

서포트 벡터 머신을 이용한 실시간 얼굴 학습 방법

이일용^o 안정호 변혜란

연세대학교 공과대학 컴퓨터과학산업시스템공학과
{hamir77^o, jungho, hrbyun}@aipiri.yonsei.ac.kr

Real Time Face Training Method Using Support Vector Machine

Ilyong Lee^o Jung-Ho Ahn Hyeran Byun
Dept. of Computer Science, Yonsei University, Seoul

요약

근래 패턴인식 분야에 서포트벡터머신(Support Vector Machine)이 많이 사용되어지고 있다. 서포트벡터머신이 전통적인 패턴인식 방법론에 비해 우수한 성능을 보이고 있지만, 적은 클래스의 숫자, 문자 인식과는 달리 클래스의 수가 많고, 고정되어있지 않은 얼굴인식에서는 새로운 클래스가 등록될때마다 학습을 반복해야 한다. 그러나, 서포트벡터의 특성상 학습시의 계산의 복잡성 때문에 실시간 학습은 사실상 불가능하다. 이에 이 논문에서는 서포트벡터머신을 이용한 실시간 얼굴인식 시스템에서의 빠른 학습방법을 제안했다. 이 시스템은 다중 클래스 인식방법 중 일대다(One Per Class)방법을 채택했으며, 캠브리지(Cambridge) ORL 얼굴 데이터를 임의적으로 11개의 실험 데이터 셋으로 변형한 후 실험 및 평가해 본 결과 빠른 학습능력을 보임과 동시에 인식률에서도 별 차이가 없는 것을 확인할 수 있었다.

1. 서론

최근 몇 년 동안 서포트벡터머신은 아주 효과적인 얼굴인식 방법으로 제안되어왔다[1][5][8]. 서포트벡터머신은 기본적으로 이진분류기이지만, 얼굴인식같은 다중클래스의 문제를 해결하기 위해 다중클래스 서포트벡터머신을 이용한다[1]. 이렇듯 서포트벡터머신은 여러 가지 학습방법에 의해 발전되어왔으며, 기존의 얼굴인식 방법보다 우수한 성능을 보이고 있다.

현대에는 통신기술이 발전하여 인터넷이라는 가상공간이 형성되면서 인터넷을 통한 서비스가 폭발적으로 증가하고 있다. 이에 인터넷상에서 화상카메라를 이용한 얼굴인식 기술도 그 필요성이 대두되게 되었고, 여러 가지 방법으로 사용되게 되었다. 하지만, 얼굴이라는 도메인은 클래스의 수가 일정히 정해져있는 숫자나 문자와는 달리 그 수가 무한하다고 할 수 있다. 따라서 어떠한 얼굴인식시스템에 새로운 클래스를 등록하려고 할때, 학습을 반복해야하는 문제가 발생한다. 이 문제는 인터넷서비스에서는 치명적인 문제로 대두된다. 학습을 시키려면 일단 온라인 서비스를 중단해야하기 때문에, 학습시간을 단축하는 것이 큰 관건이 된다[2].

이 논문에서는 서포트벡터머신을 이용하여 기존에 학습된 클래스에 새롭게 어떠한 클래스들이 추가로 등록될때 인식률을 떨어뜨리지 않고도 빠르게 학습될수 있는 방법을 제안했다.

2. 이론적 배경

2.1 서포트 벡터 머신

두 개의 클래스 분류 문제는 어떠한 함수에 의해 클래스를 positive(+1)한 클래스와 negative(-1)한 클래스로 분리하는 것이 목적이라 할 수 있겠다. 서포트벡터머신은 그러한 두 개의 클래스를 분류하는 최적의 분리 결정 평면을 구

하는데 목적을 둔다.[4]. 이 때 결정평면은 가장 큰 마진을 갖는데, 각각의 마진상에 있는 벡터들을 서포트벡터라고 하고, 이 서포트 벡터들을 이용하여 새로운 벡터들이 입력되어 졌을 때 클래스를 구분할 수 있는 것이다.

