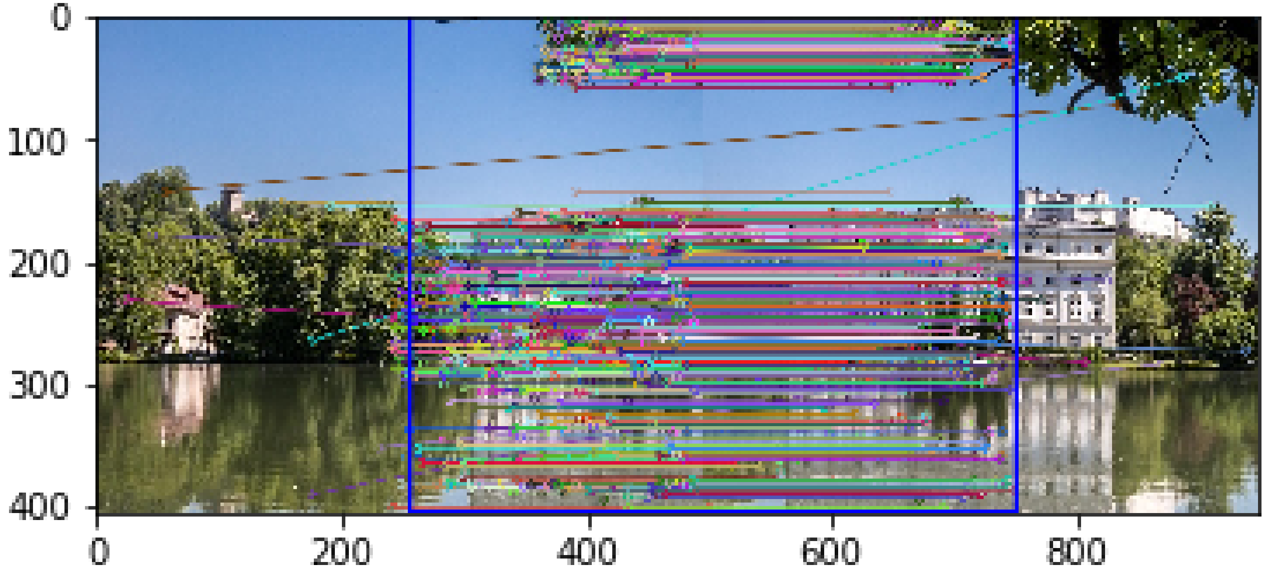
Scale Invariant Feature & Application
SIFT와 응용

Yukyung Choi

Today’s Lecture

Lecture: Local Invariant Feature
Hand on Lab: Image Stitching using OpenCV

https://colab.research.google.com/drive/1slCDHO5wK5teDO7L1be7v5qvV9kV4HY6



Agenda

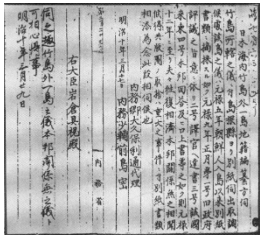
- Motivation
- SIFT & Procedure
- SIFT: Matching
- Panorama Stitching
- Reference

Motivation

In mid-april Anglesey moved his family and entourage from Rome to Naples, there to await the arrival of

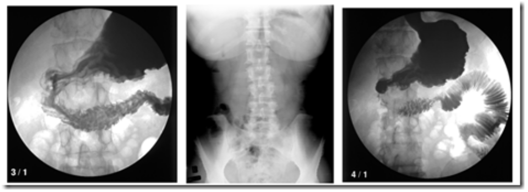
한글로 번역하라.

차량 번호를 인식하라.



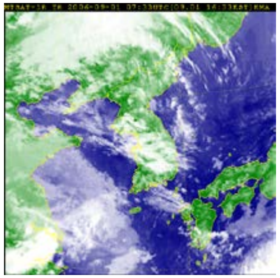
검색하라.

무슨 식물?



이상 징후를 찾아라.

뭐 하는 상점?



내일 날씨는?

몇 명인가?

BigDog

전투 로봇

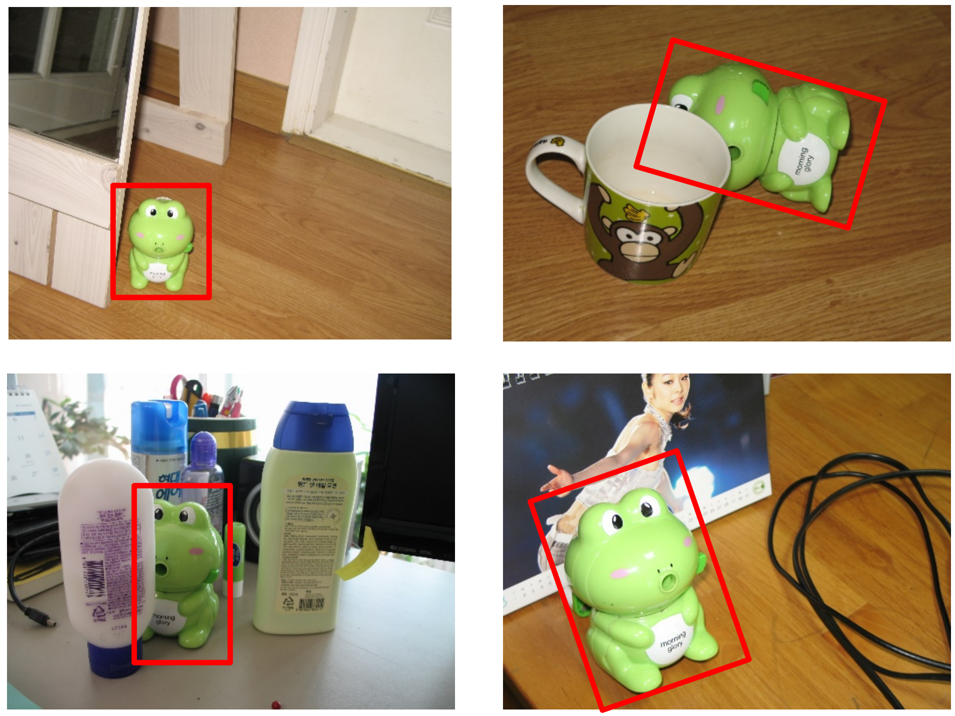
공습 목표물은 어디에?

문을 열어줄까?

웅덩이를 피하라.

