

# 지지 벡터 데이터 기술을 이용한 가려진 얼굴 요소 복원

## (Reconstructing Occluded Facial Components using Support Vector Data Description)

김 경 호 <sup>†</sup>                      정 윤 수 <sup>\*\*</sup>  
(Kyoung-Ho Kim)            (Yun-Su Chung)

이 상 웅 <sup>\*\*\*</sup>  
(Sang-Woong Lee)

**요 약** 얼굴 인식 분야는 오래전부터 꾸준히 연구되어 왔지만, 아직도 실용적인 얼굴 인식은 이루어지지 않고 있다. 이는 실제 얼굴 인식 시스템의 입력 영상의 경우, 실험실에서 획득된 얼굴 영상과는 달리 안경이나 스카프, 헤어스타일 등에 의해서 가려진 얼굴 영상인 경우에 인식 성능이 매우 저하되는 것에 기인한다. 이러한 비 얼굴 요소를 처리하기 위해, 최근 수년간 다양한 방식의 비 얼굴 요소 처리 방법이 있었으나, 만족할만한 성능을 보이지 못했다. 본 논문에서는, 최근 관련 방법 중에서 특징 공간에서 최소 거리의 볼을 찾아 근사값을 추정하는 방식인 SVDD를 이용하는 비 얼굴 요소 복원 방법을 제안하고, 실험을 통해 성능을 평가한다. 제안 방법의 실효성을 검증하기 위해, 비 얼굴 요소 부분을 점진적으로 증가시켜 복원하는 실험 등

을 통해 실험한 결과, 제안 방법은 상당한 수준의 실효성을 지니고 있음을 확인하였다.

키워드 : 얼굴 복원, SVDD, 가려진 얼굴

**Abstract** Even though face recognition researches have been developed for a long ago, there is no practical face recognition system in real life. It is caused by several real situations where non-facial components such as glasses, scarf, and hair occlude facial components while facial images in a face database are well designed. This occlusion decreases recognition performance. Previous approaches in recent years have tried to solve non-facial components but have not resulted in enough performance. In this paper, we propose a method to handle this problem based on support vector data description, which trains the hyperball in feature space to find the minimum distance estimating the approximated face. In order to evaluate its performance and validate the effectiveness of the proposed method, we make several experiments and the results show that the proposed method has a considerable effectiveness.

**Key words** : Face reconstruction, SVDD, occluded face

### 1. 서 론

얼굴 인식 분야는 오래전부터 꾸준히 연구되어 왔다. 하지만 완벽한 얼굴 인식은 아직도 이루어지지 못하고 있다. 그 원인으로는 학습된 영상과 인식할 영상의 시점 차이로 인한 조명, 색감, 또는 촬영장비의 기종 차이에 따른 문제 등 여러 가지가 있다. 2000년대에 이러한 원인들을 해결하기 위하여, 많은 방법들이 제안되었으며, 최근 완전한 얼굴영상이 아닌 헤어스타일, 안경, 스카프 등에 의해 가려진 비 얼굴 요소들을 처리하는 실용적인 얼굴 인식 기법들에 대한 관심이 고조되고 있다.

현재까지 비 얼굴 요소들을 처리하는 기술은 크게 두 가지 형태의 접근방법을 취하고 있다. 첫 번째는 비 얼굴 요소를 복원해 완성된 형태의 얼굴을 취하여 인식을 시도하는 방법이며, 다른 하나는 비 얼굴 요소를 제외하고 얼굴 인식을 시도하는 방법이다. 방법론상의 용이함에 의하여 비 얼굴 요소를 제외하고 얼굴 인식을 시도하는 방법들이 과거로부터 많이 이루어졌으며, 복원방식은 최근 들어 수학적 통계/확률 방식에 기반하여 이루어지고 있다. 복원 방식 기법은 Hwang과 Lee에 의하여 시도되고 발전되어 왔다. 이 방법은 얼굴을 형태와 질감의 프로토타입들의 선형중첩으로 모형화한 다음, 특징점에서의 형태와 질감정보만을 가지고 얼굴이 요구하는 변형의 최적화된 근사값을 찾아 복원하였다[1]. 또한 주어진 학습데이터를 새로운 선형 공간으로 이동하여 분석하는 LLE(Locally Linear Embedding) 기법을 기