$x_i \in R^n, y_i \in \{-1, +1\}$ 인 벡터집합은 $(x_1, y_1), \dots, (x_l, y_l)$ 이며 결정평면은 다음과 같다.

$$g(x) = w \cdot x + b = 0 \quad (1)$$

이때 x 는 입력벡터, w 는 조절 가능한 가중치(weight)벡터, b 는 절편(bias)이다. w 와 b 는 다음의 식 (2)를 만족시키고

$$y_i(w \cdot x_i + b) \geq 1, \quad i = 1, \dots, l \quad (2)$$

아래 비용함수를 최소화시키도록 최적화 된다[4].

$$\Phi(w) = \frac{1}{2} \|w\|^2 \quad (3)$$

제한식(2)에 대한 식(3)의 최적문제의 해는 다음 라그랑지 함수의 안장점으로 얻을 수 있다.

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i [y_i(w \cdot x_i + b) - 1] \quad (4)$$

해를 쉽게 취할 수 있도록 이원문제(dual problem)로 전환한다.

$$\max_w W(\alpha) = \max \left\{ \min_w L(w, b, \alpha) \right\} \quad (5)$$

최적의 해를 결정하는 점 α 는

$$\alpha = \arg \min \left\{ \sum_{i=1}^l \alpha_i - \sum_{i=1}^l \alpha_i \langle y_i y_j \rangle \langle x_i \cdot x_j \rangle \right\} \quad (6)$$

이때의 제한식은,

$$\alpha_i \geq 0 \quad (7)$$

$$\sum_{i=1}^l \alpha_i y_i = 0 \quad (8)$$

제한식(7),(8)을 가지고 식(5)을 풀면 모든 라그랑지 승수를 구할수 있고, 따라서 최적결정평면도 얻을 수 있다. w, b 는 최적해이다.

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (9)$$

$$b = -\frac{1}{2} w \cdot [x_r + x_s] \quad (10)$$

이때, x_r 과 x_s 는 아래를 만족하는 서포트벡터들이다.

$$\alpha_r, \alpha_s > 0, \quad y_r = 1, \quad y_s = -1 \quad (11)$$

새로운 데이터 벡터 x 에 대한 분류함수 f 는 아래와 같다.

$$f(x) = \text{sign}(w \cdot x + b) \quad (12)$$

위에서는 선형적으로 분류가능한 문제에 대해 언급했지만, 인식문제의 대부분은 결정평면이 비선형적이다. 서포트벡터머신은 그러한 경우들 또한 처리할 수 있도록 확장되었다[5][7]. 주된 아이디어는 고차원의 특징공간으로 입력 데이터 x 를 매핑(mapping)하고 그 고차원 공간에서 선형분리를 수행하는 것이다. 이때 커널(kernel)함수 K 를 이용하는데, 이 때 K 는 머서의 정리(Mercer's Theorem)를 만족하는 것이 일반적이다.

2.2 적응적 결정경계(Adaptive Decision Boundaries)

적응적 결정경계 방법은 클래스간의 초기 경계를 입력된 데이터들에 의해 조금씩 수정해가며 최적의 경계를 찾는 것이다. 그 과정은 아래와 같다[3].

