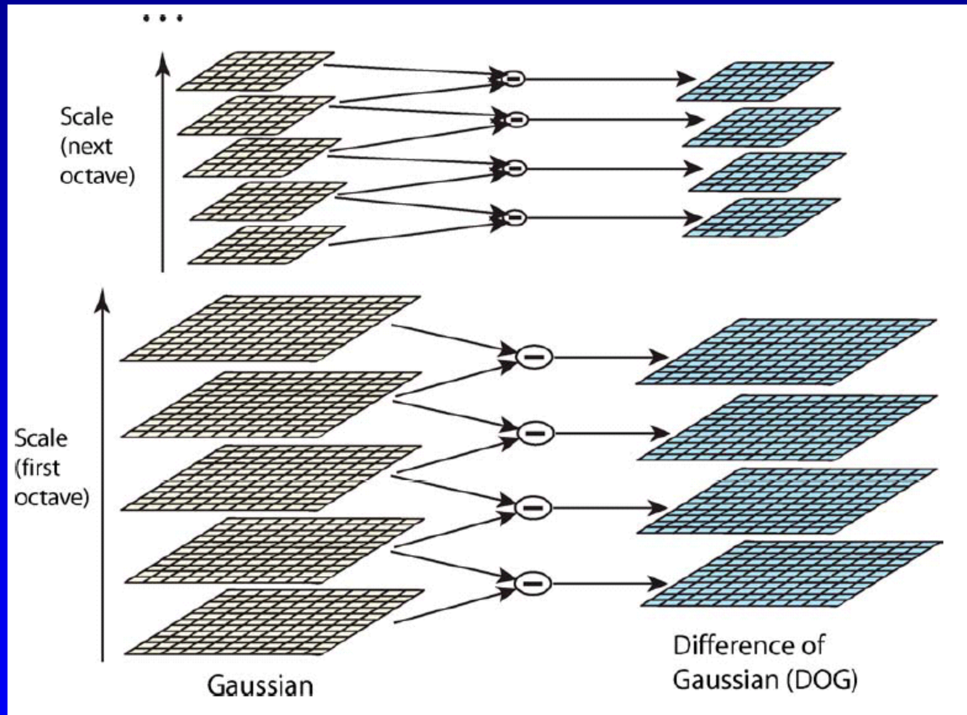
$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$
$$= L(x, y, k\sigma) - L(x, y, \sigma). \quad (1$$

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2\nabla^2 G$$

SIFT: Scale-Space Extrema Detection (II)

DOG Pyramid

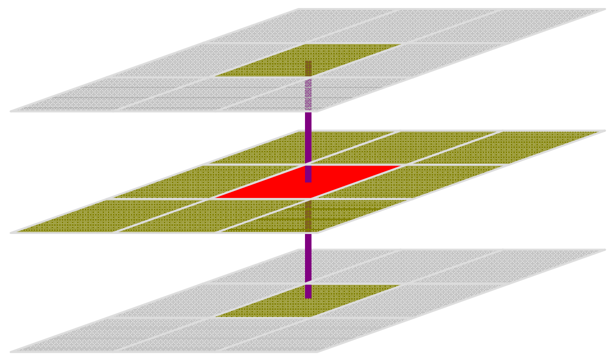
- Incremental convolution with varying variance, $k\sigma$
  - Scale-space function
  - $\mathbf{x}$ (x,y,σ)
- 2:1 subsampling
- Experimentally decided
  - k: $\sqrt{2}$
  - 3 scales per octave



SIFT: Scale-Space Extrema Detection (III)

Local extrema detection of DOG

- Decide when DOG at the location is maximum or minimum compared with 26 neighbors in the layered DOGs



26 neighbors

SIFT: Accurate Keypoint Localization (I)

- **Rejection of extrema locations**
  - Noisy
  - Along an edge
- **Interpolation of extrema location**
  - Accurate location of extreme
  - Quadratic function with Taylor expansion

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

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1} \frac{\partial D}{\partial \mathbf{x}}.$$

  - Thresholding low extreme value (T=0.03 for [0, 1])

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

SIFT: Accurate Keypoint Localization (II)

Eliminating edge response

- Keypoints along the edge in the DOG have large principal curvature ratio
- The principal curvatures are proportional to eigenvalue of Hessian matrix at the location
- Thresholding the ratio of principal curvature, r<10