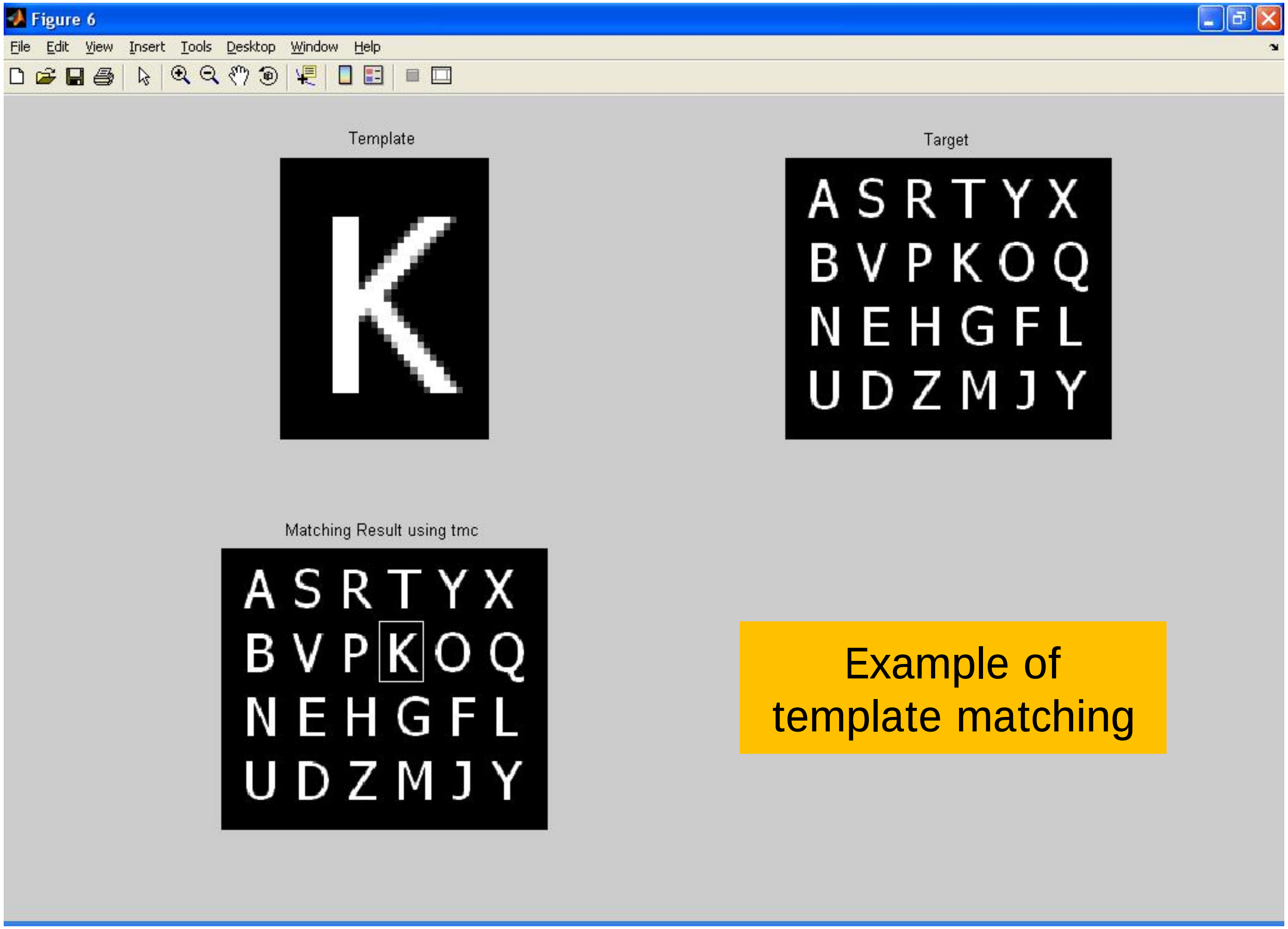
Motivation – Visual Recognition

- 찾는 물체가 영상 속에 있나? 그렇다면 어떤 포즈(pose)인가?
- 물체 인식은 전형적인 문제 (이걸 풀면 다른 문제도 풀린다.)



가림(occlusion)이 있고 clutter된 상황에서의 물체 인식

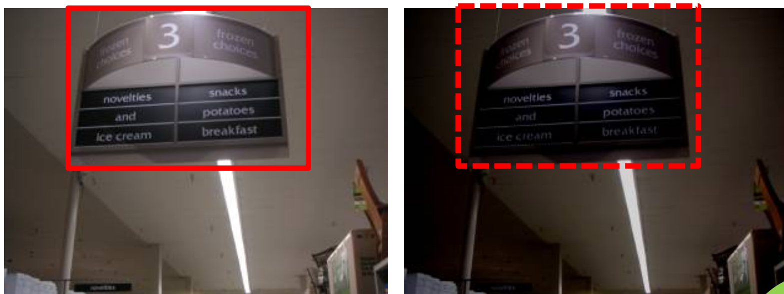
Motivation – If we can control various constraints,



Example of template matching

Motivation – Why do we need special feature?

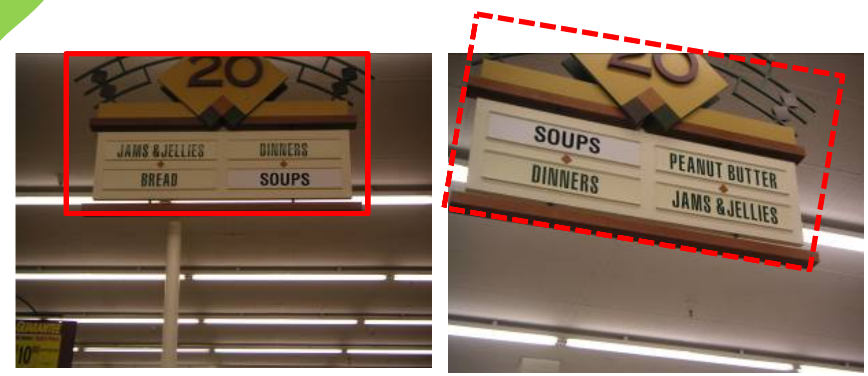
- 실제 영상에는 다양한 변화가 나타난다.