· 본 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구(KRF-2006-214-D00144)와 한국전자통신연구원에서 수행중인 "CCTV 환경에서의 비계약적 얼굴 인식 및 검색 기술 개발(과제번호 : 09BS12 00)" 과제에서 도출된 연구임  
· 이 논문은 제36회 추계학술발표회에서 'SVDD를 이용한 비 얼굴 요소 복원의 재목으로 발표된 논문을 확장한 것임

<sup>†</sup> 학생회원 : 조선대학교 컴퓨터공학부  
lovekgh1@nate.com  
<sup>\*\*</sup> 비 회원 : 한국전자통신연구원 책임연구원  
yoonsu@etri.re.kr  
<sup>\*\*\*</sup> 종신회원 : 조선대학교 컴퓨터공학부 교수  
swlee@chosun.ac.kr  
(Corresponding author)  
논문접수 : 2009년 12월 24일  
심사완료 : 2010년 1월 26일

Copyright©2010 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제16권 제4호(2010.4)

반으로 다층 퍼셉트론을 학습시키는 방법을 사용한 기법이 제안되었다[2]. 그 외에도 3D Morphable Model과 구성요소별 SVM(Support Vector Machines)을 결합하여 얼굴 인식 시스템에 적용한 방법[3]과 미리 저장된 얼굴 공간 중에 입력된 사용자의 얼굴과 가장 비슷한 얼굴 공간을 사용하여 가려진 부분을 판별해내고 원래의 얼굴을 복원하는 방법[4]들도 보고되었다.

### 2. 지지 벡터 데이터 기술

SVDD(Support Vector Data Description)는 학습 데이터의 존재 영역을 볼(Ball)을 이용하여 직접 근사하는 방법을 사용한다. SVDD는 특징 공간에서 최대한 많은 학습 데이터를 포함하는 최소 거리의 볼을 찾는 것을 목적으로 한다. 중심이  $a$ 이며, 반지름이  $R$ 인 볼의 크기는  $\xi$ (slack variable)에 의해 조절된다. 최적의 볼 크기를 구하기 위하여 다음 식이 사용된다.

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

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

여기서, 상수  $C$ (trade-off constant)는 각 항의 중요도를 조절한다. 일반적으로 입력 공간보다 주어진 입력 영상을 더 잘 표현할 수 있는 특징 공간으로 변환한다면 SVDD는 더욱 좋은 성능을 발휘할 수 있다. 이를 위해 커널을 적용하고, 얻어진 2차 프로그래밍을 풀기 위해 라그랑지 방법을 적용하여, 새로운 라그랑지 계수 알파에 대하여 식을 전개하면 식 (2)를 얻게 된다. 이 식을 풀어야 하는 우리는 원하는 구의 반지름  $R$ 을 얻게 된다[5].

$$\min_a \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i K(x_i, x_i) \quad (2)$$

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

### 3. 비 얼굴 요소 복원 방법

SVDD 학습에 의해 특징 공간에 옮겨진 얼굴 데이터들은 단위구의 표면에 위치하게 된다. 위에서 말한 최소 거리인 특징 공간에 옮겨진 학습된 데이터와 이 모든 것으로 포함하는 볼(Ball)과의 거리를 나타낸다.

