

D-SVDD를 이용한 패턴 노이즈 제거

Pattern De-Noising using D-SVDD

강대성, 박주영

고려대학교 제어계측공학과
E-mail: {mpkds, parkj}@korea.ac.kr

요 약

SVDD(support vector data description)는 one-class 서포트 벡터 학습 방법론 중 하나로 비정상 물체에서 정상 데이터를 구분하기 위해서 특징 공간(feature space)에서 정의된 구를 이용하는 전략을 쓰는 방법론이다. 하지만 SVDD는 모든 데이터에 대해서 같은 중요도를 부가하는 단점을 가지고 있다. 최근에, 이런 문제점을 보완하기 위해 데이터의 밀도 분포에 따라서 중요도를 다르게 부가하는 D-SVDD(density-induced support vector data description) 방법론이 발표되었고, 아직도 많은 연구가 진행되고 있다. 본 논문에서는 D-SVDD를 이용해서 노이즈가 섞인 비정상 데이터를 노이즈가 제거된 정상 데이터로 복원하는 방법에 대해서 논한다. 특히, 본 논문에서 제안하는 방법론을 다른 방법론과 비교하여 본 논문의 방법론의 효용성에 대해서 다룬다.

Key Words : SVDD, D-SVDD, 노이즈 제거

1. 서 론

최근 들어, 서포트 벡터 학습 방법은 이에 관련 이론이 정립되고 각종 응용 사례가 보고되면서, 지능시스템 분야에서 매우 주요한 도구 중 하나로 자리를 잡아가고 있다. 서포트 벡터 학습의 주요 응용 대상으로 최근에 주목받고 있는 분야 중 하나로는 비정상 상태 탐지 문제(abnormality detection problem)를 들 수 있다. 비정상상태 탐지문제에서는 일반적으로 학습과정에서 정상 클래스(normal class) 위주의 데이터만을 다루게 되고, 학습결과로 얻어진 결정함수(decision function)는 주어진 시험 데이터가 정상 상태 범주를 벗어나는 비정상 클래스(abnormal class)에 속하는지를 판단하게 된다. 이러한 문제는 단일 클래스 분류 문제(one-class classification problems)로 불리기도 하며, 운행 중인 기계의 고장진단이나 컴퓨터 보안과 관련한 침입 탐지 시스템(intrusion detection system)과 같은 부류의 문제에 효과적으로 사용될 수 있다. SVDD(support vector data description)는 단일 클래스 분류 문제를 다룰 때 사용하는 서포트 벡터 학습 방법 가장

대표적인 방법론이다.[1, 2] SVDD는 비정상 물체에서 정상 데이터들을 구별하기 위해서 특징 공간에서 정의된 구를 이용하는 전략을 사용한다. 즉, 구 안에 포함된 데이터들은 정상 데이터들로 인식하고, 구 밖에 있는 데이터들은 비정상 데이터들로 인식하는 알고리즘을 사용하는데 모든 데이터의 중요성을 동일하게 취급하는 단점을 가지고 있다. 최근에 이 단점을 보완하기 위해서 정상상태 내에서 데이터의 밀도 분포가 상대적으로 큰 곳에는 보다 큰 중요도를 부가하고, 데이터의 밀도 분포가 상대적으로 작은 곳에는 보다 작은 중요도를 부가하는 D-SVDD(density-induced support vector data description) 방법론이 제안되었다.[3] 본 논문에서는 D-SVDD 방법론을 활용하여 비정상 상태의 데이터를 정상 상태로 복구하는 기법을 제안하고, 제안된 기법을 이용하여 노이즈가 섞인 얼굴 영상의 복원에 응용하는 문제를 다룬다. 본 논문의 구성은 다음과 같다. 2장에서는 SVDD와 D-SVDD의 기본 이론에 대해서 간단히 기술한 후, D-SVDD를 이용해서 얼굴 이미지를 복원하는 방법에 대해서 다룬다. 3장에서는 시뮬레이션에 사용한 얼굴 영상

과 각종 파라미터, 그리고 시뮬레이션 절차에 대해서 논한 후, 4장에서는 실험 결과에 대한 고찰과 PCA 방법론[5]과 KPCA 방법론[4]과 비교하여 얼마나 더 좋은 결과를 보이는지에 대해서 다룬다.

