

2022-1 영상처리특론(BE111_0243) 중간고사

2022-1 Special Topic in Image Processing Mid-Term Take Home Exam

20211101095 김현빈

Image Segmentation. While k-means and mixtures of Gaussians use a parametric form to model the probability density function being segmented, mean shift implicitly models this distribution using a smooth continuous non-parametric model. The key to mean shift is a technique for efficiently finding peaks in this high-dimensional data distribution without ever computing the complete function explicitly.

(Problem 1) Explain the basic idea of K-means Clustering algorithm and Mean Shift algorithm for image segmentation, respectively. Describe the main difference of both algorithms in addition to explain the weak and strong point of each algorithm. Apply both algorithms to the following picture (a satellite image of Suwon area) and compare the segmentation results. (attach the segmentation results)

(Answer)

A. K-means vs Mean Shift

K-means 알고리즘은 다음과 같은 순서로 진행됩니다.

- 1) k 개의 클러스터 $\mu_i (\forall i, 1 \leq i \leq k)$ 들을 초기화합니다.
- 2) 픽셀 데이터를 각 클러스터 중심 μ_i 들 중 해당 데이터와 가장 가까운 클러스터 S_i 에 배정한다.
- 3) 각 클러스터의 데이터 무게 중심값으로 기존 μ_i 를 대체한다.
- 4) 과정 2~3을 반복한다.
- 5) 클러스터의 데이터 구성이 변하지 않으면 종료한다.

클러스터 중심 초기화 기법에는 대표적으로 random partition, forgy, macQueen, Kaufman 방법이 있습니다. 거리함수로는 주로 유클리드 거리를 활용합니다.

K-means 알고리즘은 간단하고, 빠르게 수행됩니다. 그러나 클러스터 중심 초기화 방법 및 결과에 따라 클러스터링 결과가 크게 변합니다. 또한 일반적인 픽셀값 분포에서 벗어나는 outlier가 존재하는 경우에 매우 취약합니다.

Mean-Shift 알고리즘은 데이터 분포의 무게중심을 찾고, 가장 밀집한 구간들을 찾아 군 집화하는 방법입니다.

- 1) Sliding window의 크기(kernel size, 유클리드 거리 기반 반지름)를 정의한다.
- 2) 클러스터의 중심이 되는 data point를 설정한다.
- 3) window마다 데이터 분포가 높은 방향으로 centroid를 이동한다.
- 4) 단, window끼리 중첩되는 경우 데이터 밀도가 적은 window를 제거한다.
- 5) 클러스터의 데이터 구성이 변하지 않으면 종료한다.

Mean-Shift 알고리즘은 클러스터의 개수가 자동적으로 설정된다는 장점이 있습니다. 반면 K-means에 비해 상당히 느립니다.

B. Result of K-means

다음과 같은 환경에서 구현하였습니다.

- Python (language)
- Numpy (대수연산 병렬화)
- OpenCV (영상 입출력)

구현 세부 사항은 다음과 같습니다.

- RGB to HSV 색상 변환
- Centroid 초기화 방법 forgy 및 random
- Centroid 및 Cluster 갱신
- 반복 종료 플래그로 Cluster 변화량 대신 Centroid 변화량 측정, graph로 출력
- 영상 전처리로 median 필터 적용(선택)

