얼굴 영역 추적과 입 검출을 이용한 AAM 얼굴 모양 파라미터 추정

최권택°, 변혜란
연세대학교 컴퓨터과학과
{choikt°, hrbyun}@cs.yonsei.ac.kr

Active Appearance Model Face Shape Estimation Using Face Region Tracking and Mouth Detection

KwunTaeg Choi°, Hyeran Byun
Yonsei University

요 약

얼굴의 특징점 추적은 많은 응용프로그램에서 사용된다. AAM기반의 접근방식은 정교한 얼굴 특징점 정보를 제공하지만 정확한 특징 점 추출을 위해 얼굴 모양 파라미터 초기화 문제와 연속 영상에서 얼굴의 이동이 클 경우 모션 보정에 대한 문제가 여전히 남아있다. 이러한 문제를 풀기 위해 본 논문에서는 CAMShift를 사용해 얼굴 영역을 추적하고, 얼굴 영역 내에서 입을 검출함으로써 AAM 검색을 위한 얼굴 모양 파라미터를 추정하는 방법을 제안한다. 기존 알고리즘과의 비교 실험을 통해 얼굴의 움직임이 심한 상황에서도 제안하는 알고리즘의 성능이 매우 우수함을 확인할 수 있었다.

Keywords: 얼굴 추적, CAMShift, AAM, 입 검출, 적응적 피부색 모델

1. 서 론

얼굴을 검출하고 추적하는 기술은 이동로봇, 감시 시스템, 인간과 로봇간의 상호작용 등 많은 응용시스템에서 연구되고 있는 분야 중 하나이다. 현재의 비전 기술을 이용하여 실시간으로 얼굴을 검출하고, 추적하는 일은 어려운 일임에도 불구하고, 컴퓨터 성능의 발달로 인해 영상처리 기법의 발전과 더불어 이에 관한 연구가 활발히 진행되고 있다.

얼굴 추적은 영역 추적 방법과 세부 특징점 추적 방법으로 그 유형을 나눌 수 있다. Qian[1]은 피부색 정보를 이용해 얼굴 영역을 추적하였고, Birchfield[2]는 얼굴의 색과 형태 정보 모두를 이용한 방법을 제안했다. 얼굴 영역 추적방법과 비교해 Cootes[3]가 제안한 AAM 알고리즘은 정교한 얼굴 특징점 정보를 추출할 수 있기 때문에 추적과 인식에서 많이 사용되고 있고 좋은 성능을 보여주고 있다. 그러나 초기 얼굴 모양에 대한 사전 지식이 주어져야 한다는 문제를 가지고 있다. 이런 사전 지식은 수작업이나 정면 얼굴 검출방법을 이용해 주어질 수 있지만 이를 얼굴 추적에 적용할 때 또 다른 문제가 발생한다. 얼굴의 이동이 큰 경우 이전 얼굴 모양의 결과를 사전 정보로 이용할 수 없기 때문에 재초기화가 필요한데 현재 영상이 정면 얼굴이 아닌 경우 얼굴 모양 파라미터에 대한 재초기화가 불가능하기 때문에 추적 성능은 낮아진다.

본 논문은 얼굴 모양 파라미터를 자동으로 추정하는 방법을 제안해 초기 얼굴 파라미터 설정과 얼굴의 이동이 클 경우 얼굴 모양 파라미터를 재 추정함으로써 동영상에서 정교한 얼굴 특징점을 안정적으로 추적할 수 있는 시스템을 개발하였다.

2. 칼라정보를 이용한 얼굴 영역 추적

본 논문의 얼굴 영역 추적 알고리즘은 칼라정보를 이용한다. 추적하고자 하는 얼굴의 색 분포를 추정하고 이 정보를 다음 프레임에서 객체의 후보 영역을 찾는 특징 정보로 사용해 얼굴 영역을 추적하였다.

