

평균이동 기법을 이용한 랜덤포레스트 기반 실시간 얼굴 특징점 추적

장성걸, 박종일¹

한양대학교 컴퓨터소프트웨어학과

sgzhang@mr.hanyang.ac.kr, jipark@hanyang.ac.kr¹

Real Time Face Tracking Method based Random Regression Forest using Mean Shift

Xingjie Zhang, Jong-Il Park

Dept. of Computer and Software, Hanyang University

요 약

본 논문에서는 평균이동 (mean shift) 기법을 이용하여 랜덤포레스트 (random forest) 기반 실시간 얼굴 특징점 추적 (facial features tracking) 방법을 제안한다. 우선, 눈의 위치를 이용하여 검출된 얼굴영역을 적절한 크기와 위치로 개선하여 랜덤포레스트를 이용한 얼굴 특징점 추적 알고리즘이 받는, 얼굴검출 (face detection) 과정에 얻어지는 얼굴영역 상자 (face bounding box) 크기와 위치의 영향을 감소 하였다. 또한 랜덤포레스트의 얼굴 특징점 추정결과에서 추정평균 대신 평균이동기법을 이용하여 잘못된 추정결과들을 제거하고 제대로 된 추정결과만 사용하여 얼굴 특징점 검출 정확도를 개선하였다. 따라서 제안하는 방법들을 이용하여 기존의 랜덤포레스트 기반 얼굴 특징점 검출 기법의 성능을 제고하고 실시간으로 얼굴 특징점을 추적할 수 있다.

1. 서론

최근 들어 영상 내 사람의 얼굴을 인식하고 표정을 분석하거나 변형, 증강하는 기술들에 대한 연구가 각광을 받고 있다. 만약 얼굴의 특징 점들의 위치를 정확하게 알 수 있다면 이런 기술들에 대한 연구에 중요한 정보를 제공해 주게 된다. 따라서 효과적인 얼굴 특징점 추적기법에 대한 연구가 활발하게 진행되고 있다.

랜덤포레스트를 이용한 얼굴 특징점 검출 기법은 회귀나무의 앙상블 학습방법을 이용하여 화소의 밝기값으로부터 직접 얼굴의 특징점 위치를 추정하는 기법으로써 널리 알려져 있다 [1,2,3]. 이러한 랜덤포레스트 기반 얼굴 특징점 검출 기법은 간단한 구현으로 빠른 실행속도와 정확성으로 얼굴 특징점 검출을 진행할 수 있으므로 실시간 얼굴인식이나 표정인식, 얼굴교체 등과 같은 다양한 연구의 전처리 과정으로 사용되고 있다.

이렇게 높은 효용성에도 불구하고 실제로 실시간 얼굴 특징점 추적에 랜덤포레스트를 사용하기에는 몇 가지 제약이 존재한다. 일반적으로 랜덤포레스트를 이용한 얼굴 특징점 검출 기법은 랜덤으로 생성된 회귀나무 포레스트를 캐스케이드 학습 프레임워크 (cascade learning framework) 방식으로 배치하여 단계 별로 점진적으로 얼굴의 특징점들을 검출한다 [1]. 우선 랜덤포레스트를 이용한 얼굴 특징점 검출에서는 초기 얼굴형태를 얼굴영역 공간에 대응되게 설정하기 때문에 얼굴

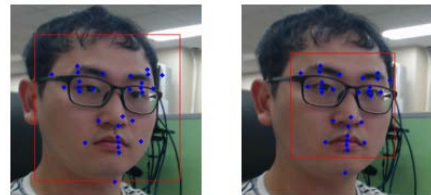


그림 1. 얼굴영역 상자 크기에 따른 얼굴 특징 점 추정 결과.