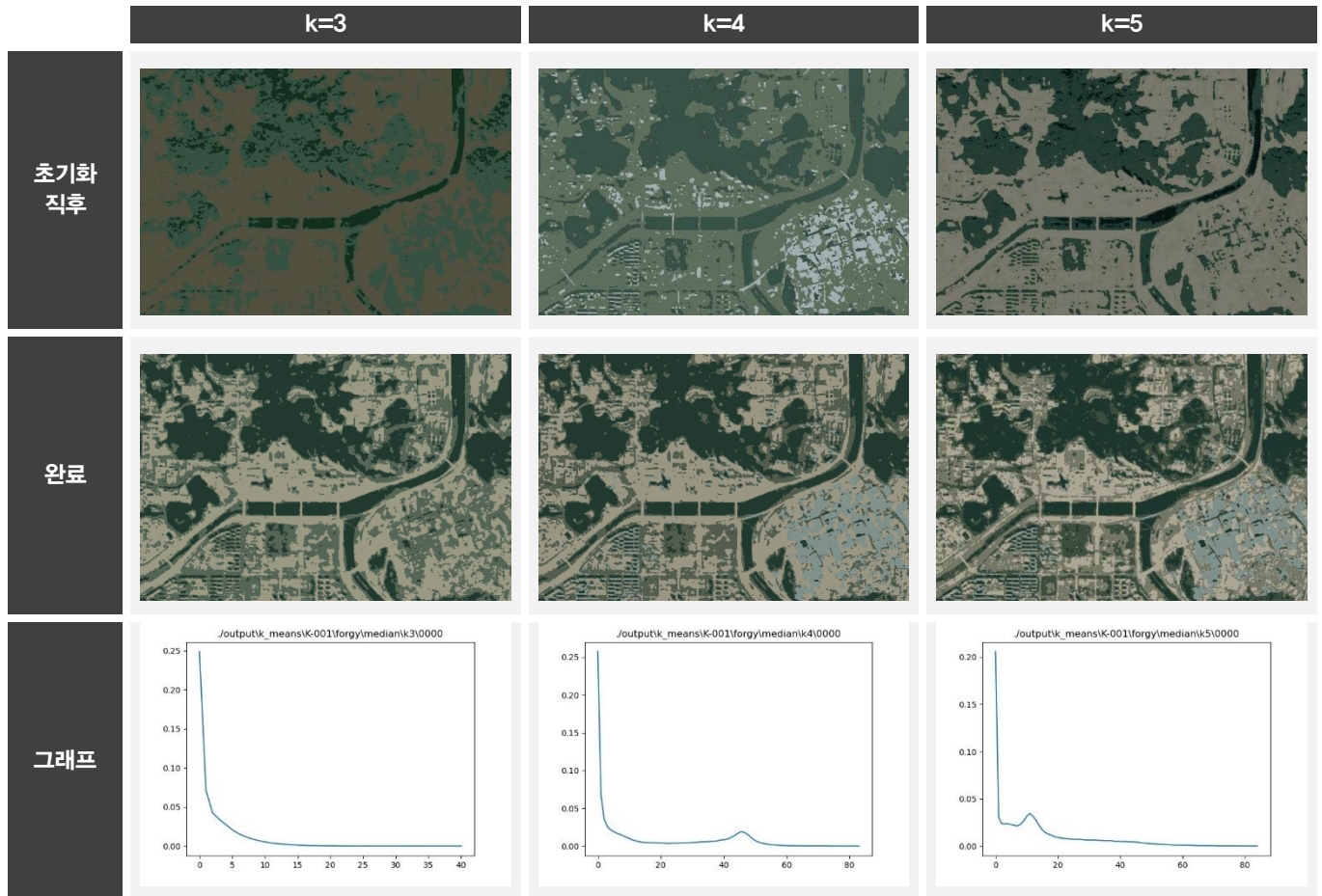


FIGURE 1. forgy 초기화를 적용한 K-Means 결과

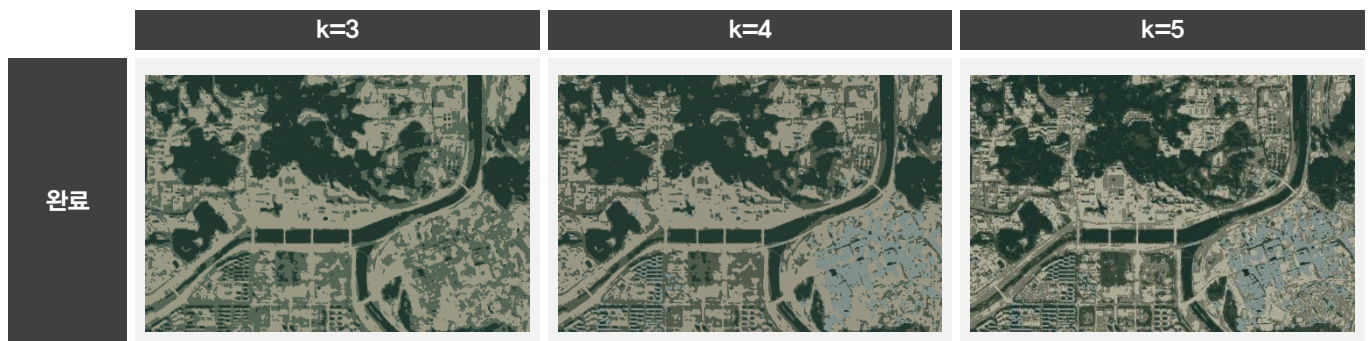