2.1. CAMShift를 이용한 얼굴영역 추적

색 정보에 대한 히스토그램 기반 추적은 많은 추적 시스템에서 사용되는 방법이다. 이러한 방법 중 MeanShift[4]와 CAMShift[5] 알고리즘이 가장 널리 사용된다. 이들 알고리즘은 추적 대상 영역을 특정 색 공간으로 바꾸고 커널 기반으로 히스토그램을 구한다. 이 분포를 가지고 다음 프레임 영상에 히스토그램 정보를 역투영시키고, 이 분포의 모드를 반복적으로 찾기 위해 Mean-Shift 방법을 사용한다. 그러나 이 알고리즘은 추정 영역의 크기가 고정되어 있다는 단점이 있다. 이 문제를 해결하기 위해 CAMShift(Continuously Adaptive Mean Shift) 방법이 제안되었다. 이 알고리즘은 커널의 크기를 조절함으로써 추적대상의 크기가 변하는 경우에도 대상 영역을 성공적으로 추적할 수 있다. 얼굴의 색 분포를 이용하는 이들 알고리즘은 첫 프레임에서 추적할 얼굴의 영역을 설정해야 하는데 본 논문에서는 정면 얼굴검출을 위해 Viola[6]가 제안한 방법을 이용해 추적 영역을 추정하였다. 이 영역을 YCrCb 칼라 공간에서 각각 16개의 빈으로 나누고 Cr 및 Cb 성분을 이용해 색 분포를 추정하고 역투영시킨 후 다음 영상에서 CAMShift알고리즘 사용해 얼굴영역의 위치를 계산했다.

2.2. 적응적 피부색 모델을 이용한 색 분포 추정

CAMShift 알고리즘은 추적할 영역의 색 분포를 이용하기 때문에 정확한 색 분포 추정은 매우 중요하다. 일반적으로 사각형 영역이나 타원 영역 정보를 이용하는데 이는 배경 잡음을 포함하기 때문에 정확한 색 분포 추정을 어렵게 만든다. 이러한 문제점을 해결

하기 위해 본 논문은 적응적 피부색 모델을 이용해 색 분포를 추정한다. 본 논문의 추적 알고리즘은 AAM을 이용해 정확한 얼굴 특징점을 추출하기 때문에 주어진 영상에서 피부색 영역을 정확하게 획득할 수 있다. 그림 1은 이를 위한 방법을 보여주고 있다. 이전 프레임에서 얻어진 영상과 얼굴 모양을 이용해 내부 영역을 채움으로써 정확한 피부색 마스크를 생성해 낼 수 있고, 이 영역에 해당하는 화소만을 이용해 정확한 색 분포를 추정할 수 있다. 즉 전 프레임에서 AAM을 통해 얻어진 얼굴 모양으로부터 피부색 마스크를 생성하고, 색 분포를 매 프레임 변경함으로써 매우 정확한 색 분포를 추정할 수 있으므로 CAMShift의 성능이 향상된다.

그림 1. 적응적 피부색 모델

3. 입 검출을 이용한 AAM 모델 파라미터 추정

AAM 알고리즘을 이용해 정확한 특징점을 추출하기 위해서는 초기 얼굴 모양 파라미터가 필요하다. 이를 위해 CAMShift 알고리즘을 이용해 얼굴 영역을 추적했고, 이로부터 영상 내에서 얼굴의 크기와 위치정보를 추정해낼 수 있다. 그러나 정확한 초기 얼굴 모양 파라미터는 좀더 정확한 얼굴 특징점을 추출하기 때문에 본 논문에서는 입 검출을 통해 AAM 검색을 위한 초기 얼굴 모양 파라미터를 좀더 정확하게 추정한다.

3.1. 입 검출

얼굴 영역 내에서 입 영역을 검출하가 위해 본 논문에서는 Jain[7]이 사용한 YCbCr 색 공간에서 Cb 및 Cr 성분을 이용한 방법을 사용하였다. 이 알고리즘은 입 주변에서 Cr 성분이 Cb 성분보다 높게 나타나는 통계적인 특성을 이용한다. 추정된 얼굴 영역을 YCbCr 색 공간으로 변환하고 Cr 및 Cb 성분을 분석해 입 영역 지도를 만든다. 입 영역은 매우 밝은 값으로 나타나기 때문에 역치 값을 적용해 입 영역의 중심좌표를 얻어 낼 수 있다(그림 2). 입 위치를 검출한 후 정교한 특징점을 추출하기 위해 AAM 검색을 수행한다.