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

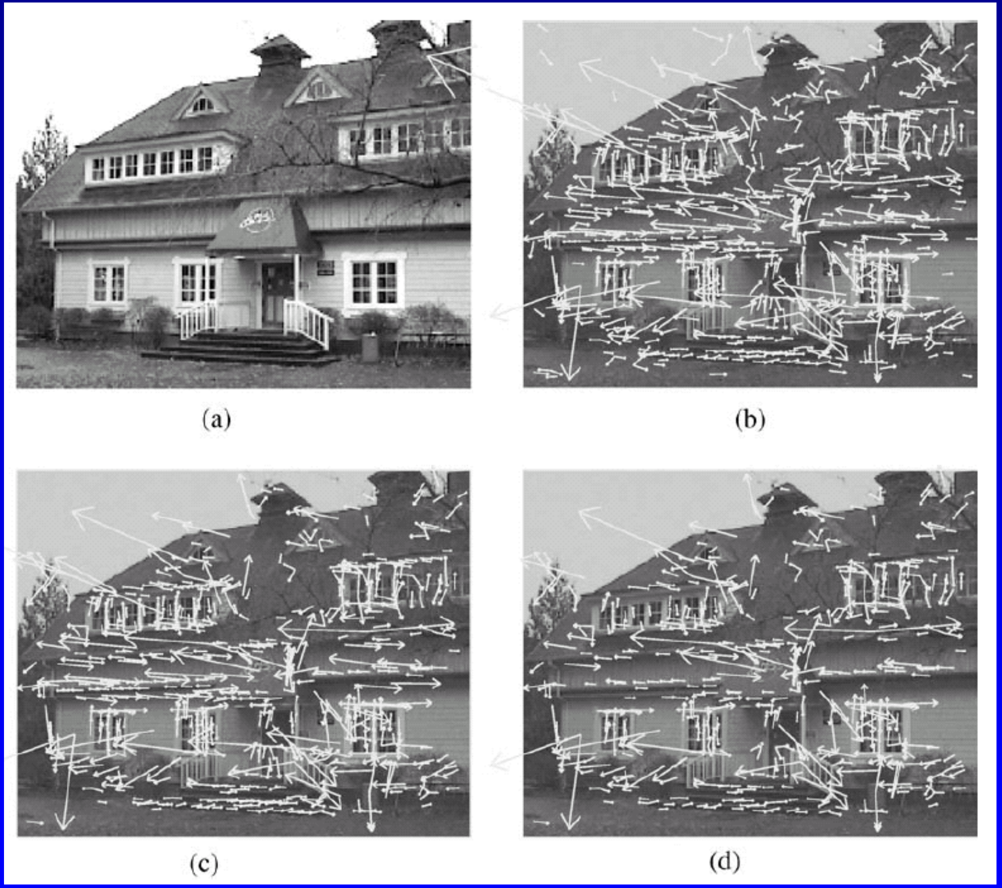
$$\mathrm{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$

$$\mathrm{Det}(\mathbf{H}) = D_{xx} D_{yy} - (D_{xy})^2 = \alpha\beta.$$

$$\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r+1)^2}{r}.$$

SIFT: Accurate Keypoint Localization (III)

- **Results**
  - (a) original image
  - (b) extreme points (832)
  - (c) low extreme threshold (729)
  - (d) large ratio of principal curvature threshold (536)



