

**SVDD 기반 노이즈 제거 기법을 이용한 얼굴 영상의 복원**

강대성, 김종호, 박주영  
고려대학교 제어계측공학과

**Reconstruction of Facial Image Utilizing SVDD based Denoising Method**

Daesung Kang, Jongho Kim, Jooyoung Park  
Dept. of Control and Instrumentation Engineering, Korea University

**Abstract** - 많은 경우, 부분 손상된 얼굴 영상을 복원해야 할 필요가 있다. 대표적인 예로는 감시카메라에 찍힌 범인의 얼굴 영상이 이에 속한다. 이런 경우 얼굴의 중요한 부분이 가려져 있기 때문에 자동 얼굴 인식 시스템이나 사람의 관찰로는 그 부분을 인식하기는 매우 어렵다. 이 논문에서는 그 어려움을 극복하기 위해 새롭게 개발된 SVDD기반 노이즈 제거 기법을 부분 손상된 얼굴 영상에 적용한 문제를 고려해 보았다.

**1. 서 론**

얼굴 인식에 대한 많은 연구가 오랫동안 계속되고 있고, 좋은 결과가 많이 나오고 있다. 하지만 아직까지 부분 손상된 얼굴을 인식하는 데에는 어려움이 많다. 예를 들면, 폐쇄회로 TV(CC-TV)화면 등을 통해 얼굴이 드러났더라도 모자, 선글라스, 복면 등으로 얼굴의 일부만 노출될 경우 신원 확인이 어렵다. 이 어려움을 해결하기 위해서 본 논문에서는 최근 지능 시스템 분야에서 많은 연구가 진행되고 one-class 서포트 벡터 학습 방법[1][2] 중 하나인 SVDD(support vector data description)[10]를 이용하여 부분적으로 가려진 얼굴을 복원하려고 한다. 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 데이터들의 노이즈를 제거하려고 한다.

본 논문의 구성은 다음과 같다. 본문에서는 SVDD에 대해서 간단히 기술하고, 시뮬레이션 결과를 보이고 분석한다. 마지막으로 결론에서는 결론과 향후과제에 대해서 논한다.

**2. 본 론**

**2.1 Support Vector Data Description**

정상 영역에 속하는 물체의 서포트를 근사하는 SVDD 방법은 다음과 같이 구해진다.[5],[6] : 중심이  $a \in R^d$ 이고, 반경이  $R$ , 그리고 데이터  $x_i \in R^d, i = 1, \dots, N$ 를 구성하는 학습 데이터 집합  $D$ 를 가지는 구  $B$ 가 있다고 하자. 학습 데이터들은 노이즈를 타기 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 작아야 하고, 또 다른 하나는 가능한 많은 학습 데이터들 포함해야 한다. 이러한 목적을 만족하는 구는 다음의 최적 문제를 풀면 얻을 수 있다.

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

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

위의 식에서, 변수  $\xi_i$ 는 구 밖으로 벗어난  $i$ 번째 학습 데이터와 관련된 벌점이다. (1)의 목적함수는 두개의 상

충되는 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 입력 공간  $R^d$ 이 아닌 특징 공간  $F$ 에서 문제를 다루려고 한다. 특징 공간에서 가능한 작고 특징 공간으로 옮겨진 학습 데이터  $D_F = \{\phi(x_i) | i = 1, \dots, N\} \subset F$ 를 가능한 많이 포함하는 구  $B_F$ 를 찾는 것은 다음의 QP(Quadratic Problem)문제를 풀면 쉽게 구할 수 있다.

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

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

특징 공간에서의 구를 이용하는 SVDD를 위한 QP문제는 다음과 같이 얻을 수 있다.

$$\min_a \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(x_i, x_j) - \sum_{i=1}^N a_i K(x_i, x_i) \quad (3)$$

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

만약 가우스 함수를 커널로 사용하게 되면 커널  $K$ 는

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

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

$$\min_a \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(x_i, x_j) \quad (5)$$

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

입력 공간에서 라그랑지 함수의 최적해를 만족시키는 안장점(saddle point)의 조건이  $a = \sum_{i=1}^N a_i x_i$ 이 되고, 특징 공간에서는 다음 식과 같이 된다.

