

새로운 커널 기반 정상 상태 복구 기법과 응용

New Kernel-Based Normality Recovery Method and Applications

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 비 정상 물체에서 정상 데이터를 구분하기 위해서 특징 공간에서 정의된 구를 이용하는 전략을 쓰는 방법론이다. 본 논문에서는 SVDD를 이용해서 노이즈가 섞인 비정상 데이터를 노이즈가 제거된 정상 데이터로 복원하는 방법에 대해서 논한다. 그리고 저해상도의 이미지를 고해상도의 이미지로 복원함으로써 본 논문의 방법론이 어떻게 실용적으로 적용되는지에 대해서 다룬다.

Key Words : SVDD, 노이즈 제거, 얼굴영상복원

1. 서 론

최근 들어, 서포트 벡터 학습 방법은 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 비정상상태 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 정상상태 범주를 벗어나는 비정상 클래스(abnormal class)에 속하는지를 판단하게 된다. 이러한 문제는 단일 클래스 분류 문제(one-class classification problems)로 불리기도 하며, 운행중인 기계의 고장진단이나 컴퓨터 보안과 관련한 침입 탐지 시스템(intrusion detection system)과 같은 부류의 문제에 효과적으로 사용될 수 있다. 비정상상태 탐지 문제에 대한 고전적인 접근 방법으로는, 학습 데이터를 이용하여 확률밀도함수(probability density function)를 추정한 후, 시험용 개체(test objects)에 대한 밀도 값(density value)이 정해진 문턱 값(threshold value) 미만인 되는 경우에 비정상 클래스로 판정하는 전략이

있다. 그러나, 이러한 전략은 확률밀도함수의 정확한 추정에 기반을 두고 있으므로, 일반적으로 다량의 학습 데이터를 필요로 한다. 따라서, 제한된 개수의 학습 데이터만이 주어진 경우에는, 확률밀도함수의 추정을 이용하여 정상 클래스의 데이터가 존재하는 영역을 간접적으로 구하는 것보다는 정상 클래스가 존재하는 영역을 직접적으로 구하고자 하는 전략이 오히려 유리할 수 있다. 이러한 전략을 활용하는 방법들 중 가장 대표적인 예로 SVDD(support vector data description)을 들 수 있다[1, 2]. 본 논문에서는 SVDD 기법을 활용하여 정상상태를 복구하는 기법을 제안하고, 제안된 기법을 저해상도 얼굴영상의 복원에 응용하는 문제를 다룬다. 얼굴 인식에 대한 연구는 오랫동안 활발히 진행되어 왔고, 이에 관한 좋은 결과도 많이 나오고 있다. 하지만 보안 시스템에 많이 사용하는 CCTV나 카메라에 잡힌 영상을 선명하고 자세하게 보기에 여전히 많은 문제점이 있다. 왜냐하면 캡처된 얼굴 이미지는 노이즈가 섞이거나, 얼굴의 일부가 가려지거나 또는 이미지의 해상도가 떨어지기 쉽기 때문이다. 이러한 문제점을 해결하기 위해 본 논문에서는 최근 지능 시스템 분야에서 많은 연구가 진행되고 있는 one-class 서포트 벡터 학습 방법 중 하나인 SVDD (support vector data description)[1, 2]를 이용하여 저해상도의 얼굴

이미지를 고해상도의 얼굴 이미지로 복원하는 문제를 다루려 한다.[3] 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 인식해서 비정상 데이터들을 정상 데이터 영역의 구의 경계면으로 사영시켜서 비정상 데이터들의 노이즈를 제거하려고 한다. 본 논문의 구성은 다음과 같다. 2장에서는 SVDD에 대해서 간단히 기술한 후, SVDD를 이용해서 얼굴 이미지를 복원하는 방법에 대해서 다룬다. 3장에서는 시뮬레이션에 사용한 얼굴 영상과 시뮬레이션 절차에 대해서 논한 후, 4장에서는 결과와 고찰에 대해서 다룬다.

2. SVDD

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. 학습 데이터들은 노이즈의 영향을 받기 쉽기 때문에 학습 데이터들 중 일부는 비정상 데이터가 될 수도 있다. 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.

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

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

위의 식에서, 변수 ξ_i 는 구 밖으로 벗어난 i 번째 학습 데이터와 관련된 벌점이다. (1)의 목적 함수는 상충되는 두 항인 반지름의 제곱 R^2 과 총 벌점 $\sum_{i=1}^N \xi_i$ 로 구성되어 있다. trade-off 상수 C 는 각 항의 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.