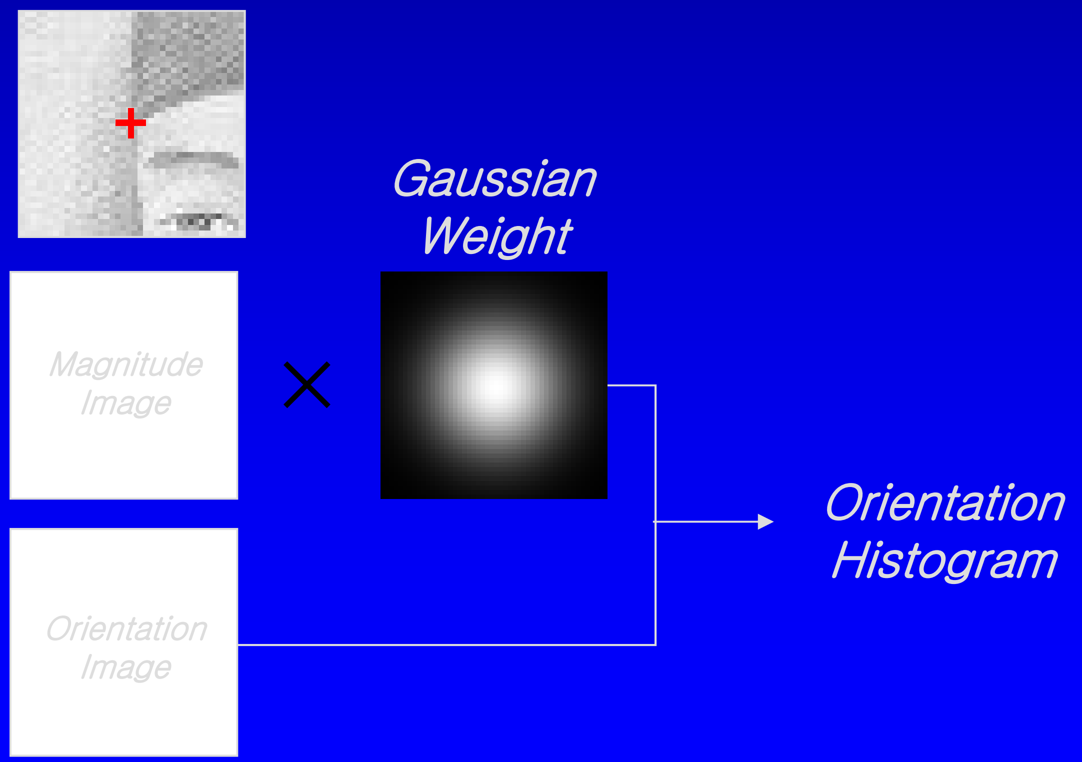
SIFT: Orientation Assignment (I)

- **Dominant direction of the keypoint**
  - Histogram of orientation in a region around the keypoint

