

## 다양체 가우시안 분해 : 얼굴 검출

양준영<sup>0</sup> 변혜란  
 연세대학교 컴퓨터 과학과  
 (gundid<sup>0</sup>, hrbyun )@cs.yonsei.ac.kr

### Decompose the Manifold Into Gaussian Densities : Face Detection

Junyeong Yang<sup>0</sup> Hyeran Byun  
 Dept. of Computer Science, Yonsei University

#### 요 약

제안하는 방법은 분산량이 큰 객체에 대하여 여러 개의 가우시안을 이용하여 다양체를 분해하는 알고리즘이다. 제안하는 방법은 단순하지만 빠르게 다양체를 근사시키는 여러 개의 가우시안을 생성한다. 또한, 가우시안 혼합 모델과 유사하나 보다 빠른 연산시간을 보장하며 Outlier에 대한 신뢰성을 향상 시켜준다. 제안하는 알고리즘은 우리가 수집한 다 인종(동양인, 흑인, 백인, 히스패닉) 얼굴 데이터 베이스 QVGA영상에서 100%의 검출률과 0개의 오분류 의 높은 성능을 도출하였다.

#### 1. 서 론

본 논문에서 다양체 가우시안 분해(DMGD:Decompose Manifold Into Gaussian Densities)방법을 제안한다. 다양체 가우시안 분해는 학습 데이터의 전역적인 특징 공간 상에서 학습 데이터를 양분하는 방법을 이용하여 다양체를 여러 개의 가우시안으로 분해하는 역할을 수행한다. 본 연구에서 제안하는 알고리즘의 타당성을 입증하기 위하여 얼굴 검출 문제에 적용하였다. 사실 제안하는 알고리즘은 다름 아니라 얼굴 검출 문제를 해결하기 위해 통계적인 확률모형을 기반으로 한 여러 논문을 토대로 몇 가지 의문점을 제시하는 것부터 시작되었다. 다년간 많은 연구결과에서 얼굴은 가우시안 분포를 이용하여 분석되어 왔다. 하지만 얼굴 검출은 같은 범주 안에 속하는 객체임에도 불구하고 조명이나 자세, 안경 등 기타 여러 요소에 의하여 외관의 변형이 다양하게 일어 날 수 있으므로 확률모형을 이용하여 표현하기가 쉽지 않다. 이와 같은 문제를 해결하기 위하여 [1][2][3]는 몇 개의 군집으로 나누어 객체의 분산량을 줄였으며, [4]는 분류성분이 높은 특징을 이용하여 가우시안 분포를 모형화함으로써 얼굴 객체에 대한 일관성을 높이고 비 얼굴 객체와의 분류 가능성을 높였다. 여기서 다음과 같은 문제를 제시하고자 한다. 1. 분산량이 큰 객체를 하나의 가우시안으로 표현하는 것이 효율적인가? 2. 비 얼굴 객체에 대해서 어떤 가정이 타당한가? 위 문제에 대해 결론적으로 학습영상이 이루는 다양체를 여러 개의 가우시안으로 분해함으로써 분산량을 줄이고, 여러 개의 분류함으로써 비 얼굴 객체

에 대하여 어떤 가설도 세우지 않는 것으로 문제를 해석하였다. 여기서, 제안하는 방법의 기여도를 간략히 언급하고자 한다.

첫째, 다양체 가우시안 분해 방법은 비 얼굴 객체를 분해 작용 요소로만 이용한다.

둘째, 다양체 가우시안 분해 방법은 가우시안 혼합모델과 유사하나 보다 빠른 연산시간을 보장한다.

셋째, 다양체 가우시안 분해 방법은 평균과 유사도가 작은 영상(Outlier)에 대해 보다 강한 신뢰성을 부여할 수 있다.