입력된 영상의 주변에 근접해 있는 학습 데이터를 이용하여, 유추하고 입력공간으로 재 이동하면, 입력영상의 가려진 부분의 주변 환경과 유사한 환경을 가진 볼 내부의 사람의 영상을 얻을 수 있다. 이 실험에서는 특징 공간상의 볼 표면에 사영된 벡터를 유추하고, 입력 공간에 재 이동하는 경우에 생성되는 pre-image를 J. T. Kwok에 의해 제안된 방법[6]으로 생성하였다. 이 과정은 그림 1에 표현되어 있다.

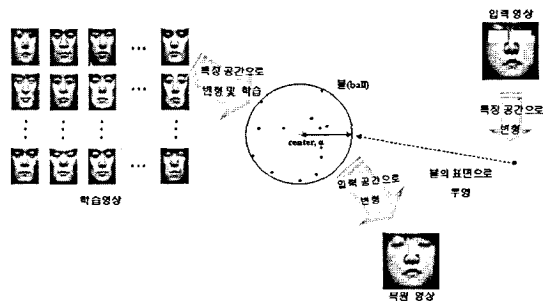


그림 1 제안 방법의 개념도

이러한 직관적인 개념을 이용한 SVDD 응용분야는 과거 저해상도 영상을 고해상도로 복원[7]하거나, 잡음에 의하여 저하된 화질을 가진 얼굴을 처리[8]하는 등의 전체적인 데이터 왜곡을 처리하는데 주로 사용되었으나, 비 얼굴 요소에 의해 가려진 일부분을 처리하는 문제에서는 일부분이 전체를 변화시켜 오히려 복원된 결과가 나빠지는 경향을 보였다[8]. 따라서 본 논문에서는 Hwang에 의해 사용되었던, 비 얼굴 요소 복원을 위해 가려진 얼굴 영역의 위치를 알고 있다고 가정하고, SVDD에 의한 복원된 전체 영상에서 가려진 영역만을 추출해 입력 영상에 정합하는 방법을 사용하여, 복원의 에러를 최소화하였다. 본 논문에서는 이러한 방법을 통해, SVDD 기반 방법이 비 얼굴 요소를 어느 정도 처리할 수 있는지를 밝히고자 한다.

### 4. 실험 방법

본 실험의 목적은 실험을 통해 SVDD를 이용한 비 얼굴 요소 복원 방법의 성능을 평가하기 위한 것이다. 따라서 성능을 평가할 수 있는 실험이 진행되며, 결과로는 객관적인 평가가 가능한 자료가 도출되어야 한다.

이를 위해 실험은 다음과 같은 실험 방법을 기본으로 수행하였다. 먼저 복원을 위한 학습영상은 200명의 정면 얼굴 영상이 담겨있는 640×480 칼라 영상에서 얼굴 영역을 추출한 후, 128×128 회색조 영상으로 정규화한 영상 200장을 사용하여 구성하였다. 그리고 200장의 영상 중 실험의 성능을 평가하기 위해 그림 2에 제시된 테스트 영상 10장을 실험에 사용하였다.

본 실험에서는 영상에 일정 비율로 흰색 사각형의 마스크를 테스트 영상에 씌우는 방법으로 비 얼굴 요소가 포함된 입력 얼굴 영상을 임의로 생성하였다. 가려진 정도에 따른 성능 평가를 위해 마스크의 크기가 전체 영상 크기의 10%가 되도록 설정하고 얼굴 구성 요소 중



그림 2 10명의 테스트 영상

주요 부분인 양 눈, 코, 입 등을 가려, 10%, 20%, 30%, 40%, 50%가 가려진 입력 영상을 생성하였다. 그림 3에서 도시한 것처럼 10% 가려짐에서는 영상 내 얼굴의 오른쪽 눈에 흰색 사각형을 씌웠다. 이때 흰색 사각형의 위치는 모든 테스트 영상의 동일 한 위치에 씌운 것이 아니라 테스트 영상에서 오른쪽 눈의 실제 위치에 씌워졌다. 때문에 이미지마다 흰색 사각형의 위치가 달라진다. 20% 가려짐에서는 양쪽 눈에 10%의 흰색 사각형을 씌웠고, 30% 가려짐에서 양쪽 눈, 그리고 코에 흰색 사각형을 씌웠다. 그리고 40% 가려짐은 양쪽 눈, 한쪽 입에 흰색 사각형을 씌웠다. 이때 입의 가려진 영역은 20%이다. 마지막으로 50% 가려짐은 양쪽 눈, 코, 입에 흰색 사각형을 씌웠다. 위에서 설명한 각 단계별 입력 영상은 아래의 그림 3과 같다.