$$dx = L_{x+1,y} - L_{x-1,y}$$
$$dy = L_{x,y+1} - L_{x,y-1}$$

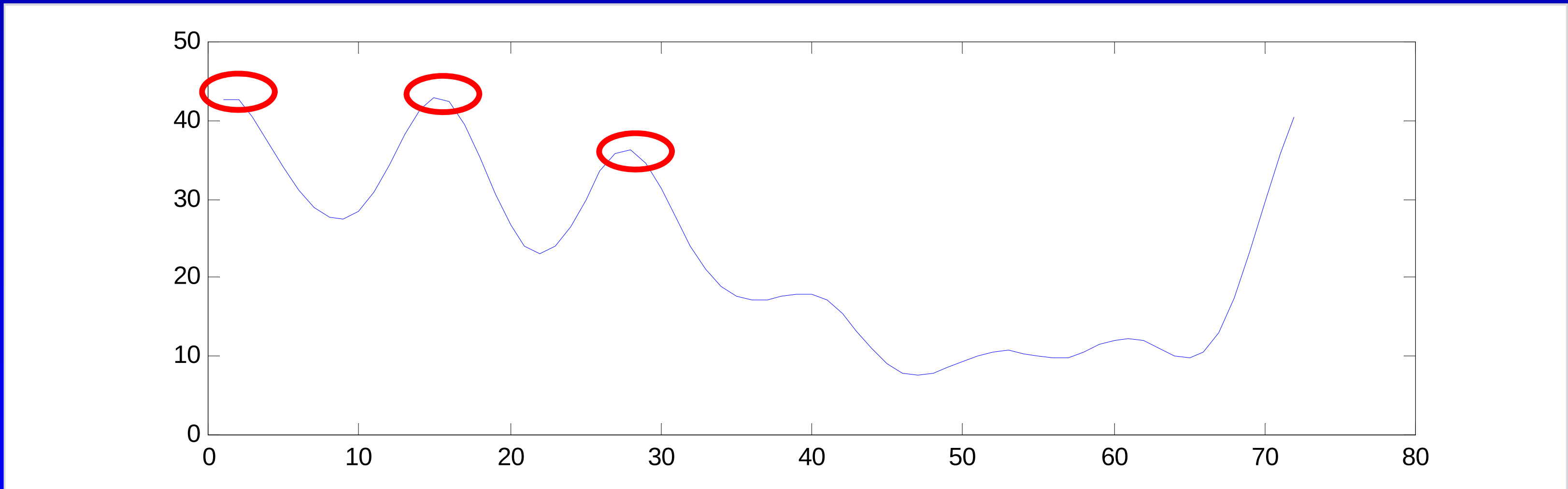
$$m = \sqrt{dx^2 + dy^2}$$

$$\theta = \tan^{-1}(dy / dx)$$



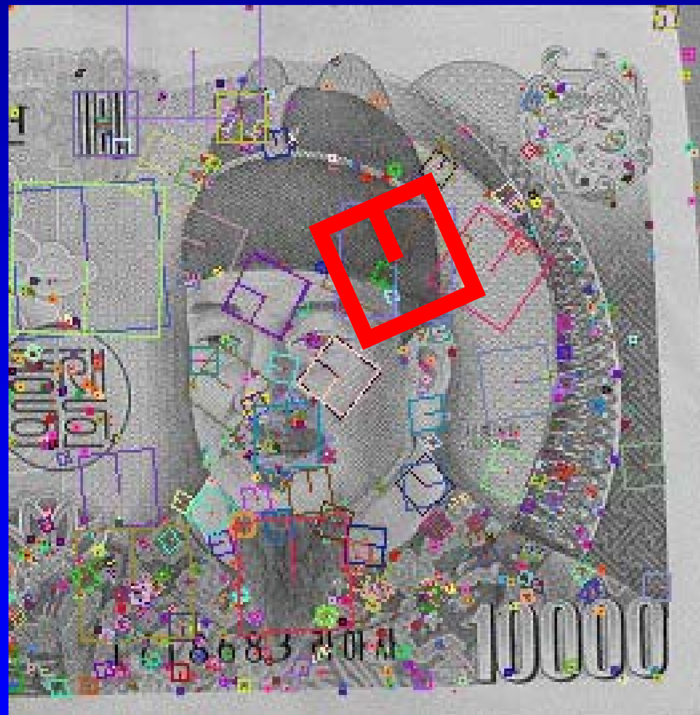
SIFT: Orientation Assignment (II)

- **3 peaks in histogram**
  - Quadratic function fitting for accurate orientation



SIFT: Orientation Assignment (III)