조명 (Brightness)

조명+크기 (Brightness + Scale)

조명+크기+회전
(Brightness+Scale+Rotation)



조명+크기+회전+어파인
(Brightness+Scale+Rotation+Affine)

Motivation – Local Invariant Feature

- 지역특징(local feature)의 조건
  - 불변성 (invariant)
    - 변화된 환경에서도 반복하여 나타나는 특성 (재현성 (repeatability))
      - 조명 불변
      - 크기 불변
      - 회전 불변
      - 잡음 불변

➔ 밝기 정규화, 크기 정규화, 방향 정규화 etc

  - 풍부한 정보 (rich description)
  - **강한 분별력 (distinctiveness)**
  - 계산 효율 (추출, 매칭)

Famous Invariant Features

SIFT, SURF, DAISY

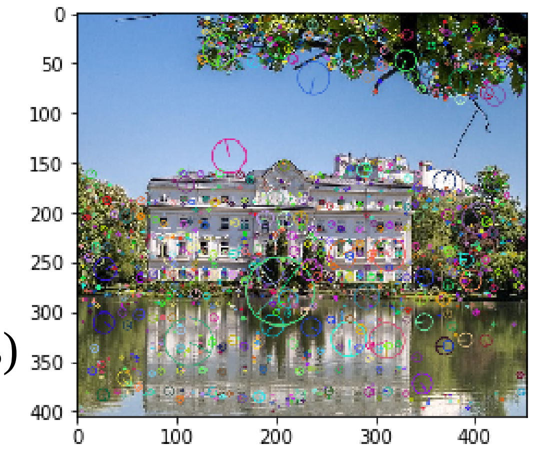
BRIEF, ORB, BRISK

...

http://multispectral.sejong.ac.kr/ykchoi/RSS2016/Features.pdf

SIFT ?

- SIFT의 창안 (ICCV➔IJCV)
  - Lowe, David G. (1999). "Object recognition from local scale-invariant features". *Proceedings of the International Conference on Computer Vision*. pp. 1150–1157.
  - 불변성이 강한 지역 특징 (local feature)



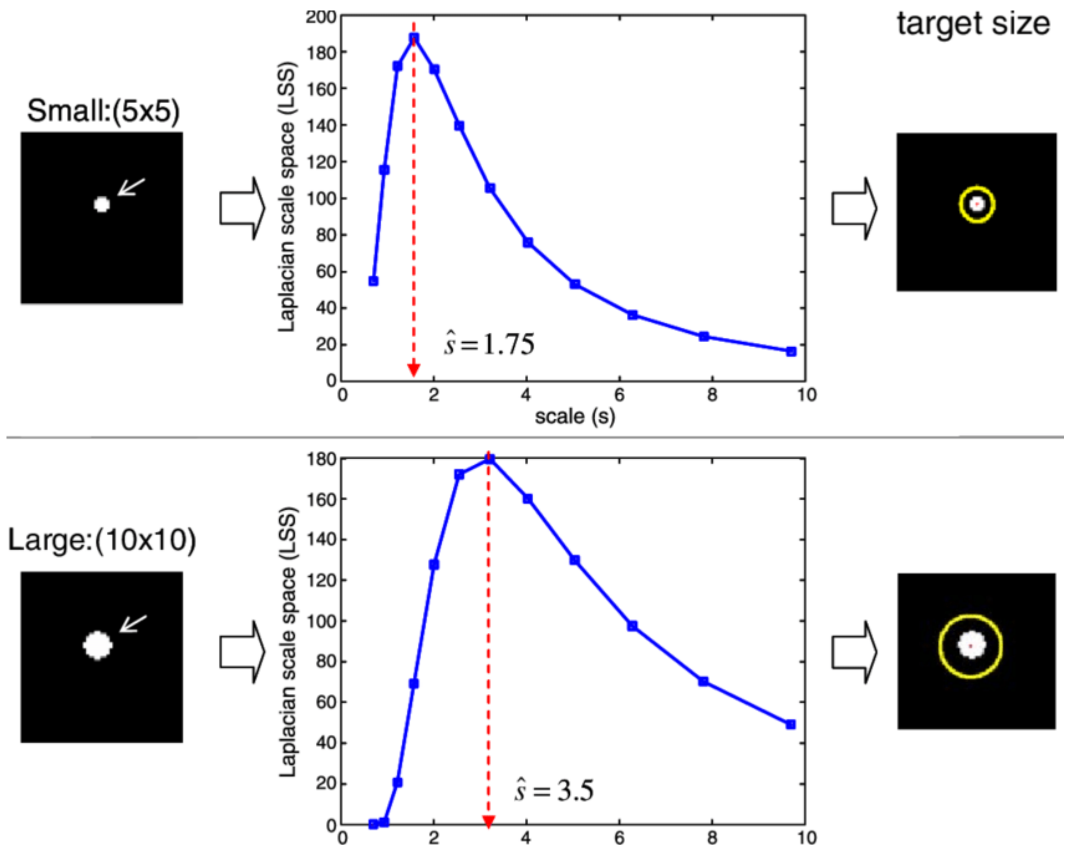
- 등장 배경
  - Scale-space theory [Lindeberg94]
  - ‘deep structure’ (feature structure between different scales)

SIFT Procedure

- 추출 과정
  1. 키포인트 탐지 (keypoints detection)➔ (x,y,scale)
     ① Scale-space에서 극점 탐지
     ② 잡음 극점 제거
  2. 기술자 추출 (descriptor extraction) ➔ (theta) ➔ 1d Vector
     ① 방향 결정
     ② 그라디언트 히스토그램 계산

SIFT : Keypoints Detection

- 고유스케일: **Lindeberg's Laplacian Operator (Scale Theory)**

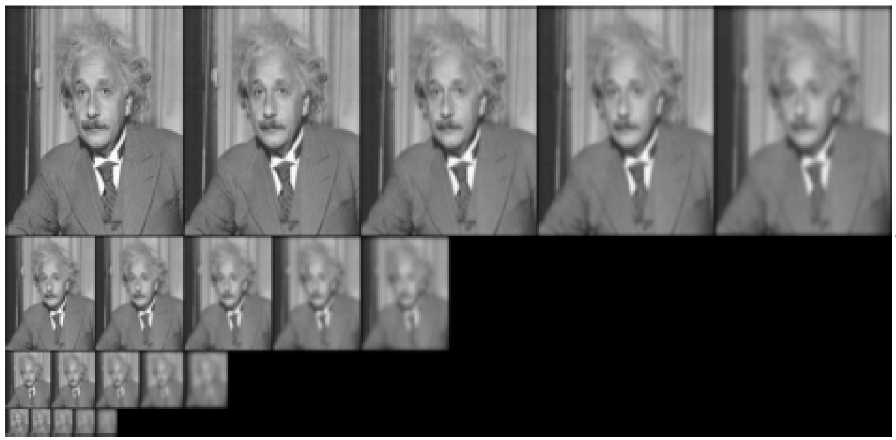


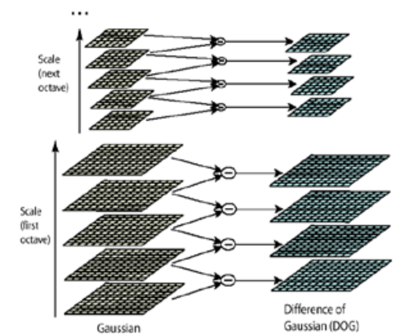
SIFT : Keypoints Detection

- SIFT는 “scale space”에서 작업함으로써 크기 불변성 달성

갈수록 큰 σ의 Gaussian으로 컨볼루션

Resampling으로
pyramid 구성





DOG
(Difference of
Gaussian)

