

적응적인 평균 모양을 이용한 동적 모양 모델 기반 얼굴 특징점 추출

김현철*, 김형준*, 황원준**, 기석철**, 김희을*

*한양대학교 전자통신컴퓨터공학과, **삼성종합기술원

hckim@vision.hanyang.ac.kr, wykim@hanyang.ac.kr

Facial Feature Extraction using an Active Shape Model with an Adaptive Mean Shape

Hyun-Chul Kim*, Hyoung-Joon Kim*, Wonjun Hwang**, Seok-Cheol Kee**, Whoi-Yul Kim*

*Division of Electrical and Computer Engineering, Hanyang University

**Samsung Advanced Institute of Technology

요 약

본 논문은 포즈가 취해진 얼굴의 정확한 특징점 추출을 위하여 적응적인 평균 모양 방법을 이용한 ASM(Active Shape Model)을 제안한다. ASM은 사람 얼굴의 모양을 모델링하기 위하여 통계학상의 모양 모델을 이용한다. 통계학상의 모양 모델의 평균 모양은 입력 영상의 얼굴 포즈와 관계없이 하나로 고정되어 있으며, 이는 모양 모델 제한 조건 검사 및 복원 과정에서 잘못된 결과를 만드는 원인이 된다. 이러한 문제를 해결 하기 위하여 입력 영상의 얼굴 모양에 적응적인 평균 모양을 제안하며, 실험을 통해 제안한 방법이 고정된 평균 모양 방법의 문제를 해결하고 특징점 추출 성능을 향상시킴을 보였다.

1. 서론

정확한 얼굴 특징점 추출은 얼굴 표정 분석, 얼굴 합성, 얼굴 인식 및 모델 기반의 부호화에서 매우 중요한 과정이며, 일반적으로 얼굴 특징점 추출은 눈, 입, 코와 같은 얼굴의 중요한 부분을 찾는 것을 의미한다.

ASM(Active Shape Model)은 Cootes에 의해 제안된 특징점 추출 방법으로써 사람 얼굴의 특징점 추출을 위한 목적으로 널리 이용되고 있다 [1]. 하지만 사람 얼굴 영상에 대해서 ASM만을 이용할 경우에는 정확한 특징점 추출을 기대할 수 없으며 [2], 이를 해결 하기 위하여 Kernel PCA [2], Wavelet [3]을 활용한 다양한 방법들이 연구되고 있다.

본 논문에서는 고정된 평균 모양의 사용으로 인해 발생하는 ASM의 문제점을 밝히고, 이를 해결 하기 위한 방법으로 입력 영상의 얼굴 포즈에 적응적인 평균 모양을 제안한다. 적응적인 평균 모양을 구하기 위하여 최적화 방법인 유전자 알고리즘(Genetic Algorithm) [4]을 사용하며 유전자 알고리즘에 사용되는 목적 함수는 ASM의 학습과정을 통해서 구한 통계학상의 외관 모델 [5] 정보로 구성된다.

2. ASM

ASM은 통계학상의 모양 모델인 식 (1)과 외관 모델을 이용하여 특징점을 추출하는 방법이다.

$$\mathbf{x}_f \approx \bar{\mathbf{x}} + \mathbf{P}_m \mathbf{b}_f \quad (1)$$

여기서 $\bar{\mathbf{x}}$ 와 \mathbf{P}_m 은 ASM의 학습 과정을 통해 생성된 평균모양과 고유 벡터의 열로 구성된 행렬이며 \mathbf{b}_f 는 모양파라미터를 원소로 하여 구성된 열벡터이다.

ASM은 특징점 추출을 위하여 모양의 조정과 모양 모델의 제한 조건 검사 과정을 반복적으로 수행하며 구체적인 과정은 다음과 같다.

특징점을 추출하기 위하여 현재의 모양을 외관 모델과 입력 영상을 이용하여 조정한다. 그 후 조정된 모양(\mathbf{x}_f)이 모양 모델 제한 조건[3]을 만족하는지의 여부를 판별하기 위하여 \mathbf{b}_f 를 구한다. 이 때 올바른 모양의 유지를 위해서는 \mathbf{b}_f 의 모든 원소들이 $\pm 3\sqrt{\lambda_i}$ (λ_i 는 ASM의 학습 과정을 통해 생성된 고유값)의 범위 안에 있어야 하므로 그 범위 밖의 값을 가지는 원소는 그 범위 안의 값을 가지도록 강제로 조정된다. 앞의 과정은 임의로 정한 수렴 조건이 만족 될 때까지 진행되며 조건이 만족되면 특징점 추출과정은 완료된다.