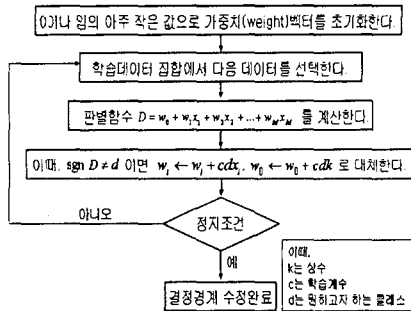


그림1. 결정경계 수정방법

3. 제안된 학습 방법

만약 웹상에서 PC카메라를 이용하여 얼굴인식 시스템에 자신의 얼굴을 등록해야한다면 어떻게 할 것인가. 자신의 얼굴을 등록한다는 것은 곧 기존의 시스템에 자신의 얼굴을 학습시키는 것을 의미한다. 설사 등록이 가능하다고 해도 학습을 시키는 데에는 시간이 오래 걸리며, 또 웹서비스를 해야 하는 서비스 업체 입장에서도 시스템을 중단하고 등록된 모든 얼굴데이터에 대해 학습을 시켜야만 한다. 이에 대한 한 가지 대안으로 실시간으로 등록(학습)을 가능하게 하는 방법을 제안해 보고자 한다.

만약 20개의 클래스가 일대다 방법을 이용한 서포트벡터머신으로 학습되어있는 시스템이 있다고 가정하자. 이때, 서포트벡터머신의 수는 20개가 생성되었을 것이고, 각각의 서포트벡터머신은 라그랑지안 승수, 절편과 서포트벡터들에 대한 정보를 갖고 있을 것이다. 따라서 이 시스템은 20개의 서포트벡터들을 이용하여 40명의 얼굴 인식을 할 수 있다. 이때, 새로운 21번째 사람을 등록(학습)시키고자 한다. 기존의 서포트벡터머신 학습방법에 의하면, 그림 2와 같이 미리 생성된 서포트벡터머신을 무시하고 21명의

얼굴들을 가지고 다시 학습을 시켜야만 한다. 이런 방법으로 학습을 시도한다면, 새로운 사람의 얼굴을 등록(학습)시킬 때 마다 서포트벡터들을 새롭게 다시 생성해야만 하기 때문에 실시간 등록은 가능하지 않다. 하지만, 20개의 서포트벡터머신이 미리 생성된 상태에서 21번째 서포트벡터머신 하나만 생성한다면, 실시간 등록(학습)은 가능해진다. 아래 그림3은 이 학습방법을 구체적으로 보여주고 있다.

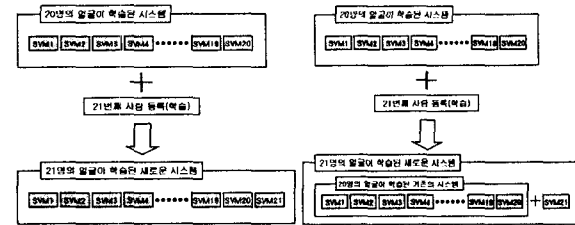


그림2. 기존의 새롭게 학습시키는 방법

그림3. 실시간으로 학습시키는 방법

하지만 그림 3과같은 방법을 취하면, 그만큼의 학습속도는 빠르게 얻을 수 있으나, 인식률이 그림 2의 학습방법에 비해 떨어지게 된다. 따라서 우리는 최종적으로 그림 4와 같은 학습방법을 제안하였다.

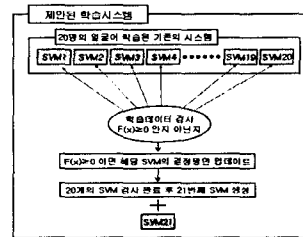


그림4. 제안된 학습시스템

기존에 학습된 서포트벡터머신에 저장되어있는 결정평면을 수정해가면서 서포트벡터머신을 하나씩 추가하는 것이다. 예를들어 21번째 클래스의 5개의 데이터를 학습시키고자 한다면, 5개의 학습데이터가 기존의 20개의 결정평면을 수정시킬 가능성여부를 아래와 같이 검사한다.