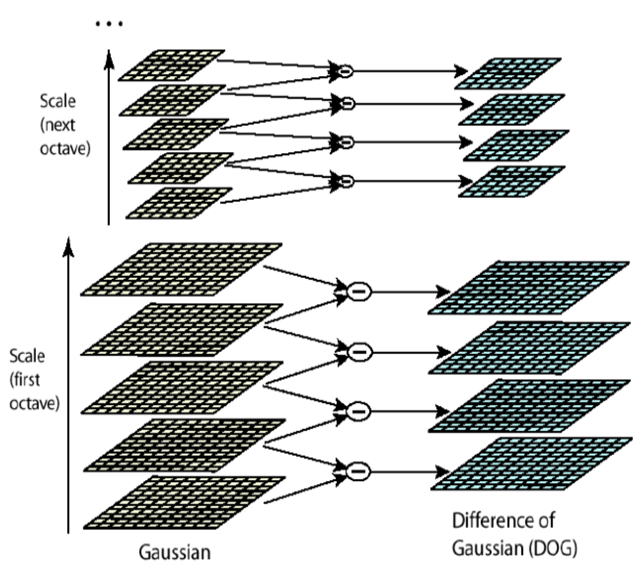
SIFT : Keypoints Detection

- 수식으로 써 보면,
  - 스케일 σ인 가우션으로 컨볼루션된 영상 *L*

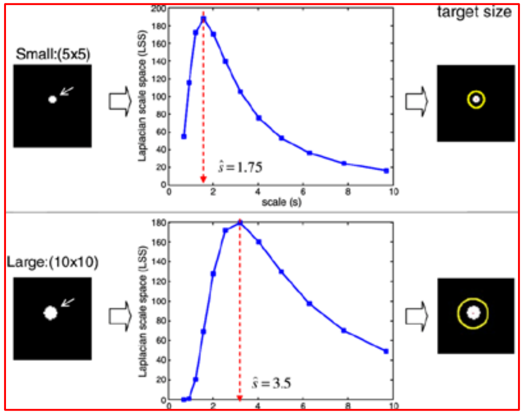
$$L(x,y,\sigma) = G(x,y,\sigma) * f(x,y) \text{ where } G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

  - DOG 영상 ➔ **Difference of Gaussian (DoG)**

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * f(x,y) = L(x,y,k\sigma) - L(x,y,\sigma)$$

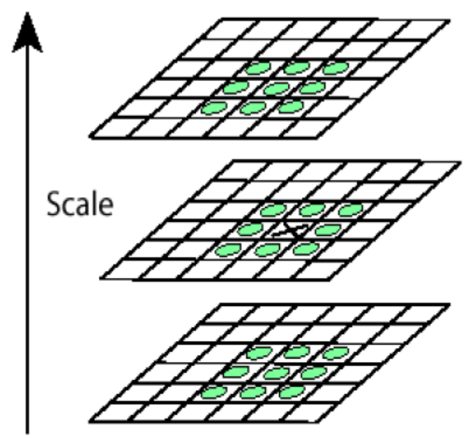


Scale Theory와 DoG의 관계

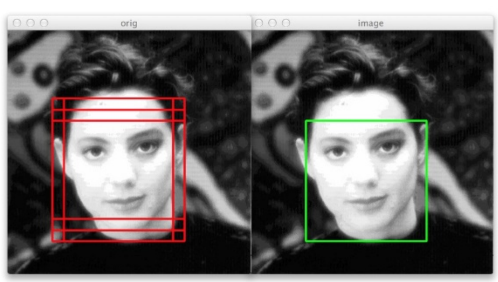


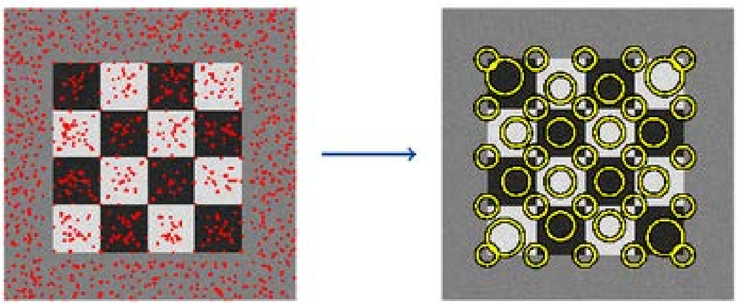
SIFT : Keypoints Detection

- 키포인트 탐지
  - 모든 DOG의 모든 점에 대해 **극점 (maxima 또는 minima) 조건 검사**
    - ➔ **Non-Maxima Suppression (NMS)**
  - 통과한 점은 후보 키포인트가 됨



인근 8개와 위 아래 영상의
18개 (총 26개) 점과 비교

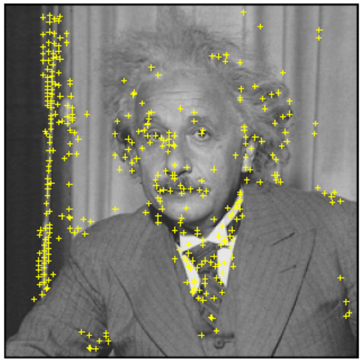




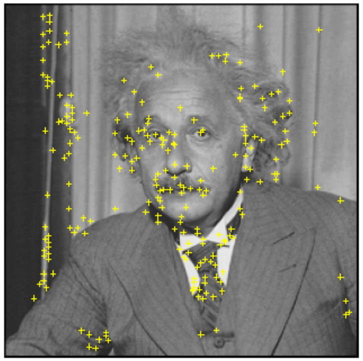
SIFT : Keypoints Detection

- 잡음 극점 제거 for **Robustness**
  - 낮은 대비(contrast) 극점 제거
  - 에지에 놓인 극점 제거

a)