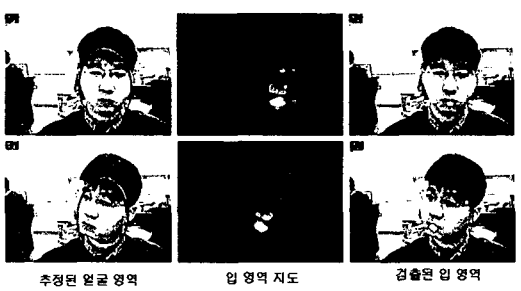

그림 2. 입 영역 검출

3.2. AAM 검색을 위한 얼굴 모양 파라미터 추정

그림 3 (a)는 얼굴 영역 정보(위치, 크기)만을 이용해서 초기 얼굴 모양을 추정한 것이고, (b)는 (a)에 대한 AAM 검색후의 결과이다. (c)는 검출된 입의 중심 좌표와 평균 얼굴 모양에서 입의 위치가 일치되도록 위치 이동을 해서 얼굴 모양을 추정한 것이다. (d)는 (c)에 대한 AAM 검색후의 결과이다.

그림 3. 입 검출을 이용한 얼굴 모양 파라미터 추정

(a)의 경우 얼굴이 정면 영상으로부터 옆으로 많이 돌아갔기 때문에 얼굴 영역만을 이용해 추정한 얼굴 모양은 실제 영상과 많은 차이를 보이고 이는 AAM 검색 실패로 이어진다. 그러나 (c)의 경우처럼 영상에서 입 부분을 일치시키면 얼굴 모양이 실제 영상에서 벗어나긴 하지만 AAM 검색을 통해 정확한 특징정을 추출할 수 있기 때문에, 입 검출을 이용함으로써 좀더 정확한 얼굴 특징점을 추출해 낼 수 있다.

4. 실험 및 결과

본 논문에서 제안한 방법은 Windows XP, Visual C++ 6.0 을 이용하여 Pentium IV 2.8GHz 하드웨어 상에서 구현되었다. 실험에서 사용된 모든 영상은 320x240 크기이며 USB 카메라를 이용해 캡쳐하였다.

4.1. AAM 학습

AAM을 학습시키기 위해 5명에 대해 각각 정면과 상하좌우, 대각선 방향과 3가지 조명 설정을 고려해 각각 27장을 사용하여 총 135장의 영상을 이용하였다(그림 4).

그림 4. 학습 데이터

일반적으로 AAM을 학습시킬 때 135장의 영상 전부를 사용하지만 본 논문에서는 각 개인별로 총 5개의 AAM을 학습시켰다. 이와 같은 학습 방법은 데이터베이스에 등록되어 있는 사람을 인식하기 위해서 얼굴을 추적하는 시스템에서 성능 향을 위해 사용될 수 있다. 비디오 영상이 입력되면 첫 프레임에서 5개의 학습된 AAM중 얼굴 영역에 대한 텍스쳐 샘플 에러가 가장 작은 모델을 하나 선택하고 그 이후는 해당 AAM 모델을 계속 사용하기 때문에 속도상의 성능저하도 발생하지 않는다. 이러한 방법은 AAM 모델 검색에 대한 문제 공간의 크기를 줄여 좀더 정확한 얼굴 특징점 추적을 할 수 있게 한다.

4.2. 실험 결과

제안된 알고리즘의 성능을 확인하기 위해 3개의 동영상을 가지고 실험하였다. 첫 번째 실험 동영상은 얼굴 움직임과 변화가 작고, 세 번째 실험 동영상의 경우는 매우 심하다. 각각의 동영상은 200 프레임으로 구성되어 있고, 정면 얼굴 영상으로부터 시작한다. 그림 5의 첫 번째 줄은 얼굴 영역 추적과 입 검출을 통해 얼굴 모양을 추정한 결과이고, 두 번째 줄은 AAM 검색 후의 결과이다. 얼굴의 움직임이 심한 경우에도 매우 안정적인 결과를 보여줌을 확인할 수 있었다.

그림 5. 실험 결과

기존 방법과의 비교를 위해 세 가지 방법을 사용하였다.