위와 같은 일반적인 ASM 과정을 포즈가 취해진 얼굴에 적용할 경우에는 모양 모델 제한조건 검사 과정에서 문제가 발생하게 되는데 그림 1은 이를 보여준다. (a)는 평균 모양이고, (b)는 모양 모델을 입력 영상에 조정한 모양이며, (c)는 모양 모델 제한 조건을 검사하고 복원된 모양이다. 여기서 (b)의 올바른 모양이 (c)와 같은 기형적인 모양으로 변하게 된 이유는 고정된 평균모양 $\bar{\mathbf{x}}$ 의 사용으로 인하여 \mathbf{x}_f 와 $\bar{\mathbf{x}}$ 의 차이가 클 때도 불구하고 모양 모델의 제한 조건 검사를 위해 사용되는 λ_i 은 변하지 않아 $\pm 3\sqrt{\lambda_i}$ 의 범위 안의 값을 가지도록 인위

적으로 조정되는 \mathbf{b}_f 의 원소들이 많아지기 때문이다. 이를 강제로 $\pm 3\sqrt{\lambda_i}$ 의 범위 안으로 조정하게 되면 (c)와 같은 기형적인 모양이 생성된다.

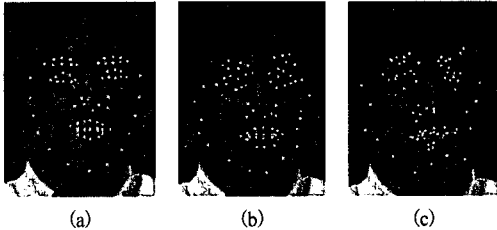


그림 1. 고정된 평균 모양에 의한 특징점의 위치 변화

3. 적응적인 평균 모양

본 논문에서는 위의 문제점을 해결 하기 위한 방법으로 입력영상의 얼굴 모양에 적응적인 평균 모양($\bar{\mathbf{x}}_a$)을 제안한다. $\bar{\mathbf{x}}_a$ 는 입력 영상의 얼굴 형태와 유사하기 때문에 \mathbf{x}_f 와 $\bar{\mathbf{x}}_a$ 의 차이를 적게 하며 \mathbf{b}_f 의 원소들이 $\pm 3\sqrt{\lambda_i}$ 의 범위 안의 값을 가지도록 한다. 따라서, 올바른 ASM 과정을 수행할 수 있게 된다.

유전자 알고리즘과 통계학상의 외관 모델을 이용하여 $\bar{\mathbf{x}}_a$ 을 구하는 방법을 단계별로 나누어서 설명하면 다음과 같다.

먼저, 단계 1에서는 길이가 $4 \times m$ 비트인 염색체를 그림 2와 같이 랜덤하게 n 개 생성한다. n 은 생성하고자 하는 후보 얼굴 모양(\mathbf{x}_i)의 개수이고 m 은 얼굴 모양을 만들기 위해 사용하는 고유벡터의 개수이며, $c_{i,k}$ 는 0 또는 1의 값을 가진다.

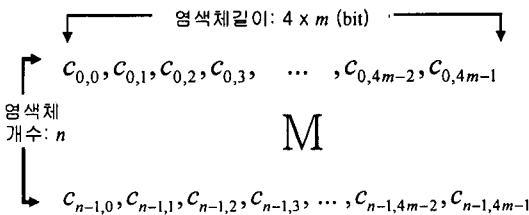


그림 2. 염색체 집합

단계 2에서는 4 비트당 하나의 $\alpha_{i,k}$ 를 생성하며 $\alpha_{i,k}$ 는 얼굴 모양을 만들기 위해 고유벡터에 가해지는 비중이다. $\alpha_{i,k}$ 는 식 (2)와 (3)을 이용하여 만들어진다.

$$\omega_{ik} = \sum_{l=4k}^{4k+3} c_{il} \cdot 2^{l-4k} \quad (2)$$

$$\alpha_{i,k} = 2 \frac{\omega_{i,k}}{3 \cdot \gamma \cdot \sqrt{\lambda_k}} - 3\sqrt{|\lambda_k|} \quad (3)$$

여기서, $i = 0, \dots, n-1, k = 0, \dots, m-1, \gamma$ 는 ω_{ik} 의 값을 정규화하는 상수이며 2^4-1 의 값을 가진다. λ_k 는 ASM 학습 과정에서 생성된 내림차순의 고유값이다 [5].

단계 3에서는 길이가 $4 \times m$ 인 n 개의 염색체를 이용하여 총 $n \cdot m$ 개의 $\alpha_{i,k}$ 를 구한다.