- Experiment

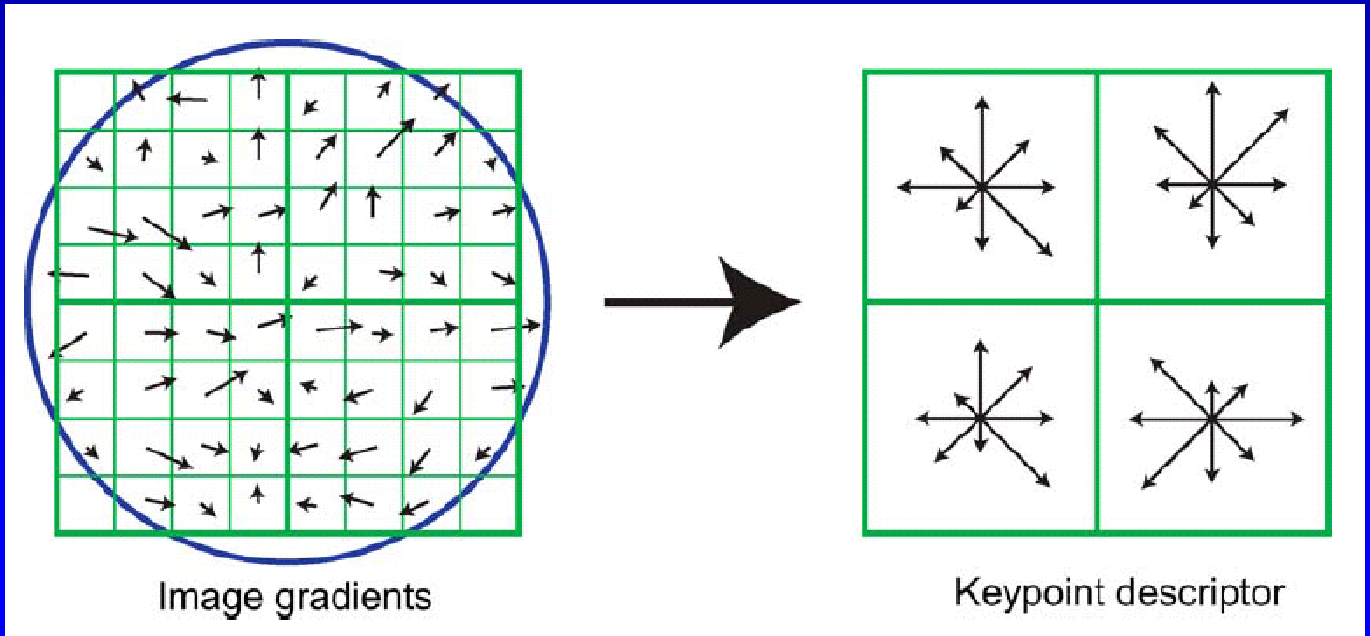


SIFT: Local Image Descriptor (I)

- **Local region description around keypoint location**
  - Real data structure to be stored and matched
  - keypoint location and scale
  - Local orientation histogram
- **Normalized orientation histogram**
  - Gradient pattern is distinctive in HVS
  - Invariant to illumination change
  - 8 directions (histogram bins) are optimal
- **Sensitivity of affine transformation**
  - 3D rotation up to 50° for reliable matching

SIFT: Local Image Descriptor (II)

- **Keypoint descriptor (2x2x8)**
  - 8x8 sample array



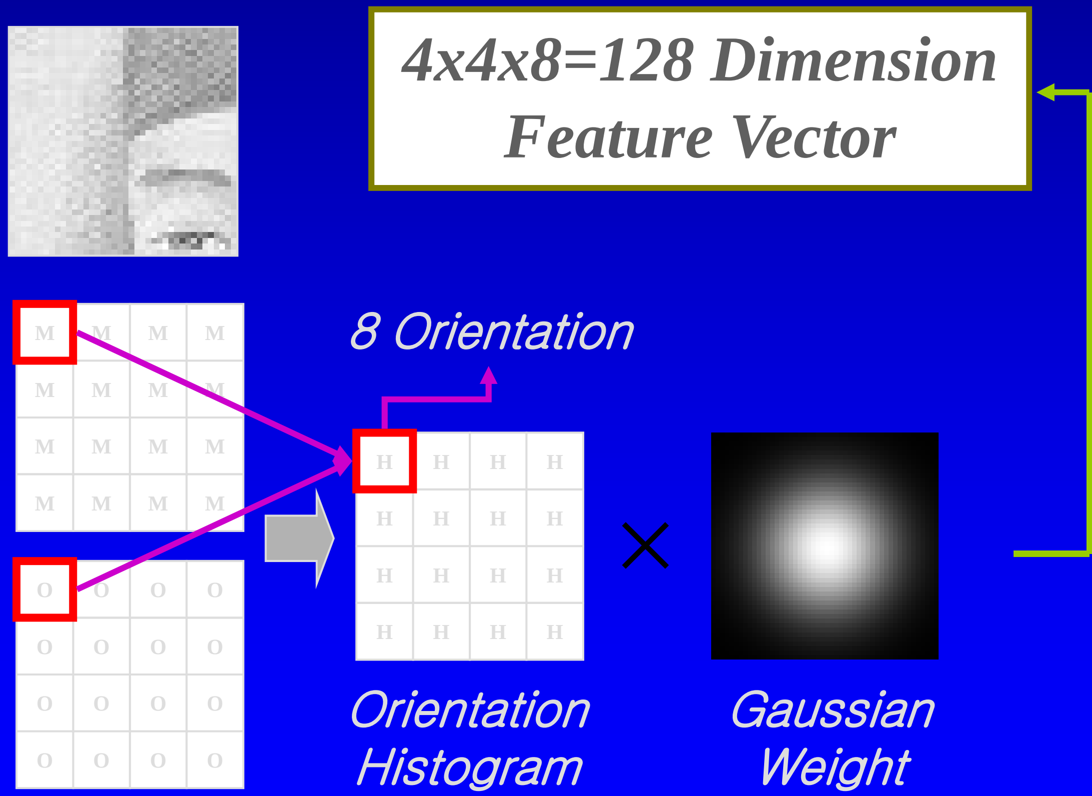
SIFT: Local Image Descriptor (III)