(a) 직접 초기 얼굴 모양 추정
(b) 정면 얼굴 검출을 이용한 AAM 얼굴 모양 추정
(c) CAMShift와 입 검출을 이용한 AAM 얼굴 모양 추정

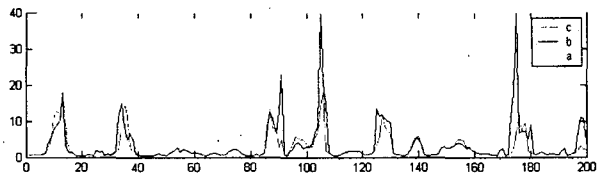

그림 6. 얼굴 영역 텍스쳐에 대한 샘플 에러

그림 6은 얼굴 이동 및 변화가 심한 테스트 동영상에 대해서 3가지 방법을 이용해 AAM 검색을 수행한 결과이다. y 축은 AAM 모델 검색 후 얼굴 영역에 대한 텍스쳐 샘플 에러값이다. 이 값이 0에 가까울수록 검색결과는 매우 좋다. 실험에서 에러가 100이 넘어가면 검색 실패를 의미한다. 에러가 40보다 클 경우 그래프에서는 40인 것으로 표현했다. (a) 방법은 모션 보정이 없기 때문에 얼굴 이동이 클 경우 AAM 모델 검색에 실패하게 되고 이후 계속 추적 실패로 이어져 처음 세 프레임 이후 얼굴 특징점 추적이 전혀 되고 있지 않다. (b) 방법의 경우 AAM 검색에 실패하더라도 해당 프레임에서 정면 얼굴이 검출될 수 있다면 정확한 추적이 가능하나 얼굴이 기울여져 있는 경우 역시 지속적인 추적이 불가능하기 때문에 제안한 방법보다 에러가 높다. 제안한 방법은 AAM 검색 실패 시 CAMShift를 이용해 얼굴 영역과 입 검출을 이용해 얼굴 모양 파라미터를 다시 추정하기 때문에 얼굴의 움직임이 매우 큰 경우에도 전체적인 추적 실패 없이 매우 안정적인 결과를 보여준다. 표1은 3개의 동영상에 대한 추적 결과이다. 각각 200 프레임으로 구성되어 있는 실험 동영상에 대해 3가지 방법에 대한 추적률을 계산했다. 얼굴의 눈, 코, 입에 대한 특징점을 제대로 찾은 경우만 추적에 성공한 것으로 판단했다.

표1. 얼굴 특징점 추적률

얼굴 이동 및 변화 정도	(a)	(b)	(c)
작음	68.5%	93.5%	95.5%
중간	54.5%	82.9%	91.5%
빠름	1.5%	70.5%	88.5%

5. 결론

본 논문에서는 CAMShift를 이용한 얼굴 영역 추적과 입 검출을 통해 AAM 검색을 위한 얼굴 모양 파라미터를 추정하는 방법을 제안했다. 정확한 얼굴 모양 초기화를 수행할 수 있기 때문에 제안하는 알고리즘은 연속 영상에서 얼굴 움직임과 변화가 매우 큰 경우에도 안정적인 추적결과를 보여주었다.

참고 문헌

[1] RJ Qian, MI Sezan, and KE Matthews, "A Robust Real-Time Face Tracking Algorithm", *Int. Conf. on Image Processing*, Vol. 1, pp. 131--135, 1998

[2] S. Birchfield, "Elliptical Head Tracking Using Intensity Gradients and Color Histograms", *IEEE Conference on Computer. Vision and Pattern Recognition*, 1998.

[3] T. F. Cootes, G.J. Edwards, and C.J. Taylor, "Active appearance model", *Fifth European Conf. Computer Vision*, 1998

[4] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift", *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 142149, 2000

[5] J. G. Allen, R.YD Xu, J. S. Jin "Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces," *Pan-Sydney Area Workshop on Visual Information Processing*, pp.3-7, 2003

[6] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", *IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawaii, USA, 2001

[7] R.-L. Hsu, M. Abdel-Mottaleb and A. Jain "Face detection in color images." *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May 2002