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

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.

$$\begin{aligned} R_F^2 - \|\phi(x_i) - a_F\|^2 \\ = R_F^2 - \left(1 - 2 \sum_{j=1}^N a_j K(x_i, x_j) + \sum_{j=1}^N \sum_{l=1}^N a_j a_l K(x_j, x_l)\right) \end{aligned} \quad (7)$$

그리고 위 식을 이용해서 특징 공간에서 주어진 테스트 데이터가 정상인지 비정상 데이터인지 판별하는 식은 다음과 같이 구할 수 있다.

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

## 2.2 SVDD를 이용한 노이즈 제거

본 논문의 목적은 특징 공간 내에서 노이즈로 왜곡된 테스트 입력을 구 모양의 서포트를 이용해서 정상 데이터로 바꾸는 것이다. (2)에서 적당한 trade-off 상수  $C$ 를 택함으로써 노이즈가 없는 정상 데이터가 있는 영역을 찾을 수 있다. 만약 정상인 테스트 데이터에 노이즈가 섞였을 경우, SVDD는 그 데이터가 비정상이라고 판단한다. 특징 공간 내에서, SVDD의 판단 영역이 구 형태이다. 구 밖에 있는 노이즈가 섞인 테스트 입력  $x$ 를 구  $B_F$ 의 중심  $a_F$ 를 향해서 판단 경계면까지 이동시키는 것은 쉽다. 물론 구 표면에 수직으로 이동된 데이터  $P\phi(x)$ 는 원래 패턴의 중요한 정보는 그대로 가지고 있다. 위에서 간단히 설명한 내용을 다음과 같은 4단계를 통해서 수식으로 구현할 수 있다. 첫 번째 단계는 주어진 학습 데이터  $D_F = \{\phi(x_i) | i = 1, \dots, N\} \subset F$ 에 대해서 SVDD를 푼다. 그 결과로 (6), (7)을 통해서 최적의  $a_i, R_F^2, a_F$  등을 구할 수 있다. 두 번째 단계는 (8)을 이용해서 각각의 테스트 입력  $x$ 가 정상 영역에 있는지 비정상 영역에 있는지 판별하는 것이다. 비정상 영역에 있을 경우, 입력의 특징 벡터  $\phi(x)$ 를 구 중심  $a_F$ 를 향해 구 표면까지 이동시킨다. 그 결과로 얻어진  $P\phi(x)$ 는 다음과 같다.

$$P\phi(x) = a_F + \frac{R_F}{\|\phi(x) - a_F\|} (\phi(x) - a_F) \quad (9)$$

이러한 이동은 일종의 프로젝션이고 특징공간에서 노이즈를 제거하는 것으로 해석할 수 있다. 프로젝션의 결과로 다음과 같은 식을 얻을 수 있다.

$$\|P\phi(x) - a_F\| = R_F \quad (10)$$

식 (9)에서  $\frac{R_F}{\|\phi(x) - a_F\|}$ 를  $\lambda$ 로 정의하면 (9)를 다음과 같이 간단히 쓸 수 있다.

$$P\phi(x) = \lambda\phi(x) + (1-\lambda)a_F \quad (11)$$

그리고  $\lambda$ 는 다음과 같이 계산할 수 있다.

$$\lambda^2 = \frac{R_F^2}{\|\phi(x) - a_F\|^2} \quad (12)$$

$$= \frac{R_F^2}{(1 - 2\sum_i a_i K(x_i, x) + \sum_i \sum_j a_i a_j K(x_i, x_j))}$$

세 번째 단계와 마지막 단계는 노이즈가 제거된  $P\phi(x)$ 의 pre-image를 구하는 것이다. 만약 특징 공간에서 입력 공간으로 가는 역 매핑  $\phi^{-1}: F \rightarrow R^d$ 가 존재한다면  $\hat{x} = \phi^{-1}(P\phi(x))$  통해서 노이즈가 제거된 패턴을 구하는 것은 쉬울 것이다. 하지만, 대개 정확한 pre-image는 존재하지 않는다[3]. 그래서 우리는 근사해를 찾아야 한다. 근사해를 찾기 위해서 입력 공간의 거리와 특징 공간의 거리간의 간단한 관계식을 사용한 [4]와 MDS(multi-dimensional scaling)[7]의 전략을 따를 것이다. 커널 트릭  $\langle \phi(y), \phi(z) \rangle = K(y, z)$ 와 (11)을 사용하면,  $\|P\phi(x)\|^2$ 과  $\langle P\phi(x), \phi(x) \rangle$ 을 다음과 같이 쉽게 계산할 수 있다.