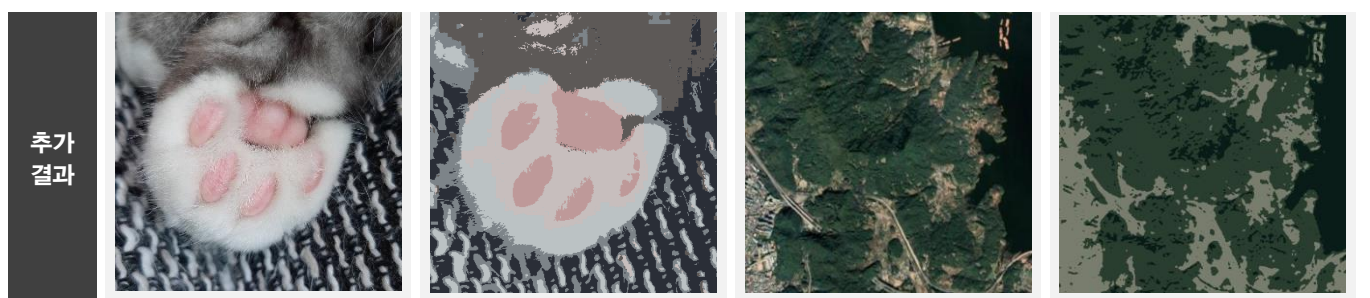
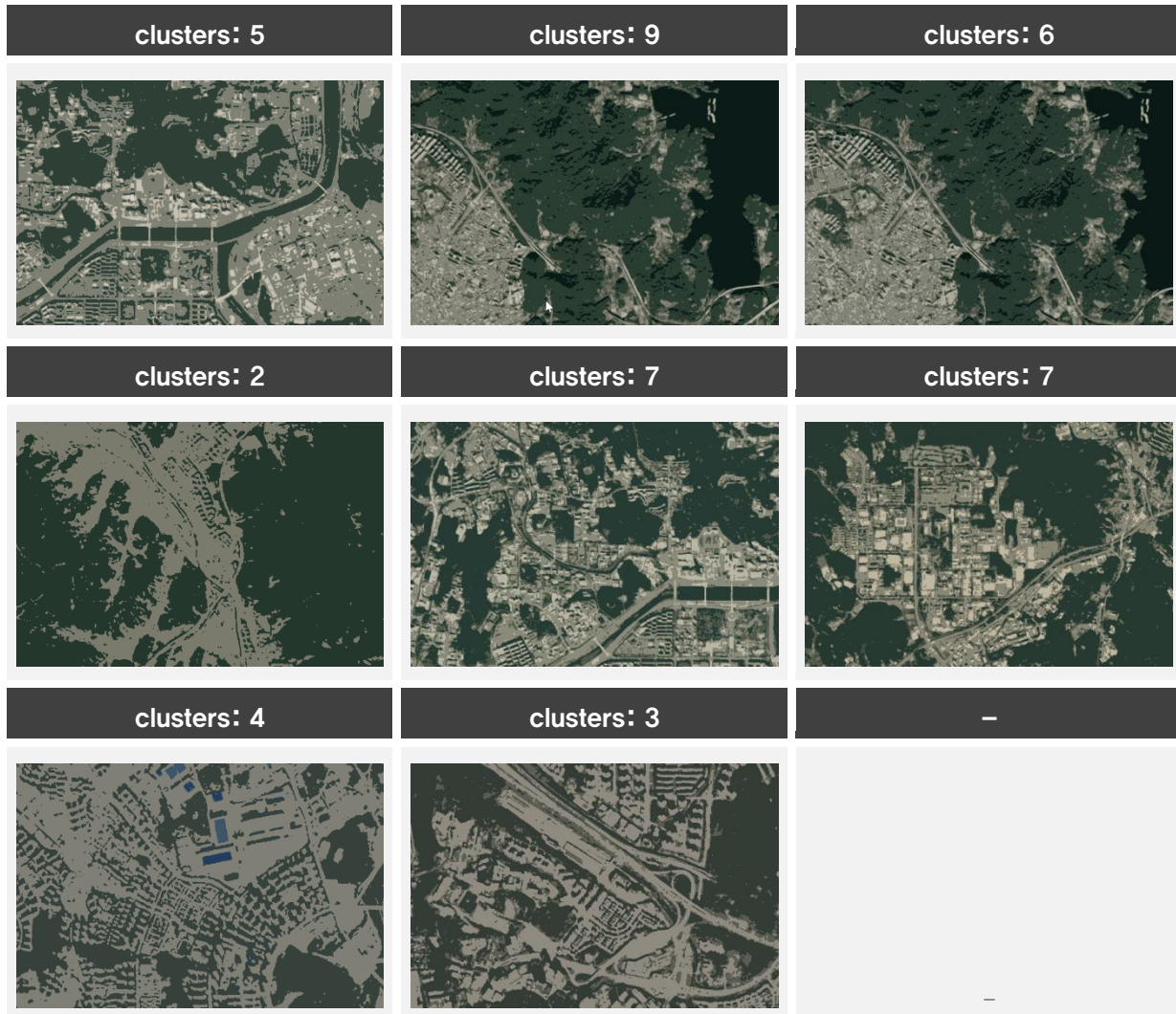