본 실험의 성능을 평가하기 위한 기준인 오차 값은 식 (3)을 통해 구해졌다.

$$\text{오차} = \left( \sum_{i=0}^{\text{size}} | \text{Img}_{\text{input}} - \text{Img}_{\text{result}} | \right) / \text{size} \quad (3)$$

여기서, size란 이미지의 크기를 의미하고, input은 가려진 입력영상, result는 SVDD를 통해 복원된 영상을 의미한다. 이 수식을 통하여 구해진 오차 값을 가지고 성능을 평가하였다.

지금까지 입력 얼굴 영상의 비 얼굴 요소 생성에 대하여 설명하였다. 그러나 마스크 비율의 증가에 대한 실험만 가지고는 SVDD를 이용한 비 얼굴 요소 복원에 대한 평가가 이루어지기는 미흡하다. 따라서 본 논문에서는 비 얼굴 요소 증가와 함께 3가지 방법을 조합하여 실험하였다. 각 3가지 방법에 대한 실험방법은 다음과 같다.

SVDD 비 얼굴 요소 전체복원 실험(이하 전체복원 실험)은 200장의 학습영상을 가지며, 10장의 영상으로 테스트한다. 10%~50%의 가려진 부분에 대한 실험이 이루어진다.

복원 부분 정합 실험(이하 부분정합 실험)은 전체복원

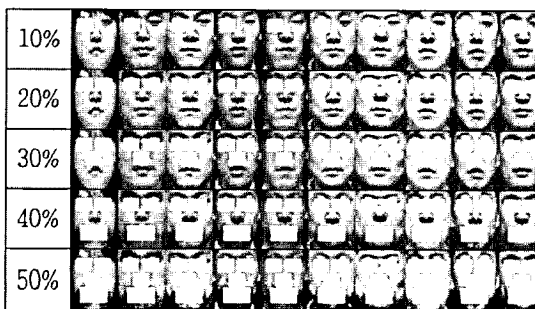


그림 3 가려진 영역의 비율별 입력 영상

실험과 같은 조건으로 이루어진다. 다만 이 실험에서는 SVDD방법이 주변 픽셀에 영향을 주기 때문에 그 영향을 최소화하기 위해 입력영상에서 가려진 부분만 복원 영상에서 추출해, 다시 입력영상에 정합하는 방법을 적용하는 실험이 이루어진다.

앞서 전체복원 실험과 부분정합 실험은 학습 데이터에 테스트 영상이 포함되어 있다. 실제 환경에서 모든 입력 영상이 데이터베이스에 항상 존재한다고 보기는 힘들다. 그러므로 존재하지 않을 때의 성능을 테스트하기 위해 입력영상을 학습영상에서 제외한 Leave-one-out 교차 검증 방식의 실험을 진행한다. 이 실험은 전체 복원과 복원부분 정합실험의 정의를 그대로 사용하되, 학습데이터에서 테스트 영상을 제외한다.

본 논문을 통해서 SVDD를 이용한 비 얼굴 요소 복원 방법의 신빙성 있는 결과를 제시하기 위해, 위에서 정의한 3가지의 실험방법에 대한 실험을 하고 그에 따른 결과분석이 이루어진다.

### 5. 실험 결과 분석

앞서 세 가지의 방법을 통한 실험결과를 토대로 그에 대한 성능 분석을 진행하도록 한다.