단계 4에서는 $n \cdot m$ 개의 $\alpha_{i,k}$ 와 식 (4) [5]를 이용하여 n 개의 모양 벡터($\mathbf{x}_i = (x_{i,0}, y_{i,0}, \dots, x_{i,m-1}, y_{i,m-1})^T$)를 구한다.

$$\mathbf{x}_i = \bar{\mathbf{x}} + \mathbf{P}_m \mathbf{b}_i \quad (4)$$

\mathbf{b}_i 는 i 번째 염색체의 m 개 $\alpha_{i,k}$ 를 원소로 하여 이루어진 열 벡터이다. 이렇게 생성된 \mathbf{x}_i 는 모델링하고자하는 객체의 모양을 유지하게 되는데, 그 이유는 단계 2에서 $\alpha_{i,k}$ 를 생성할 때 그 값들이 $\pm 3\sqrt{\lambda_i}$ 안에 있도록 제한 조건을 정하였기 때문이다 [1].

단계 5에서는 n 개의 \mathbf{x}_i 와 목적함수($f(\mathbf{x}_i)$)를 이용하여 유전자 알고리즘을 수행한다. \mathbf{x}_i 의 적합도를 평가하기 위하여 사용되는 목적함수는 통계학상의 그레이 레벨 외관 모델 [5]과 Mahalanobis 거리를 이용하여 정의하며 그 식은 다음과 같다.

$$f(\mathbf{x}_i) = \sum_{p=0}^{g-1} (\mathbf{g}(x_{i,p}, y_{i,p}) - \bar{\mathbf{g}}_p)^T \mathbf{S}_p^{-1} (\mathbf{g}(x_{i,p}, y_{i,p}) - \bar{\mathbf{g}}_p) \quad (5)$$

여기서, $\mathbf{g}(x_{i,p}, y_{i,p})$ 는 $(x_{i,p}, y_{i,p})$ 의 위치에서 수직방향으로 그레이 레벨 값을 샘플링하여 구해진 벡터이다. $\bar{\mathbf{g}}_p$ 와 \mathbf{S}_p^{-1} 는 통계학상의 그레이 레벨 외관모델의 훈련을 통해서 생성된 평균 벡터와 공분산 행렬의 역행렬이다.

단계 6에서는 유전자 알고리즘의 수렴 조건에 따라 최종 구해진 \mathbf{x}_i 를 입력 영상의 $\bar{\mathbf{x}}_a$ 로 설정하며, 이를 식 (1)의 $\bar{\mathbf{x}}$ 로 대체하여 ASM을 수행한다.

4. 실험 및 결과

ASM의 학습은 인터넷에서 받은 여러 인물의 다양한 얼굴 영상을 통해 수행되었으며, 학습 과정에 사용된 영상의 개수는 318개이고 크기는 100×120 이며 landmark점의 개수는 76개이다.

적응적인 평균 모양을 구하기 위해 염색체의 길이는 16 bit, 염색체의 개수는 120개로 설정하였다. 유전자 알고리즘의 유전 연산자로는 교배만 이용하였으며 교배 확률은 0.98로 설정하였다. 선택방법은 크기가 7인 토너먼트 선택법을 사용하였으며, 수렴 조건은 단순히 60회 반복 수행하는 것으로 설정하였다.

그림 3은 본 논문에서 제안한 방법을 이용하여 구한 적응적인 평균 모양들을 보여주며, 그 평균 모양들이 입력 영상의 얼굴 형태와 대응함을 확인 할 수 있다.

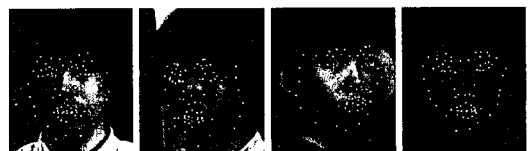


그림 3. 입력 영상의 포즈에 대응하는 평균모양

그림 4는 적응적인 평균 모양 방법에 의한 특징점의 위치변화를 보여준다. (a)는 적응적인 평균 모양, (b)는 모델 조정 후의 모양, (c)는 적응적인 평균 모양을 이용한 제한 조건 검사 및 복원 후의 모양이다. (c)를 통해 적응적인 평균 모양을 이용할 경우에는 제한 조건 검사 및 복원 과정 후 올바른 모양이 생성됨을 확인할 수 있다.