2. 기본 이론

2.1 SVDD 기본 이론

정상 영역에 속하는 물체의 서포트(support)를 근사하는 SVDD 방법론은 다음과 같이 구해진다: 중심이 $a \in R^d$ 이고, 반경이 R , 그리고 데이터 집합 D 를 가지는 구 B 가 있다고 하자. SVDD의 주된 목적은 다음의 상충되는 두 가지 목적을 동시에 달성하는 것이다. 하나는 가능한 구가 작아야하고, 또 다른 하나는 가능한 많은 학습 데이터를 포함해야 한다. 이러한 목적을 만족하는 구는 다음의 최적 문제를 풀면 얻을 수 있다.

$$\min L_0(R^2, a, \xi) = R^2 + C \sum_{i=1}^N \xi \quad (1)$$

s. t. $x_i - a^2 \leq R^2 + \xi_i, \xi_i \geq 0, i = 1, \dots, N$

위의 식에서, 변수 ξ_i 는 구 밖으로 벗어난 i 번째 학습 데이터와 관련된 벌점이다. 식(1)의 목적함수는 상충되는 두 항인 반지름의 제곱 R^2 과 총 벌점 $\sum_{i=1}^N \xi_i$ 로 구성되어 있다. trade-off 상수 C 는 이 두 항들의 상대적인 중요도를 조절한다. 하지만 본 논문에서는 많은 제약이 있는 입력 공간 R^d 가 아닌 특징 공간 F 에서 문제를 다루려고 한다. 특징 공간에서 가능한 반지름이 작고, 가능한 많은 데이터를 포함하는 구 B_F 를 구하는 것은 다음의 QP(Quadratic Problem) 문제를 풀면 쉽게 구할 수 있다.

$$\min L_F(R_F^2, a_F, \xi) = R_F^2 + C \sum_{i=1}^N \xi \quad (2)$$

s. t. $\|\phi(x_i) - a_F\|^2 \leq R_F^2 + \xi_i, \xi_i \geq 0, i = 1, \dots, N$

라그랑지 함수와 쌍대 문제(dual problem), 가우스 커널 트릭을 이용해서 식(2)을 풀면 우리는 다음과 같은 식을 얻을 수 있다.

$$\min_{\alpha} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \quad (3)$$

s. t. $\sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, C], \forall i$

이 때에 라그랑지 함수의 최적해를 만족시키는 안장점의 조건은 $a_F = \sum_{i=1}^N \alpha_i \phi(x_i)$ 이 되고, 이 식을 이용해서 특징 공간상에서 테스트 데이터가 정상인지 비정상인지 판별하는 결정함수는 다음과 같은 식이 된다.

$$f_F(x) = R_F^2 - \|\phi(x) - a_F\|^2 \geq 0 \quad (4)$$

2.2 D-SVDD의 기본 이론

D-SVDD에서는 데이터 간의 밀도 분포를 알기 위해서 다음과 같은 변수를 사용한다.

$$\rho(x_i) = \exp\left(w \times \frac{MEAN_K}{d(x_i, x_i^K)}\right) \quad (5)$$

여기서 x_i^K 는 데이터 x_i 의 K 번째 가까운 이웃 데이터이고, $d(x_i, x_i^K)$ 는 x_i^K 와 x_i 간의 거리, $MEAN_K$ 는 $d(x_i, x_i^K)$ 의 평균값을 의미한다. 그리고 w 는 0과 1사이의 임의의 수를 가지는 파라미터인데 만약 w 가 0의 값을 가지면 앞에서 설명한 SVDD와 같은 식이 된다. 그리고 식(5)에서 ρ 의 값이 클수록 높은 밀도 분포를, 작을수록 낮은 밀도 분포를 의미한다. 식(5)을 이용하여 특징 공간에서 임의의 데이터와 D-SVDD 구의 중심 a_D 간의 거리를 다음과 같이 구할 수 있다.

$$\delta(x_i) = \rho(x_i)^{1/2} \|\phi(x_i) - a_D\| \quad (6)$$

그리고 특징 공간에서 데이터간의 밀도 정보를 가지는 D-SVDD 구를 근사하는 최적화 문제는 다음과 같은 식으로 표현 할 수 있다.

$$\min \tilde{L}_0(R_D^2, a_D, \xi) = R_D^2 + C \sum_{i=1}^N \xi \quad (7)$$

s. t. $\delta^2(x_i) \leq R_D^2 + \xi_i, \xi_i \geq 0, i = 1, \dots, N$

라그랑지 함수와 가우스 커널 트릭을 이용하여 D-SVDD의 중심 a_D 를 구한 후,

$$a_D = \frac{\sum_i \alpha_i \rho(x_i) \phi(x_i)}{T} \quad (8)$$

식(7)의 쌍대 문제를 다음과 같이 구할 수 있다. 단, 여기서 $T = \sum_{i=1}^N \alpha_i \rho(x_i)$ 이다.

$$\min_{\alpha} \frac{\sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \rho(x_i) \rho(x_j) k(x_i, x_j)}{T} - T$$

$$s.t. \sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, C], i = 1, \dots, N \quad (9)$$