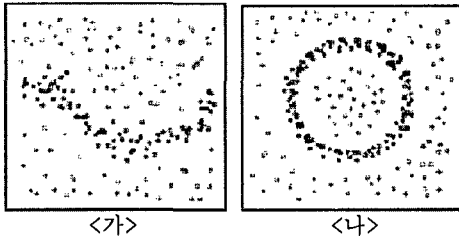
#### 2. 다양체 가우시안 분해

먼저, <그림. 1>에 가, 나 2개의 문제를 조명하는 것부터 시작하려 한다. 다변량 정규분포의 조건부 확률값을 분류식으로 이용하는데 조건부 확률식은

$$p(\text{Face} | x) = \frac{p(\text{Face})p(x | \text{Face})}{p(x)} \quad (1)$$

(1)식과 같고, (1)식에 한계치를 적용한 (2)식을 이용하여 얼굴과 유사한 패턴을 분류하였다.

$$\begin{aligned} D(x) &= \log p(\text{Face} | x) > \alpha \\ &= \log \frac{p(\text{Face})}{p(x)} + \log p(x | \text{Face}) > \alpha \quad (2) \\ &= \log p(x | \text{Face}) > \alpha \end{aligned}$$



<그림. 1> 모의 실험 영상  
 [■: 양의 학습 영상, □: 음의 학습 영상]

<그림.1>의 가, 나 의 양의 학습 영상에 대해 다변량 정규 분포를 모형화하고 (2)식을 이용하여 모든 양의 학습 영상을 포함하는 결정경계를 구했을 때 <그림. 2>와 같은 타원형의 경계가 형성된다. 모의 실험 영상에서 보듯이 양의 학습영상이 선형형태의 분포를 가지고 있지 않은 경우에 하나의 다변량 정규 분포로 확률모형을 예측하는 것은 좋지 못한 선택임을 확인할 수 있다.

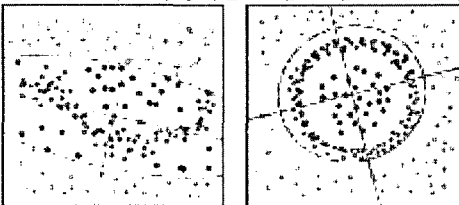
2.1. 분해방법

관측치를 2개 이상의 집단으로 나누기 위하여 전역적인 특징을 고려할 수 있는 주성분 분석(PCA)을 이용하였다. 주성분 분석을 통해 고유벡터  $\Phi$  와 고유치  $\Lambda$  를 구했을 때, 가장 큰 고유치에 대응되는 고유벡터를  $\phi_{highest}$  라 놓으면, 제안하는 분해 방법은 (3)식과 같다.

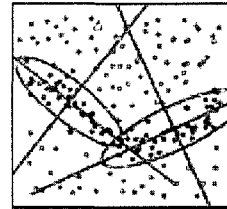
$$f(x) = \begin{cases} 1 & \text{if } \phi_{highest}^T x - A < 0 \\ -1 & \text{if } \phi_{highest}^T x - A \geq 0 \end{cases} \quad (3)$$

where  $A = \phi_{highest}^T \bar{x}$

(3)식은 결국 가장 높은 고유치에 대응되는 고유벡터를 축으로 이용하여 투영된 평균을 중심으로 공간상에 왼쪽과 오른쪽에 위치하는 학습 영상으로 분해하는 것이다. 여러 가지 가능한 방법이 있겠지만 우리는 가장 넓은 분산량을 갖는 고유벡터를 선택하여 학습영상을 나눔으로써 빠르게 분산량을 줄이게 된다. (3)식에 의하여 학습영상은  $x_i^{Left} = \{x | f(x) = 1\}$ ,  $x_i^{Right} = \{x | f(x) = -1\}$  로 분해가 일어나게 되고, 분해된  $x_i^{Left}$ ,  $x_i^{Right}$  군집에 대하여 (2)식을 이용하여 결정경계를<그림. 3 참조> 구한다.