$$\|P\phi(x)\|^2 = \lambda^2 + 2\lambda(1-\lambda) \sum_{i=1}^N a_i K(x_i, x) \quad (13)$$

$$+ (1-\lambda)^2 \sum_{i=1}^N \sum_{j=1}^N a_i a_j K(x_i, x_j)$$

$$\langle P\phi(x), \phi(x) \rangle = \lambda K(x_i, x) \quad (14)$$

$$+ (1-\lambda) \sum_{j=1}^N a_j K(x_i, x_j)$$

결국 특징 공간에서의  $P\phi(x)$ 와  $\phi(x_i)$ 사이의 거리는 (13)과 (14)를 통해서 다음과 같이 구할 수 있다.

$$\begin{aligned} \overline{d}^2(P\phi(x), \phi(x_i)) &= \|P\phi(x) - \phi(x_i)\|^2 \\ &= \|P\phi(x)\|^2 - 2\langle P\phi(x), \phi(x_i) \rangle + 1 \end{aligned} \quad (15)$$

특징 공간에서의 두 점 사이의 거리의 제곱은 가우스 커널을 사용했을 경우, 다음과 같이 표현할 수 있다.

$$\begin{aligned} \overline{d}^2(\phi(x_i), \phi(x_j)) &= \|\phi(x_i) - \phi(x_j)\|^2 \\ &= 2 - 2K(x_i, x_j) \\ &= 2 - 2\exp(-\|x_i - x_j\|^2 / \sigma^2) \\ &= 2 - 2\exp(-d^2(x_i, x_j) / \sigma^2) \end{aligned} \quad (16)$$

여기서  $d^2(x_i, x_j)$ 은 입력 공간에서의 두 점간의 거리의 제곱을 의미한다. 각 학습 패턴  $x_i$ 에 대한 특징 공간에서의 거리는 (15)에서 얻을 수 있기 때문에 입력 공간의 거리에 상응하는  $P\phi(x)$ 의 근사된 pre-image  $\hat{x}$ 를 구할 수 있다. 일반적으로 어떤 점의 위치를 결정할 때 가장 큰 역할을 하는 것이 이웃과의 거리이다. 여기서  $P\phi(x)$ 와 그와 가장 가까운  $n$ 개의 이웃들 간의 거리의 제곱은 다음과 같이 정의 된다.

$$d^2 = [d_1^2, d_2^2, \dots, d_n^2]^T \quad (17)$$

여기서  $d_i$ 는  $P\phi(x)$ 의 pre-image와  $x_{(i)}$ 간의 입력 공간에서의 거리를 나타낸다.  $P\phi(x)$ 를 입력 공간으로 되돌리기 위해서는 MDS[7] 개념을 사용한다. 이를 위해 먼저 학습 데이터  $X = [x_1, x_2, \dots, x_n]$ 의 중심을 얻

기 위해서 평균  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 을 취한 후,  $d \times n$  행렬  $X = [x_1, x_2, \dots, x_n]$ 을 만든다. 그리고  $n \times n$  centering 행렬을 다음 식과 같이 정의한다.

$$H = I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \quad (18)$$

여기서  $I_n = \text{diag}[1, \dots, 1] \in R^{n \times n}$ 인 행렬이고  $\mathbf{1}_n = [1, \dots, 1]^T \in R^{n \times 1}$ 인 벡터이다. 행렬  $XH$ 는  $x_i$  벡터들을 그들의 중심으로 평행이동 시킨다. 즉  $XH = [x_1 - \bar{x}, \dots, x_n - \bar{x}]$ 이 된다. 그리고  $XH$ 의 열 공간 내에서 좌표계를 정의한다. 만약  $XH$ 의 rank가  $q$ 라면, 우리는  $d \times n$ 행렬  $XH$ 의 SVD (singular value decomposition) [8]을 다음과 같이 구할 수 있다.

$$\begin{aligned} XH &= [U_1 U_2] \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\ &= U_1 \Lambda V_1^T = UZ \end{aligned} \quad (19)$$