하지만, 위의 쌍대 문제는 선형 제한 최적화 문제(linear constrained optimization problem)가 되기 때문에 Matlab에서 제공하는 quadprog 함수를 이용하여 구할 수 없다. 이를 구하기 위해서는 위 문제를 SDP 문제로 바꾼 후, 본 논문에서는 Matlab에서 제공하는 LMI 툴박스를 이용하여 풀었다. D-SVDD를 이용한 결정함수는 다음 식처럼 된다.

$$f_D(x) = R_D^2 - ||\phi(x) - a_D|| \geq 0 \quad (10)$$

2.3 D-SVDD를 이용한 노이즈 제거

본 논문의 목적은 특징 공간 내에서 노이즈로 왜곡된 테스트 데이터를 구 모양의 서포트를 이용해서 정상 데이터로 바꾸는 것이다. 노이즈를 제거하는 방법은 특징 공간상에 있는 비정상 데이터를 직접 정상 데이터 영역으로 사영시키면 된다. 정상 데이터를 이용해서 학습을 하면 그림 1에 있는 작은 구가 만들어지는데 이 구가 바로 D-SVDD 구이다. 그리고

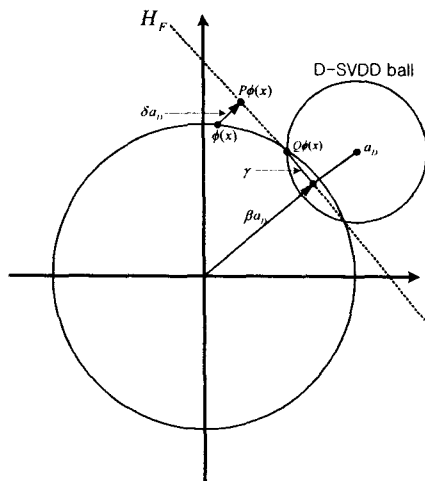


그림 1. 본 논문에서 제안하는 기본 개념

RBF 커널을 사용하면 모든 데이터들은 그림 1에 있는 것과 같이 크기가 1인 단위 구의 표면에 있게 된다. 테스트 데이터 $\phi(x)$ 는 항상 단위 구의 표면에 있지만 노이즈가 섞여있는 경우에는 D-SVDD 구에는 포함되지 않는다. 따라서, $\phi(x)$ 를 D-SVDD 구와 단위 구와 만

나는 교점인 $Q\phi(x)$ 로 옮김으로써 노이즈를 제거하게 된다. 즉, $\phi(x)$ 를 $Q\phi(x)$ 로 옮기는 것이 본 논문의 목적이며 비정상 데이터의 노이즈를 제거하여 정상 데이터로 만드는 과정이다. 이를 위해서는 다음과 같은 간단한 절차를 밟으면 된다. 그림 1처럼 비정상 데이터 $\phi(x)$ 를 초평면 H_F 로 사영시키고 그 결과 $P\phi(x)$ 를 정상 데이터 영역인 $Q\phi(x)$ 로 옮겨주는 과정을 밟으면, $Q\phi(x)$ 가 다음과 같은 형태로 표현될 수 있음을 쉽게 보일 수 있다.

$$Q\phi(x) = \lambda_1 \phi(x) + \lambda_2 a_D \quad (11)$$

단, 여기에서 λ_1 과 λ_2 는 그림 1과 위의 설명을 따라 구해지는 양수이다. 이상의 과정을 밟아 특징 공간 내의 de-noised pattern $Q\phi(x)$ 가 구해지면, 이 특징 벡터의 pre-image[4]를 구함으로써 입력 공간 위에서 정의되는 de-noised pattern을 구할 수 있다.

3. 실험

3.1 얼굴 영상 데이터

이번 실험에 사용한 얼굴 영상¹⁾은 모두 200개의 2차원 백인 얼굴 이미지를 사용하였는데 이 이미지는 레이저 스캐너(CyberwareTM)로 기록되어진 3차원 얼굴 모델의 데이터베이스에서 얻어진 것이다. 원래 이 이미지들은 얼굴에서 머리카락을 완전히 제거한 이미지로 256x256 픽셀 크기의 칼라 이미지이다. 본 논문에서는, 실험의 편리함을 위해서 256x256 픽셀의 칼라 이미지를 64x64, 8비트 흑백 이미지로 변환하여 실험에 사용하였다. 그리고 이번 실험에서 사용하는 200개의 이미지 중에서 임의로 95명의 남자와 95명의 여자 이미지를 선택하여 총 190개의 이미지를 학습 데이터로 사용하였고, 나머지 5명의 남자와 5명의 여자 이미지를 테스트 데이터로 사용하였다. 실험에서 사용한 파라미터로는 ρ 를 구할 때 필요한 K 번째 데이터를 5번째 데이터를 사용하였고 w 는 0.2를 사용하였다. 그리고 $Q\phi(x)$ 의 pre-image를 구할 때 필요한 neighborhood의 개수는 10개로 하여서 입력 공간에서의 결과를 구했다. 노이즈 이미지는 원본 테스트 이미지에 아래의 식과 같은 임의의 가우스 노이즈를 주었고 이때 사용한 noise var=600이었다.