전체복원 실험과 부분정합 실험은 상이한 실험이나 성능의 비교가 용이함으로 함께 분석한다. 입력영상을 전체복원 실험의 방법을 통하여 복원하면 그림 4와 같은 결과를 얻을 수 있다. 이 결과를 분석해보면 어느 정도의 성능을 가지고 복원이 되긴 하지만, 각 비율당 오차의 최대값과 최소값을 보면 전체 입력영상이 모두 균일하게 복원되지 않고 편차가 심하다는 것을 알 수 있다.

SVDD를 통한 대표영상 복원의 방법 자체가 볼(Ball)이라고 정의된 공간에 데이터를 투영시키고 근사한 주변 데이터들을 가지고 유추·복원하는 방법이다. 따라서 입력한 데이터의 완벽한 복원결과인 원본영상과 완전히 일치하기는 힘들다. 또한 이 복원 방법은 주변 픽셀에도

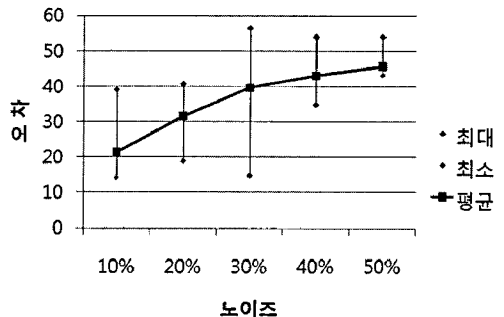


그림 4 가려진 비율별 전체복원 실험 결과

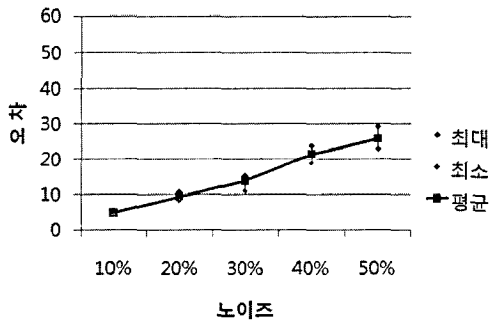


그림 5 가려진 부분만을 적용한 실험 결과

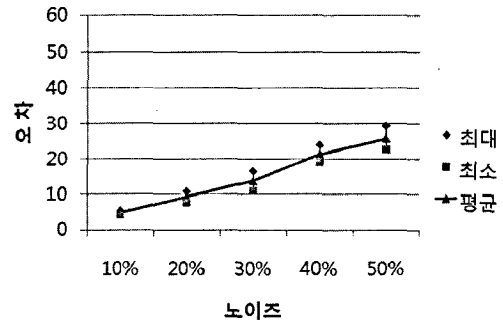


그림 7 Leave-one-out 적용 실험 결과

영향을 미치지 때문에 육안으로 관찰하기에도 흐릿한 부분이 있다.

그래서 두 번째 실험인 부분정합 실험에서는 전체복원 실험보다 결과를 좋게 하기위해 입력된 영상의 가려진 부분을 복원실험을 통해 얻어진 결과영상에서 추출, 입력영상과 정합하는 방법을 적용했다. 이 실험의 결과는 그림 5에서 볼 수 있다.

실험결과 오차가 현저하게 낮아지고 각 항의 최대와 최소값의 편차가 확연하게 줄어들었다.

가려진 부분만 결과영상에서 얻어내 입력영상에 합성하는 방법은 기본 상태에서도 가려진 부분만이 오차로 인식되므로 10%~50%의 오차율을 가진다. 가려진 부분만을 복원된 영상에서 추출해 정합하게 되면 주변 화소에 영향을 끼치는 오류요인을 최소화 할 수 있다. 두 실험의 결과 영상의 비교는 그림 6에 할 수 있다.