검출기를 통해 얻어지는 얼굴영역 상자의 크기와 위치의 영향을 비교적 심하게 받게 된다. 즉 얼굴 검출과정에 얻어진 얼굴영역 상자의 크기와 위치가 적절하면 높은 정확도를 보여 주지만, 그 반대로 얼굴영역 상자가 너무 크거나 작은 등 크기가 적절하지 못하거나, 혹은 위치가 얼굴 중심으로부터 틀어지면 잘못된 결과를 추정하게 된다 (그림 1). 또한 여러 개의 초기 얼굴형태 (initial shape) 를 입력으로 하여 얼굴 특징점의 최종 추정결과를 평균 추정치로 정하기 때문에, 그 중 소수의 초기 얼굴형태가 잘못된 추정을 하게 되면 최종 결과도 따라서 영향을 받게 된다.

본 논문에서는 이상의 문제점을 해결하여 실시간으로 얼굴의 특징점을 추적할 수 있는 알고리즘을 제안한다. 구체적으로 얼굴 검출기를 통해 검출된 얼굴 영역에서 눈의 위치를 검출하여 두 눈의 위치와 두 눈 사이 거리를 이용하여 적절한 크기와 위치의 얼굴영역 상자를 얻고, 여러 개의 추정 결과에 평균이동 알고리즘을 적용하여 잘못된 추정치의 영향을 제거하는 방법을 제안한다.

¹ 교신저자

2. 얼굴 검출 및 개선

랜덤포레스트를 이용하여 얼굴 특징점을 추정하기 위해서는 먼저 얼굴영역상자를 검출해야 한다. 여기서 얼굴영역상자란 사람의 얼굴의 안면을 표시하는 사각형 영역으로써, 랜덤포레스트에서는 안면이 최대한 중심에 위치하는, 마침 눈, 코, 입을 다 둘러싸는 안면영역 상자가 필요하다. 하지만 일반적인 얼굴검출 방법에서는 얼굴영역이 어느 정도 겹쳐있기만 하면 다 정확한 검출로 분류되기 때문에, 많은 경우에 특징점을 제대로 추정하기 어렵다.

본 논문에서는 우선 얼굴검출 기법을 이용하여 안면영역상자를 검출한 다음 안면영역 내에서 두 눈의 위치를 검출하였다. 그 다음 검출된 두 눈의 위치를 이용하여 안면영역상자의 크기와 위치를 요구에 맞게 재설정 하였다.

3. 얼굴 특징 점 추적

랜덤 포레스트를 이용하는 얼굴 특징점 추정 기법들은 일반적으로 랜덤으로 여러 개의 초기얼굴 형태를 입력 받아 cascade boosting 의 방식으로, 점진적으로 얼굴의 위치를 추정해 나간다. 즉 N 개의 초기얼굴 형태를 입력하였다면 랜덤 포레스트를 거쳐 N 개의 추정결과를 얻게 되는데 이때 얻은 결과의 추정 평균치를 최종결과로 한다.

$$S = \frac{1}{N} \sum_{i=1}^N S_i \quad (1)$$

공식 1 에서 랜덤포레스트의 얼굴 특징점 추정과정을 보여준다. 여기서 S 는 최종 추정된 얼굴 특징점의 분포이며 N 개의 추정결과와 평균을 최종 추정결과로 한다. 하지만 N 개의 추정결과 중에 한두 개의 잘못된 추정결과가 있으면 최종결과 역시 영향을 받아 잘못된 형태로 추정되게 된다. 본 논문에서는 평균이동 알고리즘을 이용하여 잘못된 추정결과를 제거하였다. 그림 2 에서는 평균이동 알고리즘을 통해 잘못된 추정결과를 제거하는 과정을 보여준다.

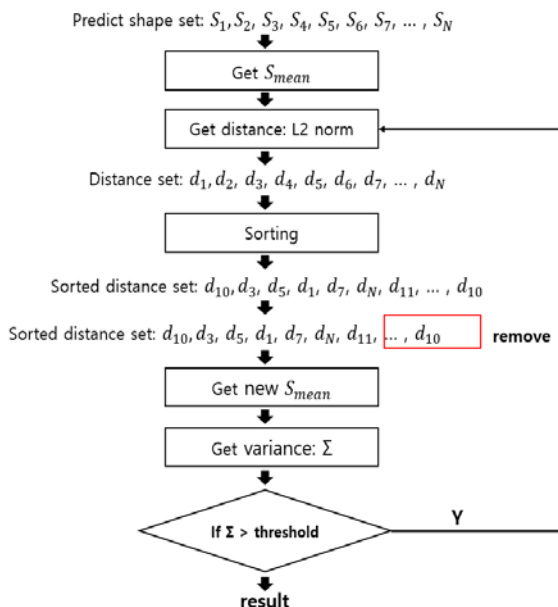


그림 2. The algorithm of remove fake prediction.

