

고급소프트웨어실습 7 주차 과제

검출된 얼굴을 이용한 응용 분야

20140938 임다운

문제 1

검출된 얼굴 이미지는 다양한 학습을 위한 데이터로 사용될 수 있다. 얼굴 이미지를 이용한 학습이 도움이 될 것 같은 문제를 세가지 이상 들어 보자. 주로 분류기의 학습을 위하여 사용되며, 그 외의 분야에 사용 될 수도 있다.

1) 범인 검거에 활용

현재 CCTV 를 이용한 범인의 검거는 사람이 수작업으로 영상들을 찾아가면서 이루어지고 있다. 얼굴 이미지를 활용한 학습이 이러한 범인 검거에 적용된다면, CCTV 를 통해서 입력된 얼굴들을 비교하여 범인의 동선을 쉽게 파악할 수 있을 것이다. 또한 CCTV 데이터에 상시적으로 접근할 수 있는 네트워크가 설정되어 있는 경우 입력된 용의자의 얼굴을 자동으로 인식하여 경찰에 알람을 줄 수도 있을 것이다.

2) 얼굴인식을 통한 보안에 활용

출입이나 기기 보안과 같이 별도의 비밀번호 입력이나 지문인식과 같은 사용자의 동작이 필요 없이 얼굴인식을 통한 암호화가 가능하다. 이미지 학습은 얼굴인식 뿐만 아니라 홍채인식에도 활용이 가능하다. 홍채 이미지를 학습하여 구분하는 과정을 거쳐 개인 간의 홍채 이미지를 구분하여 보안의 수단으로 활용할 수 있을 것이다.

3) SNS 얼굴 검색 및 자동 태그

이미지가 주로 사용되는 SNS 의 특성상 얼굴 이미지 인식 및 검색은 사용자의 편의를 높여줄 수 있을 것이다. 또한 얼굴 image 마다 유저의 아이디 및 이름을 tag 시켜주는 것이 상당히 사용자 경험을 저하시키는 일이므로, 이를 자동화하여 사진을 업로드함과 동시에 얼굴인식을 통해 태그를 달아준다면 SNS 의 연결성을 재고하면서 사용자의 불편함을 없앨 수 있다.

4) Public device 의 customization 에 활용

여러 명의 user group 의해 사용되는 디바이스의 경우 얼굴인식을 통하여 현재 사용자에 따라 자동으로 customize 된 setting 을 제공할 수 있다. 예를 들어 집안에 있는 TV 나 PC 와 같은 경우 이미지 학습을 통해 사용자를 자동으로 인식하여 사용자에게 맞는 계정을 리턴하여 주는 기능을 만들 수 있을 것이다.

문제 2

다수의 얼굴 이미지 데이터를 통계적 관점에서 본다면 어떤 분포를 따를 것인지 설명해 보자.

Gaussian 분포를 따를 것으로 예상된다. 사람들의 얼굴은 모두비슷한 형태를 가지고 있다. 눈, 코, 입을 가지고있기 때문이다. 따라서 이를 바탕으로 얼굴인식 데이터를 축적한다면 Gaussian 분포를 따를 것으로 추측할 수 있다. 논문 Statistics in Face Recognition: Analyzing Probability Distributions of PCA, ICA and LDA Performance Results 을 참조해보면, PCA, ICA, LDA 와 같은 얼굴 인식 알고리즘을 비교한 결과 데이터들이 완벽한 정규분포는 아니지만, 대체로 한 쪽이 약간 더 긴 꼬리모양을 가진 gaussian 분포가 나옴을 알 수 있다.

문제 3

얼굴 이미지는 대표적인 고차원 벡터를 형성한다. 실습에서 생성한 200'200 이미지만 해도 픽셀 값들을 칼럼 방향으로 죽 이어 붙이면 40000 차원의 벡터가 생성된다. 이와 같은 고차원 데이터를 2 차원 또는 3 차원 공간에 표현하는 것이 가능할 지와 그 방법에 관하여 대략적으로 설명해 보자.

n 차원의 벡터가 생성된다고 했을 때, 그 n 의 크기가 얼마나 커지든 상관 없이 2 차원이라면 2 개의 서로 수직인 주성분 벡터를, 3 차원이라면 서로 수직인 3 개의 주성분 벡터들을 구한다면 2,3 차원 공간에 해당 벡터를 표현하는 것이 가능할 것이다.

이렇듯 주성분을 이용해 고차원의 데이터를 저차원으로 줄이는 차원 축소 방법을 Principal Component Analysis 기법이라고 한다. PCA 는 데이터 하나하나에 대한 성분을 분석하는 것이 아니라 데이터들이 모여서 어떤 분포를 이룰 때 그 분포의 주성분을 분석하는 것을 말한다. Principal Component (주성분)은 이 좌표계에서 벡터의 움직임이 가장 큰, 즉 가장 큰 분산을 가지는 축을 말한다. 분산을 계산하여 이러한 축을 찾아내고, 이 분산들 값에 아이겐벡터를 활용하여 저차원으로 줄일 수 있다.

출처: <http://t-robotics.blogspot.kr/2014/10/pca-principal-component-analysis.html#.WgqPlhOCyRs>