실험	원본	10%	20%	30%	40%	50%
전체 복원						
부분 정합						

그림 6 전체복원, 부분정합 실험 결과 영상

위 두 가지 실험은 입력영상이 학습영상에 포함되어 있다. 현실 세계에서 모든 사람에 대한 모든 환경의 영상을 학습데이터로 가지고 있기는 불가능한 일이다. 따라서 이 부분의 성능을 테스트하기 위해 제외실험을 수행한다. 이 결과는 그림 7과 그림 8에서 확인할 수 있다.

제외실험 결과 입력영상이 테스트 영상에 포함되어 있는 부분정합 실험과 거의 유사한 결과를 얻을 수 있다. 이 실험을 통해 SVDD를 이용한 비 얼굴 요소 방법은 학습데이터의 포함 여부와 복원율의 상관관계가 크지 않다는 것을 확인할 수 있다.

실험	원본	10%	20%	30%	40%	50%
전체 복원						
부분 정합						

그림 8 Leave-one-out 적용 실험 결과 영상

### 6. 결론 및 향후 계획

실험결과 위의 방법은 30% 이하의 가려짐에서는 전반적으로 좋은 성능을 보였지만 그 이상의 가려짐과 정규화 되지 않은 입력영상(얼굴이 크거나, 기울어짐)은 실험 시 복원율이 떨어졌다. 그리고 본 실험에서 사용된 가려진영역의 자동 추출 알고리즘의 정확도가 부족해 부분정합실험이 완벽하지 않았다.

향후에는 입력영상 정규화의 전처리과정을 추가하고, 가려진영역 추출의 정확성을 높일 것이다. 또한 실제적 가림에 대한 실험을 진행 할 것이다.

이 방법은 향후 여러 분야에 사용 가능할 것으로 보인다. 예를 들어, 은행의 ATM 서비스에 사용될 수 있다. 편리한 만큼 관련 범죄가 많이 일어나는데 ATM기에는 CCTV 이외에 마땅한 방법 수단이 존재하지 않는다. 그러나 SVDD를 통한 가려진 얼굴 복원 방법을 사용하면 모자, 안경, 마스크 등으로 얼굴을 가려도 범인 추적이 가능할 것으로 기대된다.

### 참고 문헌

[1] 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*, vol.25, no.3, pp.365-372, 2003.

[2] C. Zhang, J. Wang, N. Zhao, and D. Zhang, "Reconstruction and analysis of multi-pose face images based on nonlinear dimensionality reduction,"

- Pattern Recognition*, vol.37, no.2, pp.325-336, 2004.
- [ 3 ] J. Huang, V. Blanz, and B. Heisele, "Face Recognition Using Component-Based SVM Classification and Morphable Models," *Pattern Recognition with Support Vector Machines, Lecture Notes in Computer Science*, vol.2388, pp.334-341, 2002.
  - [ 4 ] L. Goldmann, A. Rama, T. Sikora, and F. Tarres, "On the Detection and Localization of Facial Occlusions and its Use within Different Scenarios," *10th IEEE Workshop on Multimedia signal Processing*, vol.8, no.10, pp.592-597, 2008.
  - [ 5 ] S.-W. Lee, J. Park, and S.-W. Lee, "Synthesis of Face Exemplars using Support Vector Data Description," *Proc. of the KIISE Fall Conference*, vol.32, no.2, pp.835-837, 2005. (in Korean)
  - [ 6 ] J. Park, D. Kang, J. Kim, J. T. Kwok, and I. W. Tsang, "Pattern De-Noiseing Based on Support Vector Data Description," *Proc. of Intl. Joint Conf. on Neural Networks*, Montreal, Canada, pp.49-953, 2005.
  - [ 7 ] S.-W. Lee and S.-W. Lee, "SVDD-based Illumination Compensation for Face Recognition," *Advances in Biometrics, Lecture Notes in Computer Science*, vol.4642, pp.154-162, 2007.
  - [ 8 ] 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," *Advances in Biometrics, Lecture Notes in Computer Science*, vol.3832, pp.129-135, 2005.