<그림. 2> 모의 실험 영상에 대한 하나의 가우시안을 이용한 결정경계  
 [■, □: 동일, ◆: 미분류된 음의 학습 영상]

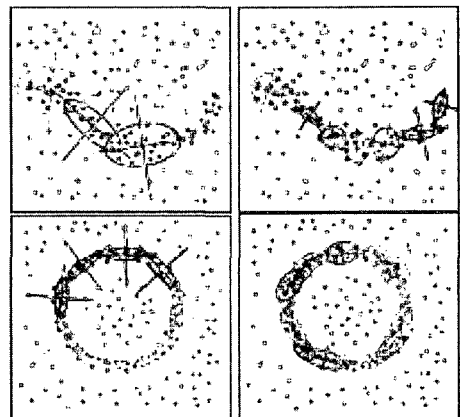


<그림. 3> 모의 실험 <가>에 대한 1번 분해 적용 결과

2.2. 분해 트리의 정의

2.1절에서 전체 학습영상을 가장 높은 고유치와 대응되는 고유벡터를 이용하여 분해하는 방법을 설명하였다. 여기서 주목할 것은 학습영상이 여러 개의 군집으로 분해되는 과정에서 트리 구조와 유사한 형태를 가지게 된다는 사실이다. 분해 방법은 단순히 학습 영상을 왼쪽과 오른쪽으로 나누는 역할을 하게 되므로 왼쪽과 오른쪽으로 나뉘어진 학습영상을 왼쪽 자식과 오른쪽 자식으로 생각할 수 있다. 이와 같이 분해 되는 과정을 트리로 유지하며 우리는 이 트리를 분해 트리라 정의하였다. 분해 트리는 다음과 같은 특성을 갖게 된다.

1. 상위노드는 하위 노드에 속해 있는 모든 양의 데이터를 수용하는 결정경계를 표현한다.
2. 각각의 리프 노드는 다양체의 특정영역의 분포를 설명한다.
3. (3)식에 의하여 양의 학습영상은 자신이 속하는 리프 노드로 찾아 갈 수 있다.



<그림. 5> 모의 실험 문제에 대한 단순 분해 결과의 예

2.3. 단순 및 적응 분해 전략

2.2절까지 분해 방법과 분해 결과를 분해 트리로 정의함을 설명하였다. 본 절에서 “어떤 경우에 분해가 일어나야 할 것인가?” 에 대한 물음에 답을 제시하고자 한다. 먼저, 모든 양의 학습영상을 통해 다변량 정규분포를 생성하고 생성된 정규분포를 이용하여 얻어진 결정경계를 루트노드로 생성한다. 단순 분해 전략은 모든 리프 노드를 m번 분해하는 전략이다<그림. 5 참조>. 하지만 단순 분해 전략은 필요이상의 분해가 일어날 가능성을 내포하고 있으므로 적응분해 전략을 새로이 구상하였다.

- 결정경계에 속하는 음의 영상을 추출한다.

$$x_j^{Neg} = \{x \in X^{Neg} \mid D(x) > \alpha\} \quad (4)$$

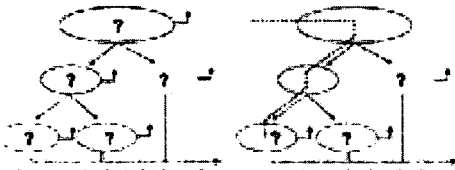
여기서  $X^{Neg} = \{x_1^{Neg}, \dots, x_m^{Neg}\}$  이고  $m$ 은 음의 영상의 개수이다.

- 추출된  $x_j^{Neg}$ 에 대하여 (3)식에 의하여 자신이 속하는 리프노드로 찾아간다.
- (2)식에 의하여 음의 영상으로 분류하지 못하는 경우, 리프노드를 분해한다.

#### 2.4. 수용 / 거부, 리프 노드 분류기의 정의

수용 / 거부 분류기는 리프 노드를 찾아가는 도중 거치는 모든 노드에서 분류 시도를 하는 방법이다. 수용 / 거부 분류기는 음의 영상이 대부분 양의 영상과 유사할 확률이 낮은 경우, 리프 노드를 찾아가기 전에 거부하게 된다.

리프 노드 분류기는 적응 분해 전략의 방법과 유사하다. 모든 영상은 자신이 속하는 리프 노드를 찾아서 해당되는 리프 노드의 결정경계에 의해 분류된다.