❑ Histogram of orientation

$$dx = L_{x+1,y} - L_{x-1,y}$$
$$dy = L_{x,y+1} - L_{x,y-1}$$

$$m = \sqrt{dx^2 + dy^2}$$

$$\theta = \tan^{-1}(dy / dx)$$

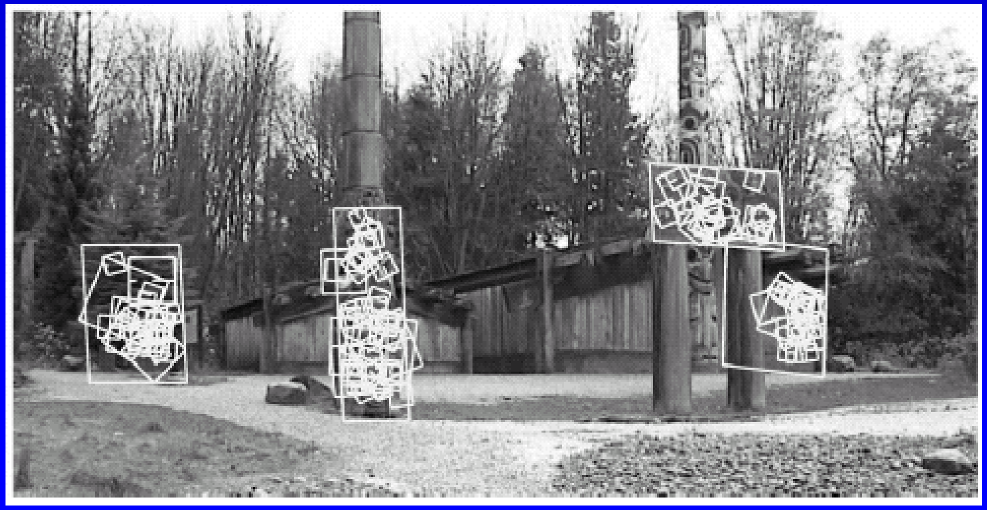


SIFT-Experiments (I)

- **Object recognition**
  - Occluded (partially visible) objects recognition
  - Under varying illuminations

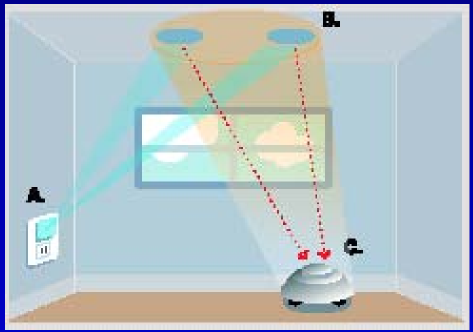
SIFT-Experiments (II)

- **Place identification**
  - Indoor/ourdoor
  - SLAMB
    - Simultaneous Localization And Map Building