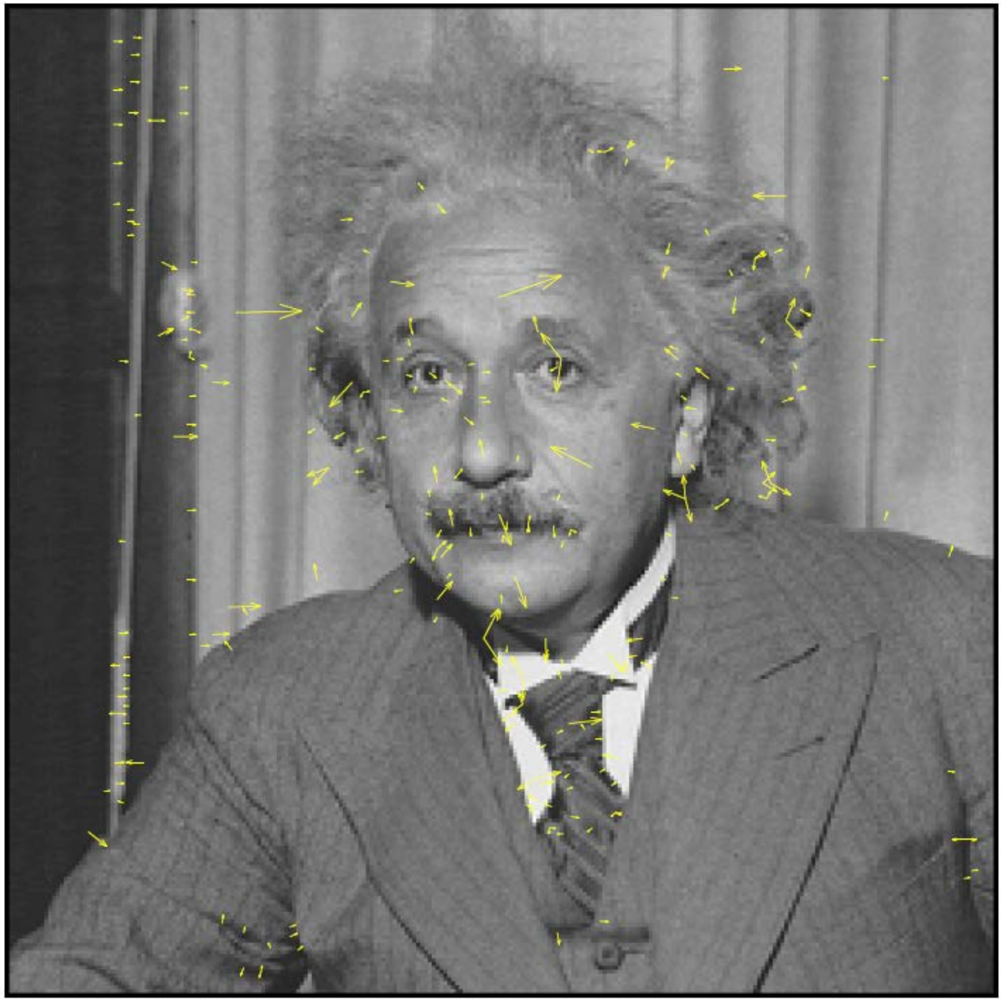


b)



(a) 극점
(b) 낮은 대비 극점 제거
(c) 에지 극점 제거

SIFT : Descriptor



회전 불변성을 위해서는
노란 화살표와 같이
각 Local Feature의 **“고유 방향”**이
필요하다.

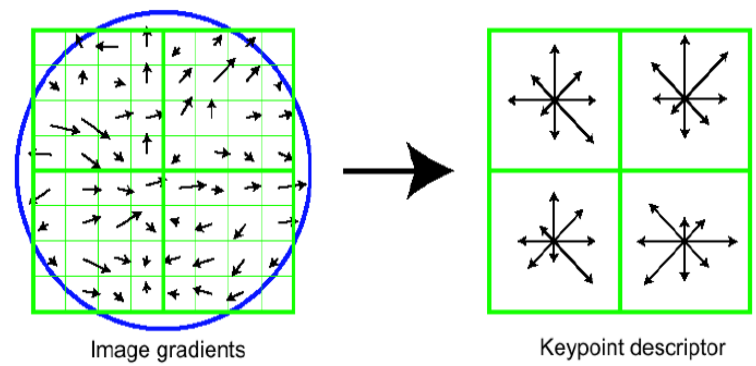
어떻게 구하는가?

SIFT : Descriptor

- <u>풍부한 정보력 & 뛰어난 구별성:</u> Histogram of Gradient ➔ HOG
  - 키포인트 주위에서 4*4 크기의 영역을 n^2개 (n* n) 만듦
    - (아래 그림에서 n=2)
  - 각 영역에 대해 8-방향 계산하고 (이때 키포인트 방향을 기준으로 계산하여, 방향 불변성 달성) 히스토그램 구함

오른쪽과 같이 n=2인 경우, 4 영역 각각이 8 방향의 히스토그램을 가지므로, 기술자는 32개의 특징을 갖는 특징벡터로 표현된다.

n=4이면, 4*4*8=128개의 특징.

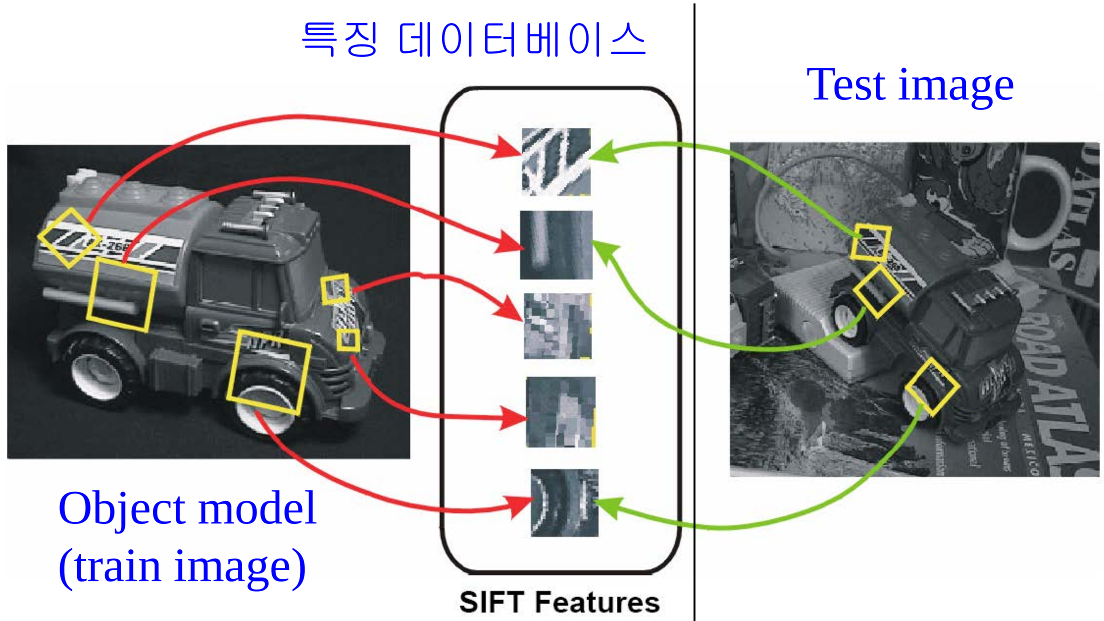


SIFT : Descriptor

- SIFT를 통해 제공되는 키포인트 기술자 정보
  - 위치, 스케일, 방향 ➔ (x,y,scale,theta)
  - 그라디언트 히스토그램 (아래 경우 4*4*8 특징벡터) ➔ 128dim 1d vector