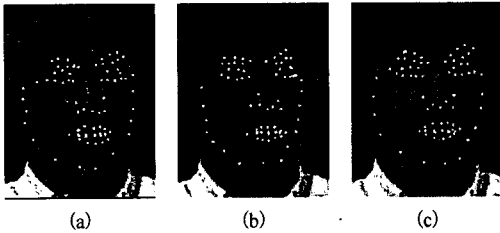
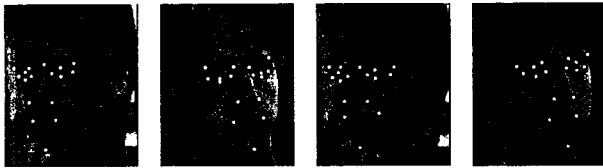


그림 4. 적응적인 평균 모양에 의한 특징점의 위치 변화

특징점 추출 성능 평가를 위한 테스트 데이터 셋은 CMU PIE [6] 얼굴 데이터 셋을 이용하여 구성하였으며, 안경 쓴 사람을 제외한 40명에 대해서 각 인물당 그림 5와 같은 4장의 영상으로 총 160장으로 이루어졌다.



(a) B_W_05 (b) N_W_05 (c) N_W_11 (d) S_W_11

그림 5. 성능 평가를 위해 선택된 landmark점들

특징점 추출 성능은 그림 5에 나타낸 17개 특징점들의 평균 위치 오차를 이용하여 평가하였으며, 표 1은 본 논문에서 제안한 방법(A)을 사용한 경우와 고정된 평균 모양(B)을 사용한 경우의 특징점 추출 성능을 비교하여 보여준다. 표 1을 통해서 적응적인 평균 모양 방법의 성능이 고정된 평균 모양 방법보다 뛰어남을 확인할 수 있다. 그림 6은 특징점 추출의 결과를 보여준다.

표 1. 고정된 평균 모양과 본 논문에서 제안한 방법의 성능 비교

포즈	B_W_05	N_W_05	N_W_11	S_W_11
방법 A의 평균 위치 오차(pixel)	2.01	1.51	1.47	1.95
방법 B의 평균 위치 오차(pixel)	4.08	5.37	3.54	5.16

특징점 추출 시간은 팬티엄 4, 2.4Ghz 컴퓨터를 사용하여 측정하였으며, 500~900ms의 시간이 소요된다. 이중 적응적인 평균 모양의 계산을 위해 사용되는 시간은 400~800ms이다.

5. 결론

본 논문에서는 포즈가 취해진 얼굴에 ASM이 적용될 때 발생하는 문제점을 분석하고 이를 해결하기 위한 방법으로 적응적인 평균 모양을 제안하였다. 적응적인 평균 모양은

최적화 방법인 유전자 알고리즘을 이용하여 구하였으며 목적 함수는 통계학상의 외관 모델 정보를 이용하여 구성하였다.

실험 결과, 적응적인 평균 모양 방법을 ASM에 이용하여 특징점 추출을 수행한 경우에는 고정된 평균 모양 방법을 이용함으로써 생성되던 모양 모델의 제한 조건 및 복원 과정 문제점을 해결 할 수 있었다. 하지만 적응적인 평균 모양을 구하기 위해 사용된 유전자 알고리즘으로 인하여 특징점 추출의 수행 시간이 다소 많이 증가하여 속도 향상을 위한 향후 연구가 필요하다.

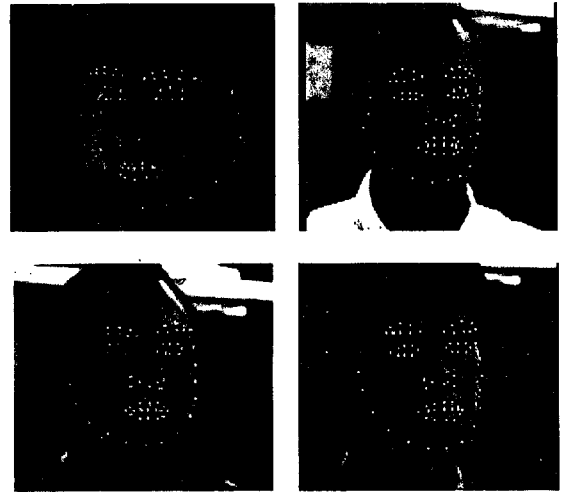


그림 6. 본 논문에서 제안한 방법에 의한 특징점 추출 결과

참고 문헌

- [1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active Shape Models - Their Training and Application," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, January 1995.
- [2] S. Romdhani, S. Gong, and A. Psarrou. "A multi-view nonlinear active shape model using kernel pca," *British Machine Vision Conference*, pp. 483-492, 1999.
- [3] Changbo Hu, Rogerio Feris and Matthew Truk, "Active Wavelet Networks for Face Alignment," *British Machine Vision Conference*, 2003.
- [4] D.E.Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
- [5] T.F. Cootes and C.J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," *Proceedings of SPIE 4322*, pp. 236-248, 2001.
- [6] T.Sim, S.Baker, and M.Bast, *The CMU Pose, Illumination, and Expression (PIE) database of human faces*. The Robotics Institute, Carnegie Mellon University, 2001.