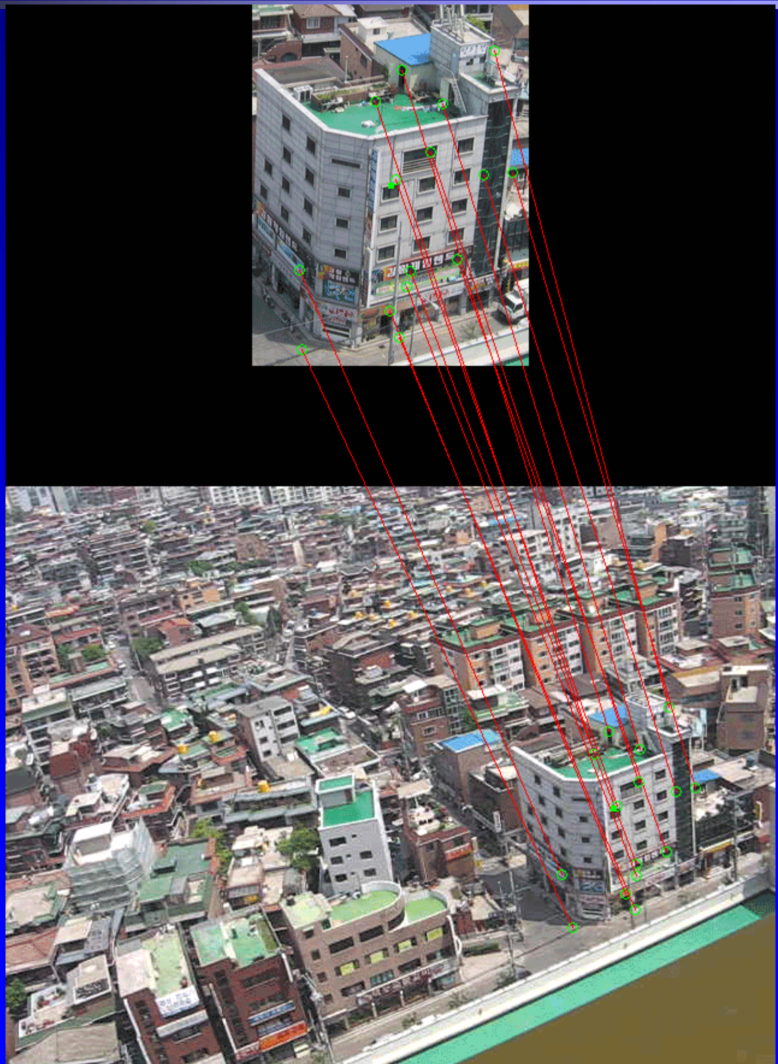


Applications Using SIFT (I)

- **Evolution Robotics**
  - Northstar
    - Self localization
  - Robotics software
- **SONY**
  - AIBO



Applications Using SIFT (II)



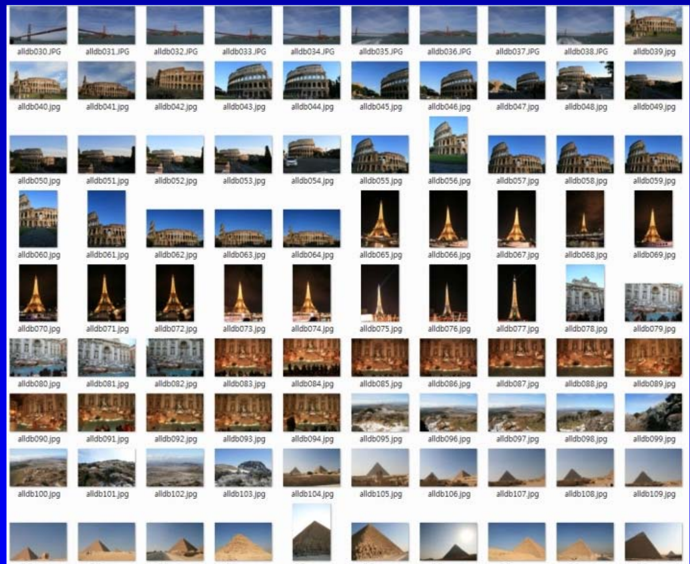
- Target detection and recognition:
  - Position, pose
  - Sensor error compensation

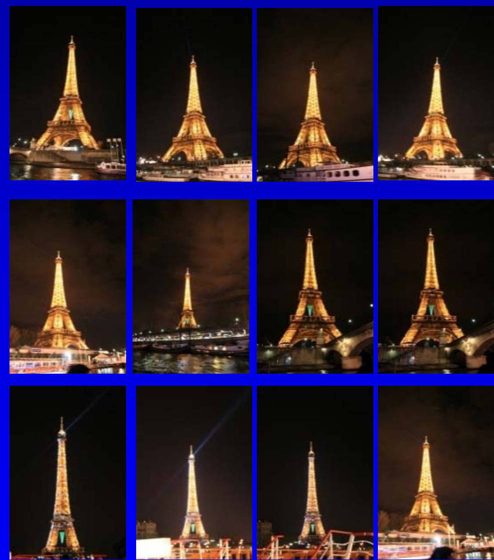
Applications Using SIFT (III)

- Building recognition in broadcasting contents

Applications Using SIFT (IV)

- **Image retrieval**





Query

Database, BOF models

Result