SIFT : Matching

- 특징 매칭은 물체 인식의 핵심 연산
  - 물체 모델 영상 (training image)에서 SIFT특징을 추출하여 데이터베이스에 저장 (오프라인 과정)
  - 테스트 영상이 입력되면, 거기에서 SIFT특징을 추출하고, 데이터베이스의 특징과 매칭 (온라인 과정)

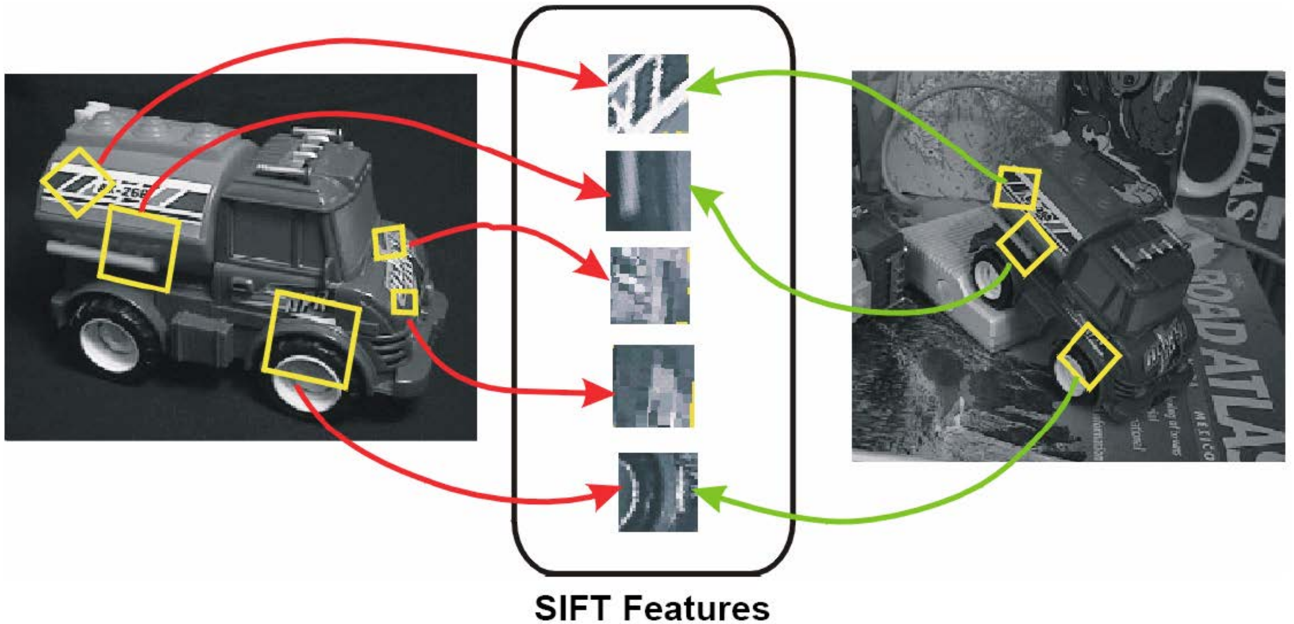


오프라인으로 이루어짐 | 온라인으로 이루어짐

SIFT : Matching

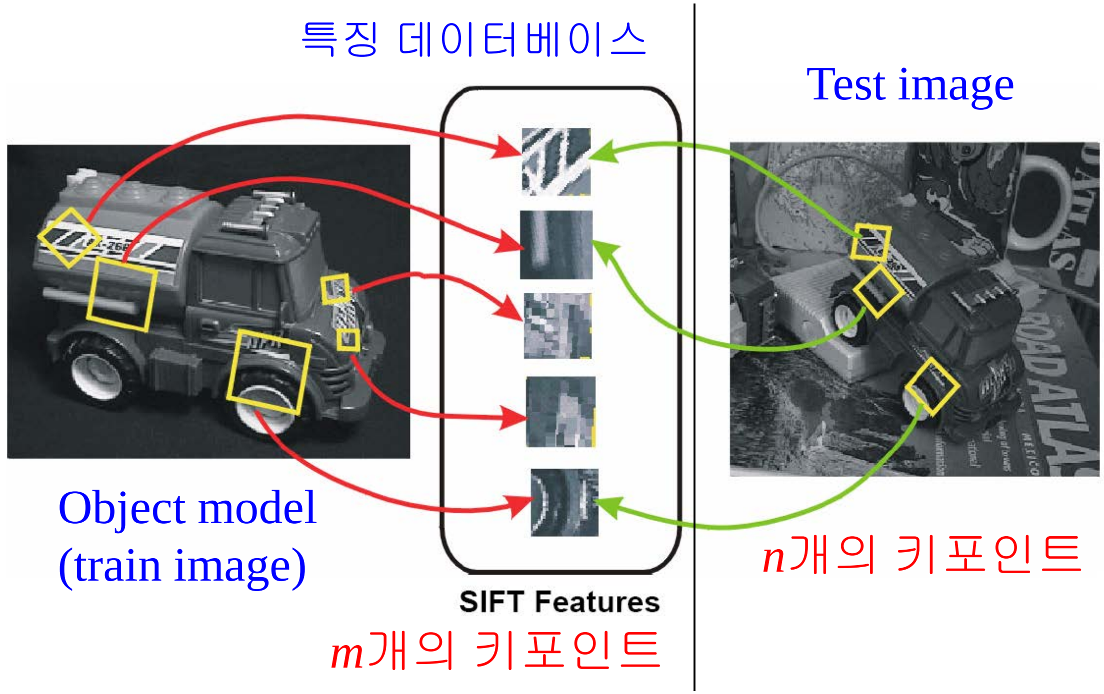
- 매칭 과정
  1. **최근접 이웃 탐색 (nearest-neighbor search) : KD-Tree**
  2. Hough 변환을 이용하여, 매칭된 키포인트 클러스터 찾기
  3. Linear least squares로 검증
  4. Outlier 처리