<그림. 6> 수용/거부, 리프 노드 분류기의 정의

#### 3. 얼굴검출

제안하는 방법 뿐만 아니라 모든 다양체 분석의 기본은 샘플링이다. 다양체를 잘 분석하기 위해서는 다양체의 형태를 따라 널리 분포된 관측치가 필요하다. 가장 직관적인 방법은 많은 얼굴 데이터 베이스를 통해 다양한 얼굴 영상을 가능한 한 많이 수집하는 것이다. 얼굴 객체에 대해 적용하기 전에 다음과 같은 3가지 가정을 두었다.

- 모든 음의 영상들은 어떤 양의 영상과도 같을 수 없다.
- 수집된 양의 영상은 풍부하며, 다양체 전반을 거쳐 분포되어 있다.
- 각각의 군집의 분포는 가우시안 분포와 유사하다.

##### 3.1. 정규화 및 특징추출

실험에 사용한 정규화 및 특징추출은 [4]와 동일하다. 정규화를 위해 학습 영상에 (5)식을 적용하고,

$$\tilde{x} = \frac{x - \mu}{\sigma} \quad (5)$$

학습영상을 정규화 한 후 (6)식에 의해 특징을 추출한다. 추출된 특징을 (5)식에 의하여 평균을 0으로, 분산을 1로 다시 정규화 하였다.

##### 3.2 다양체 분해 방법의 통계적 학습

학습 방법은 [1][4]와 유사하며, 음의 영상을 획득하기 위하여 Bootstrap[4]을 이용하였다.

$$\tilde{I} = (\tilde{I}^T \tilde{I}_h \tilde{I}_v^T)^T \quad \tilde{I} \in \mathfrak{R}^{(2n^2-2n) \times 1}$$

$$\hat{I}(i, j) = \frac{I(i, j) - \mu}{\sigma}, \quad I(i, j) \in \mathfrak{R}^m, \hat{I}(i, j) \in \mathfrak{R}^m \quad (6)$$

$$\hat{I}_h(i, j) = \hat{I}(i+1, j) - \hat{I}(i, j) \quad 1 \leq i < n, 1 \leq j \leq n$$

$$\hat{I}_v(i, j) = \hat{I}(i, j+1) - \hat{I}(i, j) \quad 1 \leq i \leq n, 1 \leq j < n$$

- 얼굴 객체에 대한 결정경계를 생성하여 루트노드로 설정한다.
- 정지영상(얼굴을 제외한 자연영상)을 입력으로 리프 노드 분류기로부터 추출되는 분류하지 못한 비 얼굴 객체를 수집한다.
- 수집된 비 얼굴 객체를 통해 적응 분해 전략을 적용한다.
- 2~3번 과정을 반복한다.

#### 4. 실험결과 및 결론

수집한 다인종 얼굴(동양인, 백인, 흑인, 히스패닉) 데이터 베이스에서 QQVGA영상 397장(각 인종 당 약 100장씩)을 실험하여 100%의 검출율과 0개의 오분류의 성능을 얻었다. 앞으로 더 많은 데이터 베이스에서 실험 및 매개 변수 조정이 필요하며, 분해방법 및 Overfitting에 대한

	검출율	오분류
동양인(100장)	100%	0
백인(100장)	100%	0
흑인(100장)	100%	0
히스패닉(97장)	100%	0

<다인종 얼굴 DB에 대한 실험결과>

연구가 더 진행되어야 할 것으로 사료된다.

#### 참고문헌

- [1] K.K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998
- [2] M.H. Yang, N. Ahuja, and D. Kriegman, "Face Detection Using Mixture of Linear Subspaces," Proc. Fifth Int'l Conf. Automatic Face and Gesture Recognition, pp. 70-76, 2000
- [3] H. Schneiderman and T. Kanade, "Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition," Proc. Conf. Computer Vision and Pattern Recognition, pp. 45-51, June 1998
- [4] C. Liu, "A Bayesian Discriminating Features Method for Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no. 6, pp. 725-740, June. 2003