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

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

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

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

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

만약 가우스 커널을 사용하면 커널 K 는

$$K(x, z) = \exp(-\|x - z\|^2 / \sigma^2) \tag{4}$$

가 되고, $K(x, x) = 1$ 이 된다. 이 관계를 이용하면 (3)은 다음과 같이 간단하게 표현된다.

$$\min_{\alpha} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \tag{5}$$

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

입력 공간에서 라그랑지 함수의 최적해를 만족시키는 안장점의 조건이 $a = \sum_{i=1}^N \alpha_i x_i$ 이 되고, 특징공간에서는 다음 식과 같이 된다.

$$a_F = \sum_{i=1}^N \alpha_i \phi(x_i) \tag{6}$$

판단 경계면(decision boundary) 위에 있는 각각의 서포트 벡터 x_i 와 구 중심 a_F 간의 거리는 구의 반경이 되기 때문에 다음과 같은 식이 성립된다.

$$R_F^2 - \|\phi(x_i) - a_F\|^2 = \tag{7}$$

$$R_F^2 - \left(1 - 2 \sum_{i=1}^N \alpha_i K(x_i, x) + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \right) = 0$$

그리고 위 식을 이용해서 특징 공간상에서 테스트 데이터가 정상인지 비정상인지 판별하는 식은 다음과 같이 구할 수 있다.

$$\begin{aligned}
 f_F(x) &= R_F^2 - ||\phi(x) - a_F|| \quad (8) \\
 &= R_F^2 - \left(1 - 2 \sum_{i=1}^N \alpha_i K(x_i, x) + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \right) \\
 &\geq 0
 \end{aligned}$$

2.2 SVDD를 이용한 노이즈 제거

본 논문의 목적은 특징 공간 내에서 노이즈로 왜곡된 테스트 입력을 구 모양의 서포트를 이용해서 정상 데이터로 바꾸는 것이다. 바꾸는 방법은 특징 공간 상에서 비정상 데이터를 직접 정상 데이터 영역으로 옮기면 된다. 정상 데이터를 이용해서 학습을 하면 그림 1에 있는 작은 구가 만들어지는데 이 구가 바로 SVDD 구이다.

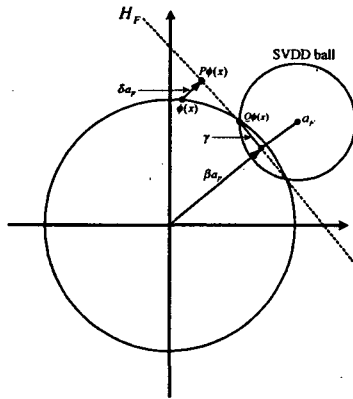


그림 1. 본 논문에서 제안하는 기본 개념

RBF 커널을 사용하면 모든 데이터들이 그림 1에 있는 것과 같이 크기가 1인 단위 구의 표면에 있게 된다. 테스트 데이터 $\phi(x)$ 는 항상 단위 구의 표면에 있지만 노이즈를 가지고 있는 데이터의 경우에는 SVDD 구에는 포함되지 않는다. 노이즈를 제거하기 위해서는 단위 구에 있는 $\phi(x)$ 를 SVDD 구와 만나는 교점인 $Q\phi(x)$ 로 옮기면 된다. 즉, $\phi(x)$ 를 $Q\phi(x)$ 로 옮기는 것이 본 논문의 목적이며 비정상 데이터의 노이즈를 제거하여 정상 데이터로 만드는 과정이다. 위해서는 다음과 같은 간단한 절차를 밟으면 된다. 그림 1처럼 비정상 데이터 $\phi(x)$ 를 초평면 H_F 로 사영시키고 그 결과를 정상 데이터 영역인 $Q\phi(x)$ 로 옮겨주는 과정을 밟으면, $Q\phi(x)$ 가 다음과 같은 형태로 표현될 수 있음을 쉽게 보일 수 있다.

$$Q\phi(x) = \lambda_1 \phi(x) + \lambda_2 a_F \quad (9)$$

단, 여기에서 λ_1 과 λ_2 는 그림 1과 위의 설명을 따라 구해지는 양수이다. 이상의 과정을 밟아 특징공간 내의 denoised pattern $Q\phi(x)$ 가 구해지면, 이 특징벡터의 pre-image를 구함으로써 입력공간 위에서 정의되는 denoised