RANSAC



SIFT : Matching (KD-Tree)

- 키포인트 매칭을 어떻게 할 것인가? – 계산 효율이 핵심
  - m과 n은 보통 수천~수만 개
  - n개 각각에 대해, m개 중에서 최근접 이웃을 탐색해야함
  - **$O(mn)$ 계산 복잡도**



Comparison

- **Related Works**
  - SURF, RIFT, PCA-SIFT, ASIFT, GLOH 등
- **Comparison Paper [Mikolajczyk2005]**
  - "… we observe that the ranking of the descriptors is mostly independent of the interest region detector and that the SIFT-based descriptors perform best. …"

Famous Invariant Features

SIFT, SURF, DAISY

BRIEF, ORB, BRISK

…

http://multispectral.sejong.ac.kr/ykchoi/RSS2016/Features.pdf

Panorama Stitching

(a) Image 1

(b) Image 2

(c) SIFT matches 1

(d) SIFT matches 2

(e) Images aligned according to a homography

Algorithm: Panoramic Recognition

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbours for each feature using a k-d tree

III. For each image:
   (i) Select m candidate matching images (with the maximum number of feature matches to this image)
   (ii) Find geometrically conisistent feature matches using RANSAC to solve for the homography between pairs of images
   (iii) Verify image matches using probabilistic model

IV. Find connected components of image matches

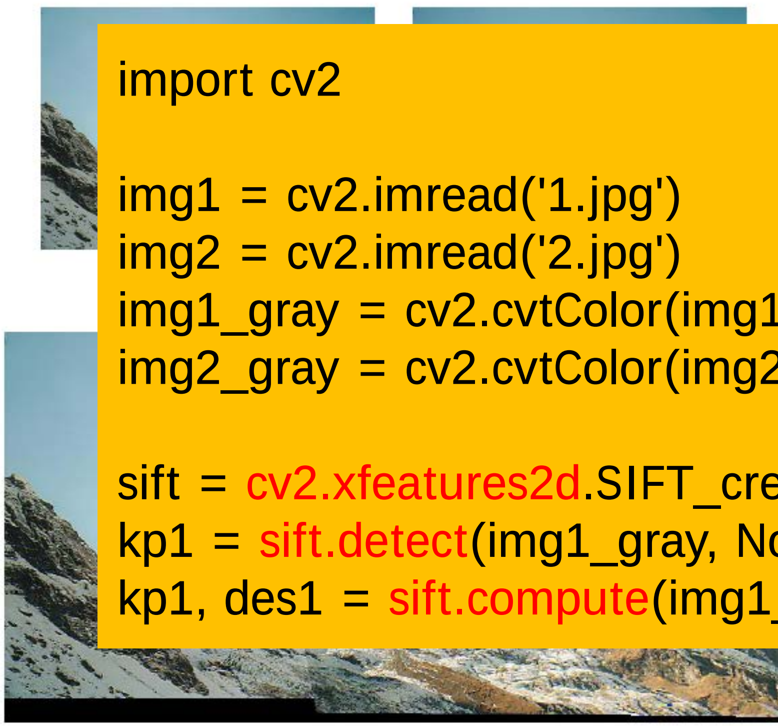
V. For each connected component:
   (i) Perform bundle adjustment to solve for the rotation $\theta_1, \theta_2, \theta_3$ and focal length f of all cameras
   (ii) Render panorama using multi-band blending

Output: Panoramic image(s)

Panorama Stitching

(a) Image 1

(b) Image 2



(e) Images aligned according to a homography

Algorithm: Panoramic Recognition

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbours for each feature using a k-d tree

h the
s im-

tches
y be-

odel

rota-
eras

(ii) Render panorama using multi-band blending

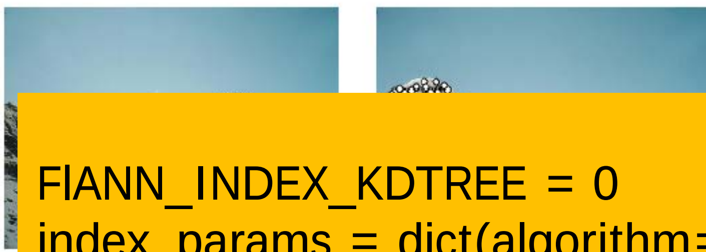
Output: Panoramic image(s)

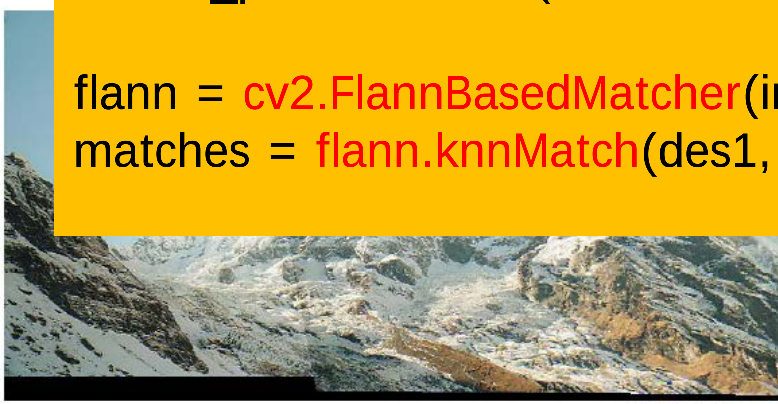
```
import cv2

img1 = cv2.imread('1.jpg')
img2 = cv2.imread('2.jpg')
img1_gray = cv2.cvtColor(img1, cv2.COLOR_BGR2GRAY)
img2_gray = cv2.cvtColor(img2, cv2.COLOR_BGR2GRAY)

sift = cv2.xfeatures2d.SIFT_create()
kp1 = sift.detect(img1_gray, None)
kp1, des1 = sift.compute(img1_gray, kp1)
```

Panorama Stitching

(a) Image 1

(b) Image 2





(e) Images aligned according to a homography

Algorithm: Panoramic Recognition

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbours for each feature using a k-d tree

