5월 1째주 연구노트 – SVM-KNN2

# 모양 및 텍스처 거리

SVM-KNN을 적용 할 때 우리는 시각적 대상 인식의 두 가지 주요 단서 인 모양과 질감에 기반한 분류에 주력합니다. 다음과 같이 여러 가지 우수한 거리 함수를 소개합니다.

## 텍스처에 대한 χ2 거리

Leung과 Malik에 따르면, 텍스처의 이미지는 "textons"의 히스토그램에 맵핑 될 수 있습니다.이 텍스 트는 다양한 유형의 텍스처 요소의 분포를 캡처합니다. 거리는 두 개의 텍스트 히스토그램 사이의 피어슨의 χ2 검정 통계로 정의됩니다.

# 질감의 한계 거리

통계적 관점에서 텍스처 위의 χ2 거리는 텍스처 응답의 두 조인트 분포 사이의 차이를 측정하는 것으로 볼 수 있습니다. 텍스처 조각이 필터 뱅크를 통과하고 응답의 공동 분포가 텍스트로 벡터 양자화되며, 텍스트의 히스토그램을 비교합니다. Levina et al. 은 관절 분포의 차이점 (즉, 각 필터 응답의 히스토그램)을 간단히 관찰함으로써 관절 분포가 종종 서로 잘 구별 될 수 있다는 것을 발견했다. 따라서 텍스처에 대한 또 다른 거리 함수는 각 필터의 응답 히스토그램 사이의 거리를 합산하는 것입니다. 이것은 안정적으로 양자화 할 수있는 너무 많은 유형의 텍스트를 포함 할 수있는 실제 이미지에 대한 실험에서 사용됩니다.

## 접선 거리

숫자의 그레이 스케일 이미지 쌍에 정의 된 접선 거리는 두 개의 선형 부분 공간 (픽셀 영역 Rn에서 n은 픽셀 수) 사이의 최소 거리로 정의되며, 작은 아핀 변환으로부터의 섭동을 포함하여 이미지에서 파생됩니다 펜 스트로크의 두께 변화 (7 차원 선형 공간 형성).

## 모양 기반 컨텍스트 기반 거리

셰이프 컨텍스트의 기본 개념은 다음과 같습니다. 셰이프는 컨트롤 포인트에서 해당 포인트 주변의 "가로"를 캡처하는 설명자를 사용하여 점 집합으로 나타냅니다. 이러한 설명자는 변형 모델을 사용하여 반복적으로 일치됩니다. 거리는 최종 일치 모양에 남아있는 불일치와 어 피연 변이에서 얼마나 변형되었는지를 나타내는 점수에서 파생됩니다.

## 기하학적 블러 기반 거리

예를 들어 에지 맵상의 형상 문맥 기술자 또는 SIFT 기술자 또는 기하학적 흐림 기술자와 같은 다수의 형상 기술자가 그레이 스케일 이미지 상에 정의 될 수있다. 우리의 실험에서, 우리는 기하학적 흐림 기술자에 초점을 맞 춥니 다. 일반적으로 에지 포인트에 정의 된 기하학적 흐림 설명자는 에지 응답의 주변 패치에 공간적으로 변화하는 흐림 효과를 적용합니다. 중심에서 멀리 떨어진 지점은 변형시 공간적 불확실성을 반영하기 위해 더 흐려집니다. 이 블러 링 후에, 디스크립터는 L2 norm 1을 갖도록 정규화된다.

## Kernelizing the distance

비대칭 : (모양 컨텍스트를 기반으로 거리와 기하학적 흐림 기반 거리의) 우리는 단순히 대칭 거리를 정의합니다 : D (X, Y) + D (Y를 X) 때문에 실제로 불일치 | D (X, Y) - D ( y, x) | 작다.

삼각 부등식 : 즉, 부등식 D (X, Y) + (D) (Y, z) ≥ d를 (X, Z) (접선의 거리, 형상 컨텍스트 기반 형 거리 기하학적 흐림 기반 거리) 항상 보유하지 않는 거리가 긍정적 인 명확한 커널로 변환되는 것을 방지합니다. 이 문제에 대한 여러 가지 해결책이 제안되었습니다. 여기서 우리는 커널 행렬의 최소 고유치를 계산하고, 그것이 음수라면 커널 행렬의 대각선에 절대 값을 더한다. 직관적으로, 우리가 커널 행렬을 일종의 "유사성 척도"로 본다면, 대각선에 양수 상수를 추가한다는 것은 자기 유사성을 강화시키는 것을 의미하며, 이는 예제들 사이에 표현 된 유사성에 영향을 미치지 않아야한다.

# Performance on benchmark data sets 4.1. MNIST

MNIST의 수기 디지트 세트는 훈련을위한 60,000 개의 예제와 테스트를위한 10,000 개의 예제가 포함되어 있습니다. 각 세트에는 센서스 국 직원과 고등학생 두 개의 별개의 인구 수의 동등한 숫자가 들어 있습니다. 각 자릿수는 모양 컨텍스트 계산을 제외하고 28x28 이미지입니다. 각 자릿수는 70x70 이미지로 조정됩니다. 다수의 최첨단 알고리즘이 1 % 오류율 미만에서 수행되며, 그 중 형상 컨텍스트 기반 방법은 .67 %에서 수행됩니다.