1) 본 논문에서 사용한 얼굴영상은 Max Plank Institute for Biological Cybernetics로부터 제공받은 것임.

$$I_{noise} = I_{original} + randn * \sqrt{noise\ var} \quad (12)$$

3.2 시뮬레이션 절차

본 논문에서 제안한 방법론을 통해서 얼굴 영상에 섞여있는 노이즈를 제거하는 방법의 절차는 다음과 같다:

- ① 학습 데이터로 선택된 190개의 얼굴 이미지와 노이즈가 섞인 테스트 데이터를 64×64 픽셀의 8비트 흑백 이미지로 변환한다.
- ② 변환된 학습 데이터를 LMI를 이용하여 특징 공간에서 학습을 한 후, 결정함수를 구한다.
- ③ 테스트 데이터 $\phi(x)$ 를 본 논문에서 제안한 노이즈 제거 방법론을 이용해서 각각 $Q\phi(x)$ 를 구한다.
- ④ 특징 공간상에 있는 $Q\phi(x)$ 를 입력공간으로 옮기기 위해서 pre-image[4]를 이용해서 노이즈가 제거된 이미지를 얻는다.

4. 결과 고찰

본 논문에서는 노이즈가 섞인 64×64 픽셀 8비트의 비정상 이미지를 노이즈가 제거된 정상 이미지로 복원하는 방법에 대해서 논하였다. 먼저 본 논문의 결과를 결과 이미지와 테스트 데이터의 원본 이미지간의 각 픽셀에 대한 distance를 이용한 성능지수를 통해서 비교하면 PCA가 본 논문에서 제안한 방법론에 비해서 조금 더 좋은 값을 보인다. 하지만 이런 결과는 결과 이미지의 평균 밝기가 큰 영향을 미치기 때문에 PCA의 성능지수가 더 좋게 나왔다고 생각된다. 오히려 육안 상으로 결과들을 비교해 보면 본 논문에서 제안한 방법론이 다른 방법론에 비해서 더 좋은 결과를 나타냈다고 볼 수 있다. 물론 학습 이미지의 얼굴 크기나 피부 색, 눈, 코, 입 등의 위치가 모두 다르기 때문에 지금과 같은 이미지만을 기반으로 한 방법론에서는 결과 이미지가 선명하게 나오지 않는 것은 당연하다. 하지만 본 논문에서 제안한 방법론이 다른 결과 이미지와 비교해서 블리링의 효과가 덜하고 더 복원이 잘 되었다는 것을 알 수 있다. 하지만, 본 논문에서 제안한 방법론에서 선형 제한 최적화 문제를 풀기 위해서 Matlab LMI 툴박스를 이용하는데 이때 솔루션을 구하는 시간이 다른 방법론에 비해서 많이 걸린다. 이 문제를 해결하는 것이 향후 과제가 될 것이다.



그림 2. 테스트 데이터의 원본 이미지



그림 3. 테스트 데이터의 노이즈 이미지



그림 4. PCA 방법론을 이용했을 때 복원된 이미지



그림 5. KPCA 방법론을 이용했을 때 복원된 이미지



그림 6. 제안한 방법론을 이용했을 때 복원된 이미지

참 고 문 헌

[1] D. Tax and R. Duin, "Support Vector Domain Description," Pattern Recognition Letters, Vol. 20, pp. 1191-1199, 1999
 [2] D. Tax, "One-Class Classification", Ph.D. Thesis, Delft University of Technology, 2001
 [3] K. Lee, D. Kim, D. Lee and K. Lee, "Improving support vector data description using local density degree," Pattern Recognition, Vol, 38, pp. 1768-1771, 2005
 [4] J. T. Kwok and I. W. Tsang. "The pre-image problem in kernel methods," IEEE Transactions on Neural Networks, Vol. 15, pp. 1517-1525, 2004.
 [5] S. Lee, J. Park and B. Hwang, "How can we reconstruct facial image from partially occluded or low-resolution one?," Lecture Notes in Computer Science, Vol. 3338, pp. 386-399, 2004.