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b, \quad n: \text{기존에 학습된 데이터수} \quad (13)$$

이때 $\text{sign}(f(x_{\text{new}}))$ 가 -1이 아니면, 적응적 결정경계 방법에 의해 결정평면을 수정하게 된다.

$$w_{\text{new}} = w_{\text{old}} + \eta y_{\text{new}} \Phi(x_{\text{new}}) \quad (14)$$

$$b_{\text{new}} = b_{\text{old}} + \eta y_{\text{new}} k$$

여기서, η 는 학습계수이고 k 는 상수이다. 따라서, 분류함수 f 는 다음과 같이 수정(adapt)된다.

$$f_{\text{new}}(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + \eta y_{\text{new}} K(x_{\text{new}}, x) + b_{\text{old}} + \eta y_{\text{new}} k \quad (15)$$

$$= f_{\text{old}}(x) + \eta y_{\text{new}} K(x_{\text{new}}, x) + \eta y_{\text{new}} k$$

위와같이 미리 생성된 결정평면에 직접적으로 영향을 미칠 것으로 예상되는 새로운 데이터들을 이용하여 기존의 결정평면들을 수정하였다. 또한 아래 그림 5와 같이

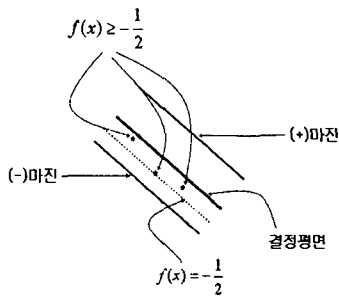


그림5. 소프트마진(soft margin)

$f(x) \geq -1/2$ 인 경우까지도 x_{new} 가 확률변수이므로, 결정 평면에 영향을 줄 수 있을거라고 가정하고 아래와 같이 소프트마진에 의해 결정평면을 수정하였다.

$$f_{new}(x) = f_{old}(x) + \eta y_{new} m_{new} K(x_{new}, x) + \eta y_{new} m_{new} \quad (16)$$

여기서, 마진 m_{new} 는 $m_{new} = f_{old}(x_{new}) + 1/2$ 로 정의된다. 이러한 수정과정을 마친 후 새로운 클래스에 대한 서포트 벡터머신 생성하여 빠르고 인식률에 안전한 시스템을 구축할 수 있다.

4. 실험 결과

제안된 학습 방법의 타당성을 보이기 위해서 우리는 기존 그림2와 같은 전통적인 학습방법, 그림3과 같은 아이디어가 삽입되어있지 않은 실시간 학습방법과 성능비교를 해보았다. 학습과 테스트 데이터는 캠브릿지 대학의 ORL 얼굴 데이터베이스를 사용하였다. ORL데이터는 총 400개의 얼굴 데이터(92x112)로 구성되어 있다. 총 사람(클래스)수는 40명이며, 1명당 10개의 이미지로 구성되어있는데, 이 중 1명당 5개의 이미지 총 200개로 학습을 했으며, 나머지 200개의 이미지로는 테스트를 해보았다. 이와 같은 실험을 총 11번 반복하였으며, 그때마다의 데이터셋은 완전 랜덤(random)하게 구성되도록 하였다.

우선 기존의 전통적 학습 방법을 이용하여 학습시켰을 때의 인식률과 학습시간을 측정해보았다. 이때 실험은 20개의 클래스를 미리 학습시키고, 새로운 클래스가 하나씩 증가할때마다 기존의 학습된 서포트벡터머신들은 모두 무시된 채 학습된 클래스의 개수를 n개라하면 n+1개의 클래스를 다시 학습하게 되는 것이다. 이때, 추가되는 클래스의 수는 20개이다. 그리고 두 번째 실험은 20개의 클래스를 미리학습시키고, 추가로 새로운 클래스가 하나씩 들어오게 되면, 기존의 학습된 서포트벡터머신들은 그대로 둔채 n+1 번째 서포트벡터머신 하나만 생성하는 것이다. 마찬가지로 추가되는 클래스의 수는 20개이다. 마지막으로 제안된 학습방법을 이용한 실험은 위의 두 번째 실험과 똑같이 시행하되, 서포트벡터머신을 추가하기전에 그전에 생성된 서포트벡터머신들의 결정평면을 수정한다.