4. 실험결과

실험에 사용된 얼굴검출 기법은 비교적 널리 알려진 Haar 특징 기반 cascade face classifier 기법이며 갈만필터를 이용하여 얼굴의 위치를 추적하였다. 랜덤포레스트는 CVPR 2013 에 발표된 explicit shape regression (ESR) 방식을 사용하였으며 초기얼굴 형태의 개수는 20 개로 정하고 실험을 진행하였다. 평균이동 알고리즘에서 사용된 임계값은 초기얼굴 형태 개수의 70%인 14 개를 취하였다. 실험결과에서 위쪽은 기존의 ESR 방식으로 얼굴 특징점을 추정한 결과이고 아래쪽은 제안하는 방법으로 얼굴 특징점을 추정한 결과이다. 실험결과에서 보여 주듯이 기존의 방식이 실패하는 경우에도 제안하는 방법을 이용하여 얼굴의 특징 점의 위치를 정확하게 추정할 수 있음을 알 수 있다.

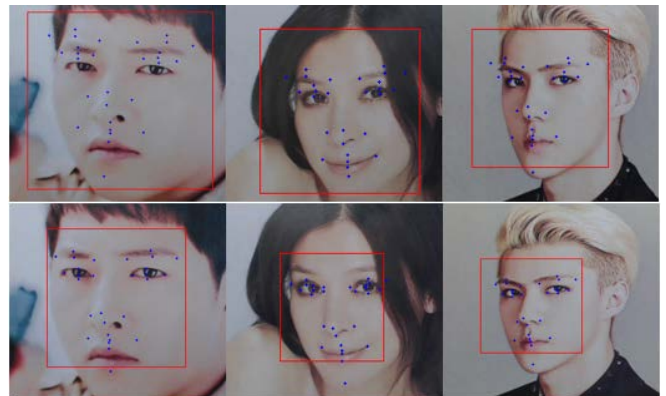


그림 3. 실험결과

5. 결론

본 논문에서는 평균이동 기법을 이용한 랜덤포레스트 기반 실시간 얼굴 특징점 추적 방법을 제안하였다. 두 눈의 위치를 이용하여 얼굴영역 상자의 크기와 위치를 적절히 조절함으로써 랜덤포레스트의 추정에 적합한 입력을 생성 하였으며, 평균이동 기법을 이용하여 잘못된 추정결과와 영향을 제거하였다. 실험을 통해 제안하는 방법이 기존 방법에 비해 개선 된 부분을 보여 주었다. 그러나 두 눈이 가려짐 등 원인으로 검출에 실패하는 경우에는 여전히 얼굴 특징점 추정도 불안정하게 되는 문제가 있다. 이 부분은 향후 연구에서 해결하려고 한다.

감사의 글

"본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학 ICT 연구센터육성 지원사업의 연구결과로 수행되었음" (IITP-2017-2012-0-00628)

참고문헌

- [1] Xudong Cao, Yichen Wei, Fang Wen and Jian Sun. "Face alignment by explicit shape regression." International Journal of Computer Vision. 107.2 (2014), pp 177-190.
- [2] Shaoqing Ren, Xudong Cao, Yichen Wei and Jian Sun. "Face alignment at 3000 fps via regressing local binary features." The IEEE Conference on Computer Vision and Pattern Recognition, 2014. pp 1685-1692.
- [3] Kazemi, Vahid, and Josephine Sullivan. "One millisecond face alignment with an ensemble of regression trees." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014, pp 1867-1874.