FIGURE 2. Random partitioning 초기화를 적용한 K-Means 결과



C. Result of MeanShift



D. Comparison

우선 두 방법의 가장 큰 차이는 클러스터링의 자동화 여부입니다. 실험데이터로 활용한 위성 사진의 경우, 지역에 따라 적으면 2개의 클러스터만 필요하기도 하며 훨씬 많은 수의 클러스터를 필요로 하는 경우도 있습니다. Mean-Shift 방법의 경우 자동으로 데이터의 밀집도를 기반으로 클러스터링하기 때문에 사람의 개입 없이도 도메인에 대한 일정 수준의 작업 성능을 유지합니다. 반면 K-means는 클러스터의 수를 사람이 직접 지정해주어야 하기 때문에 같은 도메인의 데이터라고 하더라도 작업 성능이 일정하지 않습니다. 예를 들어, {강, 산}만 있는 지역을 담은 위성사진과 {강, 산, 평야, 바다, 도심}이 복합적으로 있는 위성사진에 대해서는 k값을 결정하기 어렵습니다.

반면 구현 및 실행속도 차이에서는 K-means가 확실한 강점을 가집니다. K-means는 수렴 시점에 따라 0.3초 ~ 30초 사이에 처리 가능합니다. 100회 반복 실험 결과 평균적으로는 7초가량 소요됩니다. 동일 실행 환경 및 영상에 대해서 mean-shift는 약 210초가 소요되었습니다.

Pattern Recognition/ Feature Detection (matching). Feature detection and matching are an essential component of many computer vision applications. The concept of Feature matching is how can we extract local descriptors that are invariant to inter-image variations and yet still discriminative enough to establish correct correspondences?

(Problem 2) Selecting good features in pattern recognition is very important issue. Discuss the general ways (or methods) to select good features in developing an image classification algorithm. What types of the invariant feature in pattern recognition are used in pattern recognition? Explain the invariant features and importance of such features.

(Answer)

형태학적(morphological) 특징을 추출하기 위해 커널 내 픽셀간 색상 변화의 미분값을 활용합니다. 보통 이런 특징은 커널 영역을 변, 모서리, 평면으로 구분됩니다. 이 중에서도 컴퓨터 비전의 고전적 알고리즘 기반 특징 추출 기법에서 활용하는 대중적인 특징은 모서리(corner)입니다. 대부분의 특징점 추출 알고리즘은 이러한 모서리 특징 추출에 기반하고 있다는 점에서 영상 특징 선택에 있어 가장 주요한 방법이라고 할 수 있습니다. 모서리 특징을 활용하는 방법들은 다음 문단에서 설명합니다.

해리스 코너 탐지는 Moravec's corner detector 에서 활용된 방법을 사용했습니다. Moravec's corner detector 는 각 픽셀에 대해 원도우를 이동시키며 영상변화량(SSD)을 측정하여 이 값이 지역적으로 극대가 되는 지점을 코너로 인식하는 방법을 활용했습니다. 해리스 코너 탐지는 이 방법을 수정, 보완한 방법으로, 테일러 근사를 적용하여 변화량에 벡터를 적용해 평행이동, 회전이동에 대해 불변인 특징을 검출합니다. 또한 affine 변화와 밝기/대비 변화에도 어느정도 강인(robust)한 성능을 가집니다. 그러나 크기(scale)에 대해 매우 민감하게 영향을 받는다. 따라서 해리스 코너 검출 방법을 실용하기 위해서는 여러 크기에서 특징점을 검출할 필요가 있습니다.

SIFT 는 DoG(Difference of Gaussian)를 기반으로 크기(scale)불변인 특징을 검출하는 방법입니다. 해리스 코너 검출은 영상의 변화량과 방향 벡터만을 활용하기 때문에 크기에 매우 민감하게 반응하는데, SIFT 는 코너 검출을 위해 Laplacian 함수를 활용한다는 데에 큰 차이가 있습니다. Laplacian 은 밝기 변화에 대한 2 차 미분값으로, 밝기 변화가 일정한 곳에서는 0 에 가까운 값, 영역의 경계와 같이 밝기 변화가 급격한 곳에서는 높은 값을 나타냅니다. 다만 SIFT 는 속도를 개선하기위해 Gaussian Filter 를 활용한 각 영상들의 차를 이용해 Laplacian 을 근사적으로 계산하는 방법을 채택했습니다.