위 세 실험을 통해 우리는 아래 그림 5, 6과 같은 결과를 얻어낼 수 있었다. 20개의 클래스를 미리 학습시킨 후 클래스를 하나씩 추가하여 마지막 40번째클래스가 등록될때 학습시키는 시간은 평균 전통적학습방법에서 6.07초, 실시간학습에서 0.26초, 제안된 학습에서 평균 2.23초가 소요되었다. 하지만, 인식률에 있어서는 그림 6과 같이 각각

89.91%, 88.27%, 88.27($f(x) \leq 0$)%, 89.18($f(x) \leq -0.5$)%의 결과를 얻을 수 있었다. $f(x) \leq 0$ 경우에는 인식률을 향상시키지 못했지만, $f(x) \leq -0.5$ 의 경우 만족할만한 인식률을 보여주었다.

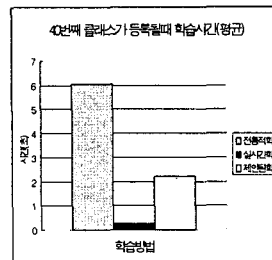


그림5. 평균학습시간 비교

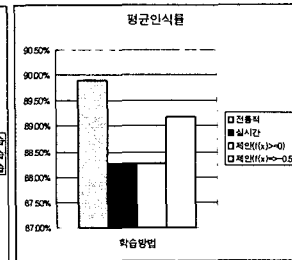


그림6. 평균인식률 비교

5. 결론

우리는 실험을 통해 학습시간을 단축시킬 수 있었다. 클래스의 수가 증가될수록 학습시간의 차이는 더 커졌을 것이다. 이것은 비단 학습시간을 줄이는 데에만 의의가 있는 것이 아니다. 실시간으로 패턴을 인식하는 시스템에서 짧은 시간 안에 등록(학습)을 가능케 할 수 있다는데에 더 큰 비중을 둘 수가 있을 것이다. 현재 우리는 다양한 웹서비스를 받고 살아가고 있다. 앞으로 웹 서비스에서도 패턴인식의 응용분야에 대한 서비스가 대두될 것으로 예상된다. 철저한 보안정책이나 멀티미디어, 엔터테인먼트에 많이 사용될 것으로 보이는 패턴인식 시스템에 높은 인식(테스트)뿐만 아니라 우리의 알고리즘으로 실시간 학습까지 가능케 될 수 있을 것이다.

하지만, 계속적인 추가 학습으로 인식률 저하의 문제점을 해결해야만 한다. 어느 정도 추가학습을 허락하고, 그 이후에는 전체적으로 학습을 시켜야만 할 것이다.

5. 참고문헌

- [1] Guodong, San Z. Li, and Kapluk Chan, "Face Recognition by Support Vector Machines", *Image and Vision Computing, Volume 19, Issues9-10, Pages 631-638*
- [2] K.W.Lau O.H.Wu, "Online training of support vector classifier", *Pattern Recognition, Volume 36, Issue 8, August 2003, Page1913-1920*
- [3] E.Gose, R. Johnsonbaugh, S.Jost, "Pattern Recognition and Image analysis", *Prentice Hall PTR, Page 176-182*
- [4] V. N. Vapnik, "Statistical Learning Theory", *Wiley, New York, 1998*
- [5] E.Osuna, R.Freund, F.Girosi, "Improved Training Algorithm for Support Vector Machins", *Neural Networks for Signal Processing 97, 1997*
- [6] E.Osuna, F.Girosi, "Reducing the run-time complexity of Support Vector Machines", *In International Conference on Pattern Recognition, 1998*
- [7] John C. Platt, "Sequential Minimal Optimization : A Fast Algorithm for Training Support Vector Machines", *Technical Report 98-14, MicroSoft Research, 1998*
- [8] B.Heisele, P.Ho, T.Poggio, "Face Recognition with Support Vector Machines:Global versus Component based Approach", *International Conference on Computer Vision(ICCV01), 2001*