$U_1 = [e_1, \dots, e_q]$ 은 orthonormal 열벡터  $e_i$ 로 이루어진  $d \times q$ 행렬이고,  $Z = \Lambda V_1^T = [z_1, \dots, z_n]$ 는  $e_i$  위로  $x_i - \bar{x}$ 을 투영한 열벡터  $z_i$ 로 이루어진  $q \times n$  행렬이다.  $x_i$ 와  $z_i$ 는 다음과 같은 식을 만족한다.

$$\|x_i - \bar{x}\|^2 = \|z_i\|^2, \quad i = 1, \dots, n \quad (20)$$

이를  $n$ 차 벡터까지 집합시키면 다음과 같다.

$$d_0^2 = [\|z_1\|^2, \dots, \|z_n\|^2]^T$$

앞에서 언급한대로 pre-image  $\hat{x}$ 의 위치는  $d^2(\hat{x}, x_i)$ 를 계산해서 얻을 수 있다.

$$d^2(\hat{x}, x_i) \approx d_i^2, \quad i = 1, \dots, n$$

이상적인 경우, 이들 간의 거리는 반드시 일정해야 한다. 하지만, 대개 입력 공간에서 정확한 pre-image가 존재하지 않고, 심지어는 이들 간의 거리조건을 만족하는 해답이 존재하지 않을 수도 있다. 그렇기 때문에  $\hat{x}$ 를 구하기 위해서는 최소자승법(Least Squares)을 이용해야 한다. 최소자승법을 이용해서  $\hat{z}$ 를 구하면 다음과 같다.

$$\hat{z} = -\frac{1}{2} \Lambda^{-1} V_1^T (d^2 - d_0^2) \quad (21)$$

$\hat{z}$ 와  $\hat{x}$ 는 다음과 같은 관계를 가지고 있다.

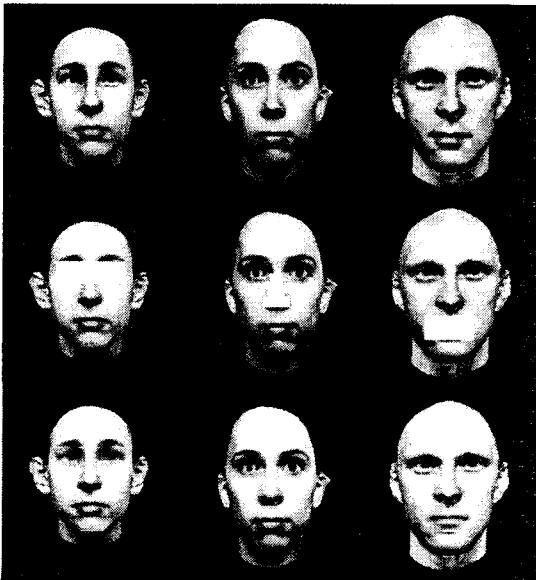
$$\hat{x} - \bar{x} = U_1 \hat{z} \quad (22)$$

결국 노이즈가 제거된 패턴의 위치는 다음과 같이 구할 수 있다.

$$\hat{x} = U_1 \hat{z} + \bar{x} \quad (23)$$

### 2.3 시뮬레이션 결과 및 분석

이번 시뮬레이션에 사용한 얼굴 영상<sup>1)</sup>들은 모두 200개의 2차원 백인 얼굴 이미지를 사용하였는데 이 이미지는 레이저 스캐너(*Cyberware<sup>TM</sup>*)로 기록되어진 3차원 얼굴 모델의 데이터베이스(database)에서 얻어진 것이다. 이 이미지들은 256×256 픽셀 크기의 칼라 이미지이다. 그리고 얼굴의 머리카락을 이미지에서 완전하게 제거되었다. 시뮬레이션에서는 칼라 이미지를 모두 8비트 흑백 이미지로 변환하였다. 시뮬레이션에 사용한 이미지는 모두 200개인데 임의로 99개의 데이터를 고른 후, 학습하는데 사용하였다. 그리고 테스트 입력은 나머지 101개의 데이터 중 하나를 골라서 인위적으로 눈, 코, 입 부위를 제거한 이미지를 사용하였다. 아래에 있는 얼굴 영상이 시뮬레이션 결과인데, 이 결과에 대해 설명을 하면 먼저 첫 번째 행에 있는 그림들은 각 시뮬레이션에 사용한 얼굴 영상들의 원본을 보여준 것이고, 두 번째 행에 있는 그림들은 각 원본 얼굴 영상의 눈, 코, 입 부위를 인위적으로 제거한 이미지이다. 그리고 마지막 행에 있는 그림들은 SVDD기반 노이즈 제거 기법을 이용해서 노이즈를 제거한 뒤, 복원한 결과 이미지들이다. 원래 결과를 복원하게 되면 이미지가 대체적으로 흐리게 나온다. 그래서 그 이미지를 사용하지 않고, 복원된 결과 이미지 중에서 가려진 부분에 대응하는 부분 이미지를 두 번째 행에 있는 그림의 가려진 부분 이미지에 대체시켰다. 가려진 부분이 대체적으로 원본과 비슷하게 복원되었지만, 시뮬레이션 도중 그렇지 않은 얼굴도 복원되는 경우도 있었고, 흐리게 나온 부분도 있었다. 하지만 복원된 얼굴 영상이 대체적으로 자연스럽게 나왔다.