III. For each image:

(i) Select m candidate matching images (with the

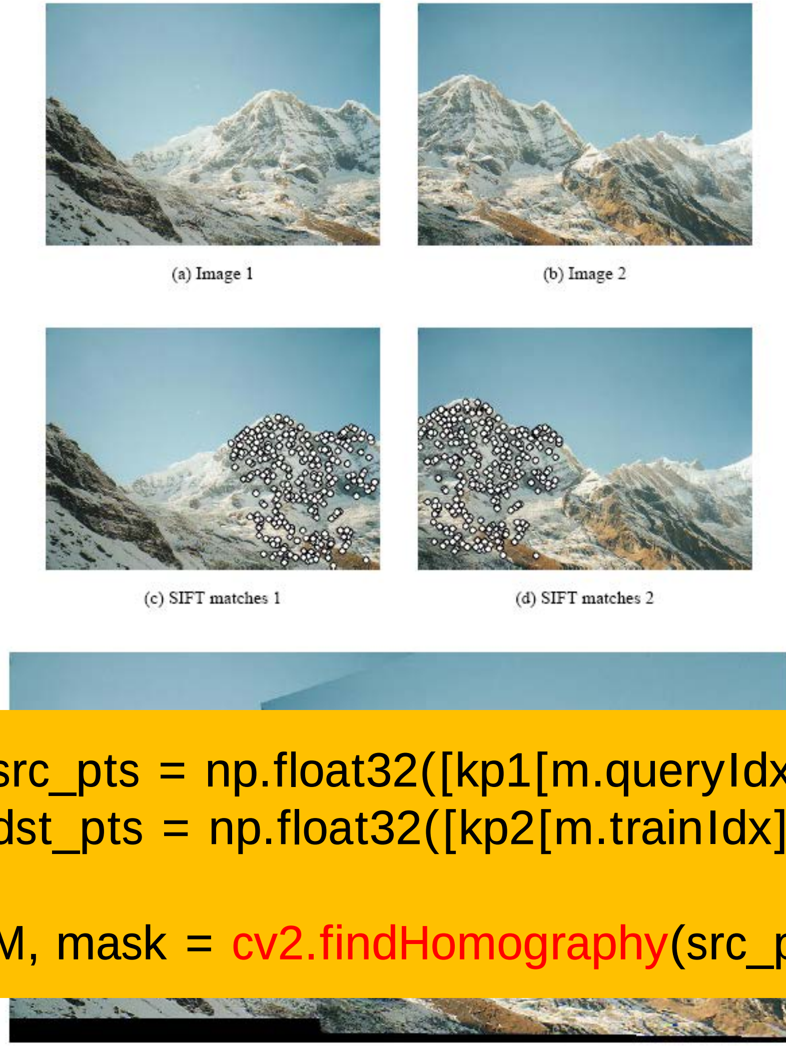
(i) Perform bundle adjustment to solve for the rotation $\theta_1, \theta_2, \theta_3$ and focal length f of all cameras

(ii) Render panorama using multi-band blending

Output: Panoramic image(s)

```
FlANN_INDEX_KDTREE = 0
index_params = dict(algorithm=FlANN_INDEX_KDTREE, trees=5)
search_params = dict(checks=50)

flann = cv2.FlannBasedMatcher(index_params, search_params)
matches = flann.knnMatch(des1, des2, k=2)
```

Panorama Stitching



(a) Image 1 (b) Image 2

(c) SIFT matches 1 (d) SIFT matches 2

(e) Images aligned according to a homography

Algorithm: Panoramic Recognition

Input: n unordered images

I. Extract SIFT features from all n images

II. Find k nearest-neighbours for each feature using a k-d tree

III. For each image:
   (i) Select m candidate matching images (with the maximum number of feature matches to this image)
   (ii) Find geometrically conisistent feature matches using RANSAC to solve for the homography between pairs of images
   (iii) Verify image matches using probabilistic model

Output: Panoramic image(s)

```
src_pts = np.float32([kp1[m.queryIdx].pt for m in good]).reshape(-1, 1, 2)
dst_pts = np.float32([kp2[m.trainIdx].pt for m in good]).reshape(-1, 1, 2)

M, mask = cv2.findHomography(src_pts, dst_pts, cv2.RANSAC, 5.0)
```

Panorama Stitching

… We have used local invariant features to allow panorama matching to be fully automated without assuming any ordering of the images or any restriction on focal lengths, orientations, or exposures. The matching time is linear in the number of images, so all panoramas can be automatically detected in large sets of images. We have also developed approaches for seamlessly blending images even when illumination changes or there are small misregistrations.

Reference

- Lowe가 SIFT를 제안
  - D.G. Lowe, "Object recognition from local scale-invariant features," *Proceedings of the International Conference on Computer Vision,* pp. 1150–1157, 1999.
  - D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110, 2004.
- Scale-space와 지역 기술자 튜토리얼
  - Tony Lindeberg “Scale-space theory: A basic tool for analysing structures at different scales,” *J. of Applied Statistics*, vol.21, No.2, pp. 224–270, 1994.
  - T. Tuytelaars and K. Mikolajczyk, “Local invariant feature detectors,” *Foundations and Trends in Computer Graphics and Vision*, Vol.3, No.3, pp177-280, 2007.
- 지역 기술자의 성능 비교
  - K. Mikolajczyk, and C. Schmid, "A performance evaluation of local descriptors", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.1615-1630, 2005.
- 응용
  - D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, D. Schmalstieg, “Real-time detection and tracking for augmented reality on mobile phones ,” *IEEE Transactions on Visualization and Computer Graphics*, 2010.
  - S. Se, D.G. Lowe, J. Little, "Vision-based mobile robot localization and mapping using scale-invariant features," *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp.2051, 2001.
  - M. Brown, D.G. Lowe, "Recognising panoramas," *Proceedings of the ninth IEEE International Conference on Computer Vision*, pp.1218–1225, 2003.
  - Iryna Gordon and D.G. Lowe, "What and where: 3D object recognition with accurate pose," in *Toward Category-Level Object Recognition*, Springer-Verlag, pp. 67-82, 2006.
  - I. Laptev and T. Lindeberg, "Local descriptors for spatio-temporal recognition". *ECCV'04 Workshop on Spatial Coherence for Visual Motion Analysis,* (*LNCS Volume 3667*), pp. 91–103, 2004.
  - Matthew Toews, William M. Wells III, D. Louis Collins, Tal Arbel, "Feature-based morphometry: discovering group-related anatomical patterns," *NeuroImage*, Vol.49, No.3, pp.2318–2327, 2010.

Recommendation

- 데이터분석과 이미지처리
  - 비디오
    - https://www.youtube.com/watch?v=V8Lpf3WCZ4g&list=PLRx0vPvlEmdBx9X5xSgcEk4CEbzEiws8C
  - 코드
    - https://github.com/ndb796/Python-Data-Analysis-and-Image-Processing-Tutorial

- 다크프로그래머
  - 영상처리
    - https://darkpgmr.tistory.com/category/%EC%98%81%EC%83%81%EC%B2%98%EB%A6%AC