이 실험에서는 두 거리, 즉 L2 및 형상 상황 거리를 사용합니다. 모양 컨텍스트의 경우 오류 비율이 베이 즈 최적에 가까울 수 있으므로 처음 10,000 개의 학습 예제 만 사용하여 개선의 여지를 남깁니다 (10,000 개의 예제에서 10 배 교차 유효성 검사 수행). 이미지의 강도가 아닌 모양의 컨텍스트에만 의존하기 위해 "모양"도 삭제합니다.

결과의 요약은 표 2에 나와 있습니다. L2 거리는 본 방법에서 간단하지만 형상 컨텍스트 기반 거리에는 여러 가지 해결 방법이 필요합니다. 그러나 두 경우 모두 성능이 크게 향상됩니다.

|  |  |  |
| --- | --- | --- |
|  | L2 | SC (limited training) |
| SVM-KNN | 1.66 (K = 80) | 1.67 (±0.49) (K = 20) |
| NN | 2.87 (K = 3) | 2.2 (±0.77) (K = 1) |

표 1. MNIST의 오류율 (%) : 각 알고리즘의 매개 변수 K는 최상의 성능 (NN의 경우 [1,10], SVM-KNN의 경우 [5, 10, .., 100]의 범위)에 따라 선택됩니다. ). SVMKNN에서 Ksl ≈ 10K 매개 변수는 Ksl이 클수록 경험적 결과를 향상시키지 않습니다.

## USPS

USPS 데이터 세트에는 버팔로의 우편 봉투에서 수집 한 9298 자필 자릿수 (교육용 7291 개, 테스트 용 2007 년)가 포함되어 있습니다. 각 자릿수는 16x16 이미지입니다. 무작위 테스트 샘플 모음. USPS 테스트 세트는 다소 어려우며 인간의 오류율은 2.5 %입니다.

L2와 접선 거리의 두 가지 유형의 거리를 시도합니다. 이미지가 너무 작아 변형을 추정하기에 충분한 세부 정보가 포함되어 있지 않기 때문에 모양 컨텍스트가 시도되지 않습니다. 접선 거리의 경우, 각 이미지는보다 안정적인 접선을 얻기 위해 너비가 σ = 0.75 인 가우스 커널로 부드럽게 처리됩니다.

|  |  |  |
| --- | --- | --- |
|  | L2 | tangent distance |
| SVM-KNN | 4.285 (K = 10) | **2.59** (K = 8) |
| NN | 5.53 (K = 3) | 2.89 (K = 1) |
| DAGSVM | 4.4 (Platt et al. [31]) | Intractable |
| HKNN | 3.93 (Vincent et al. [41]) | N/A |

표 2. USPS의 오류율 (백분율) : SVM-KNN 및 NN의 매개 변수 K는 최상의 성능에 따라 각각 선택됩니다.

표 2은 L2 사례에서 SVMKNN과 DAGSVM의 오류율이 비슷 함을 보여줍니다. 그러나 SVM-KNN은 각 SVM이 10 개 샘플의 지역 인접성 만 포함하기 때문에 교육하기가 훨씬 빠르며 클래스의 수는 이웃에서 거의 4 개를 초과하지 않습니다. (우리의 실험에서 4 개의 클래스에서 10 개의 예제를 학습하는 비용은 일반적인 가장 가까운 이웃 탐색의 비용보다 훨씬 적습니다.) 반면, DAGSVM은 45 개의 모든 클래스 (예 : 10x9 / 2) 모든 훈련 예에서 쌍 거리의 계산. 보다 비용이 많이 드는 탄젠트 거리 함수를 사용하면 DAGSVM이 실험에서 훈련하기가 어려워지는 반면 최적의 SVM-KNN (K = 8)은 일반적인 NN 분류기만큼 빠릅니다. SVM을 8 가지 예제로 훈련하는 추가 비용이 무시할 만하다. 이것은 2 절의 점근선 복잡성의 비교를 반영한다.

또한 L2 사례에서 HKNN (Vincent et al.)과 같은 또 다른 적응 형 최근 접 이웃 기법이 매우 잘 수행된다. 불행히도 입력 공간에서 작동하므로 L2 이외의 거리 함수로는 확장 할 수 없습니다.

SVM-KNN이 매우 적은 추가 비용으로 NN의 성능을 개선 할 수 있다는 것은 상당히 주목할 만하다. 비록 후자가 이미 (인간의 성능과 비교하여) 매우 잘 수행되고있다. 그러므로 우리는 SVMKNN이 기본 데이터 세트의 적절한 "불변"구조가 거리 함수에서 캡처 될 때 이상적인 분류 방법이라고 생각하는 것이 좋습니다.