1) 논문에서 사용한 얼굴 영상은 'Max Plank Institute for Biological Cybernetics' 에서 제공받았다

### 3. 결 론

본 논문에서는 부분적으로 손상된 얼굴 영상을 복원하는 기법에 대해서 소개했다. 만족스러운 결과가 나왔던 반면 문제점 역시 있었다. 현재 이 기법에서는 각각의 얼굴 이미지를 하나의 패턴 벡터로 학습하였는데, 이렇게 하면 앞에서 언급한대로 복원되는 부분의 이미지는 흐릿하게 나온다. 그 이유는 학습을 할 때, 픽셀 단위로 학습을 하였기 때문이다. 픽셀 단위로 학습했을 때 나쁜 특성이 나타나는 이유는 선형 공간의 중요한 특징 때문이다. 선형 공간의 가장 중요한 특징은 선형 결합을 정의하는 덧셈 법칙과 곱셈 법칙이 성립한다는 것이다. 간단한 예로 서로 다른 두 사람의 얼굴에 대해서 평균을 취하면, 나오는 얼굴은 두 얼굴을 겹쳐 놓은 것처럼 보인다. 비슷한 이유로, 본 논문에서 사용한 얼굴의 크기와 눈, 코, 입, 귀 등의 위치와 크기가 모두 다른 얼굴 영상에 대해 픽셀 단위로 학습하였기 때문에 결과 이미지가 흐리게 나온다. 이 문제점을 수정하기 위해서 향후에 얼굴 영상을 shape와 texture로 각각 나누어서 SVDD를 기반으로 한 노이즈 제거 기법을 적용하려 한다. 그리고 실험에 사용한 얼굴이 모두 정면을 바로보고 있는데 위나 아래, 측면 등을 바라보는 얼굴에 대해서도 연구하려 한다.

### [참 고 문 헌]

- [1] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods", Cambridge, U.K.: Cambridge University Press, 2000.
- [2] B. Schölkopf and A. J. Smola, "Learning with Kernels", Cambridge, MA: MIT Press, 2002
- [3] S. Mika, B. Schölkopf, A. Smola, K. R. Müller, M. Scholz, and G. Rätsch, "Kernel PCA and de-noising in feature space", *Advances in Neural Information Processing Systems*, vol. 11, pp. 536-542, Cambridge, MA: MIT Press, 1999
- [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] D. Tax and R. Duin, "Support Vector Domain Description," *Pattern Recognition Letters*, vol. 20, pp.1191-1199, 1999
- [6] D. Tax, *One-Class Classification*, Ph.D. Thesis, Delft University of Technology, 2001
- [7] T. F. Cox and M. A. A. Cox, "Multidimensional Scaling", *Monographs on Statistics and Applied Probability*, vol. 88, 2nd Ed., London, U. K.: Chapman & Hall, 2001.
- [8] T. K. Moon and W. C. Stirling, "Mathematical Methods and Algorithms for Signal Processing", Upper Saddle River, NJ: Prentice Hall, 2000
- [9] B. W. Hwang and S. W. Lee, "Reconstruction of partially damaged face images based on a morphable face model" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(3) :365-372, March 2003
- [10] J Park, D. Kang, J. Kim, J. T. Kwok, and I. W. Tsang, "Pattern de-noising based on support vector data description", To appear in the *Proceedings of International Joint Conference on Neural Networks 2005*.
- [11] <http://faces.kyb.tuebingen.mpg.de/>