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 256×256 픽셀 크기의 칼라 이미지를 8비트 흑백 이미지로 변환하고 얼굴의 머리카락을 이미지에서 완전히 제거한 이미지이다. 시물레이션에 사용한 이미지는 모두 200개인데 임의로 100개의 데이터는 학습 데이터로 나머지 100개의 데이터는 테스트 데이터로 사용하였다. 실험에 앞서 각 이미지를 backward warping[4]을 이용해서 원본 이미지를 shape와 texture로 각각 분리해서 따로 실험을 하였다. 만약 이미지를 shape와 texture로 나누지 않고 원본 이미지를 그대로 실험에 사용하면 학습을 할 때, 픽셀 단위로 학습하기 때문에 결과 이미지가 얼굴이 겹쳐서 나오는 것과 같이 흐리게 보인다. 이미지를 분리한 후, 본 논문에서 제안한 방법론을 이용해서 shape와 texture에 각각 있는 노이즈를 제거한 후 forward warping[4]을 이용해서 본 논문에서 제안한 결과 이미지가 나온다.

3.2 시물레이션 절차

본 논문에서 제안한 방법론을 통해서 저해상도의 얼굴 영상을 고해상도의 얼굴 영상으로 복원하는 절차는 다음과 같다. 다음의 절차는 참고문헌 [6]에서 본 논문의 저자 등에 의해 제안된 순서를 따르고 있다:

- ① 학습 데이터로 선택된 N개의 얼굴 이미지를 backward warping을 통해서 N개의 texture 벡터 T_1, T_2, \dots, T_N 와 x, y 방향을 가진 shape 벡터 $S_{1x}, S_{2x}, \dots, S_{Nx}, S_{1y}, S_{2y}, \dots, S_{Ny}$ 2N개를 나눈 후, 각각 학습을 한다.
- ② 테스트 데이터에 해당하는 저해상도의 이미지를 절차 ①처럼 각각 shape와 texture로 나눈 후, bicubic 보간법을 이용해서 고해상도로 각각 변환시킨다.

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

③ 고해상도로 변환된 shape와 texture로, 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[5]를 이용한다.

⑤ 입력 공간으로 옮겨진 shape와 texture를 forward warping을 통해서 이미지로 합성한다.

4. 결과 고찰

이번 시뮬레이션에서 사용한 저해상도 이미지는 32×32 픽셀의 이미지이다. 이 이미지를 고해상도의 이미지 256×256 픽셀의 이미지로 변환하면 그림2와 같이 알아보기 힘든 이미지로 변환된다. 그림 3은 저해상도의 이미지를 shape와 texture로 나누어서 Bicubic 보간법을 사용해서 고해상도로 복원한 후, forward warping을 이용해서 shape와 texture를 합성했을 때의 이미지이다. 그림 2보다는 얼굴 영상을 알아보기 쉽지만 여전히 흐릿하고 불명확하게 보인다. 그림 4는 본 논문에서 제안한 방법론을 이용해서 저해상도의 얼굴 영상을 고해상도의 얼굴 영상으로 복원한 결과이다. 그림 3과 비교했을 때 눈, 코, 입 등이 명확하게 보일 뿐 아니라, 원본 얼굴 영상인 그림 5와 비교했을 때 매우 비슷하다는 것을 알 수 있다. 비록 시뮬레이션의 결과가 좋게 나왔지만 여전히 문제점이 있다. 가장 큰 문제점은 저해상도의 얼굴 영상의 shape와 texture를 각각 구하는 것이다. 본 논문에서는 저해상도 이미지에서 직접적으로 shape와 texture를 구하진 못했다. 대신 고해상도 이미지를 각각 shape과 texture를 나눈 후, 각각 nearest neighborhood 보간법을 이용해서 32×32 로 바꾸어서 사용하는 트릭을 사용했다. 물론 처음 사용한 고해상도의 shape와 texture의 값과 트릭으로 바꾼 저해상도에서 고해상도로 변환된 shape과 texture의 값은 확연히 차이가 있지만 저해상도의 shape과 texture를 구하는 문제점은 여전히 남아있고, 향후에 개선해야 할 문제점이다.



그림 2. 학습에 사용된 얼굴 영상



그림 3. Bicubic 보간법을 이용했을 때 복원된 얼굴 영상



그림 4. 제안한 방법론을 이용했을 때 복원된 얼굴 영상



그림 5. 원본 얼굴 영상

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] S.-W. Lee, J.-S Park and S.-W. 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.

[4] T. Vetter and N.E. Troje, "Separation of texture and shape in images of faces for image coding and synthesis," *Journal of the Optical Society of America A*, vol. 14, pp. 2152-2161, 1997

[5] 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.

[6] J. Park, D. Kang, J. T. Kwok, S.-W. Lee, B.-W. Hwang, and S.-W. Lee, "Facial image reconstruction by SVDD-based pattern de-noising," *Lecture Notes in Computer Science*, vol. 3832, pp. 129-135, 2005.