위 두 방법에서 알 수 있듯, 불변(Invariant) 특징이란, 여러 형태 변화에도 불구하고 그 특징을 유지할 수 있어야 합니다. 기본적으로 이러한 형태 변화에는 일차적인 영상 변환으로 수행되는 것들을 의미합니다. 회전, 반전, 이동이 있습니다. 추가로, 조도, 대비, 명도, 채도 등의 변화와 같은 색상 변화나 scale 변화 등이 있으며, 모서리 특징은 다른 특징에 비해 이러한 변화들에 대해 비교적 강인하게 추출될 수 있습니다.

Image Classification discussion.

(Problem 3) What is the difference between parametric classification and non-parametric classification? Show one representative method of each category and explain how each method works in detail. And apply each the algorithm to above image(City of Suwon).

(Answer)

머신 러닝을 기반으로 하는 방법들은 입력 x 와 출력 y 에 대해 다음과 같이 일반적인 함수 f 로 표현할 수 있습니다.

$$y = f(x)$$

머신 러닝의 목적은 위의 함수 f 를 생성하는 것입니다. 컴퓨터 비전에서 머신러닝 기반의 방법들은 다음과 같은 단계를 요구합니다.

1. 개체 특징 정의 및 추출
2. 개체 특징 일반성 학습

다시 돌아와서, 컴퓨터 비전이 아닌 일반적인 데이터에 대한 머신러닝 함수는 다음과 같이 극히 간단할 수 도 있습니다.

$$b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 = 0$$

Parametric 머신 러닝 기법은 위 함수의 b_1 , b_2 와 같은 파라미터를 구체적으로 명시하는 방법을 의미합니다.

대표적으로 로지스틱 회귀, 원시적 퍼셉트론 기반의 방법, Naïve Bayes 등이 있습니다. 이러한 prametric 방법들은 다음과 같은 특징을 지닙니다.

- 장점: 구조적 단순성, 빠른 속도, 적은 데이터
- 한계: 특징 제약, 기능적 단순성, 불일치성

non-parametric 방법은 위와 반대로 학습에 parameter를 갱신하는 과정을 수반하지는 않는 방법들을 의미합니다.

SVM, kNN, decision tree와 같은 방법들이 있습니다. 이러한 방법들은 다음과 같은 특징을 지닙니다.

- 장점: 유연성, 일반성, (상대적)고성능, 특징을 가정하지 않음.
- 한계: 다량의 데이터 요구, 과적합, 느린 훈련 속도

non-parametric 알고리즘 중 하나로 SVM을 활용, parametric 방법으로 random forest 기반의 방법을 위성 사진에 적용해보았습니다. segmentation 방법은 다양한 방법으로 수행되지만, 픽셀에 대한 이진 분류로도 수행될 수 있습니다. 따라서 pixel-wise classification을 수행하여 이진 마스크를 생성한 결과를 아래 그림으로 보입니다.





위 그림은 입력 이미지와 RF(Random Forest), SVM(Support Vector Machine) 분류기를 기반으로 영상을 binary segmentation한 결과입니다.

훈련 데이터로 정확한 이진 마스크를 Label로 생성해야하기 때문에 3장의 훈련데이터로만 훈련을 진행했기 때문에 전체적인 성능이 낮습니다.

가장 오른쪽 이미지는 훈련 데이터와 가장 유사한 위성사진입니다. 건축물을 target으로 하는 segmentor라고 볼 수 있습니다. 유사한 이미지에 대해서는 상당히 잘 작동하는 모습을 보여줍니다.

정상적인 작동을 하지는 않았지만 대략적으로 Parametric 방법과 Nonparametric 방법의 특성이 드러납니다.

RF 기반 방법은 비교적 선명하고 세밀한 mask가 출력되었습니다. 간단한 구조의 분류기로 인해 추론과정에서 데이터가 크게 가공되지 않은 것으로 보입니다. 반면 SVM mask는 비교적 추상적인 형태를 띄고 있습니다. 결과적으로 매우 부정확한 mask지만, 데이터 부족으로 인한 학습 미달 현상이 저와 같이 드러나는 것으로 보입니다.