

SIFT 특징을 이용하여 중첩상황에 강인한 AAM 기반 얼굴 추적

엄 성 은^{*} · 장 준 수^{**}

요 약

얼굴추적은 3차원 공간상에서 머리(head)와 안면(face)의 움직임을 추정하는 기술로, 얼굴·표정·감정인식과 같은 상위 분석단계의 중요한 기반기술이다. 본 논문에서는 AAM 기반의 얼굴추적 알고리즘을 제안한다. AAM은 변형되는 대상을 분할하고 추적하는데 광범위하게 적용되고 있다. 그러나 여전히 여러 가지 해결해야 할 제약사항들이 있다. 특히 자체중첩(self-occlusion)과 부분적인 중첩, 그리고 일시적으로 완전히 가려지는 완전중첩 상황에서 보통 국부해에 수렴(local convergence)하거나 탈산하기 쉽다. 본 논문에서는 이러한 중첩상황에 대한 AAM의 강인성을 향상시키기 위해서 SIFT 특징을 이용하고 있다. SIFT는 일부 영상의 특징점으로 안정적인 추적이 가능하기 때문에 자체와 부분중첩에 효과적이며, 완전중첩의 상황에도 SIFT의 전역적인 매칭성능으로 별도의 재초기화 없이 연속적인 추적이 가능하다. 또한 추적과정에서 큰 자세 변화에 따른 움직임을 효과적으로 추정하기 위해서 다시점(multi-view) 얼굴영상의 SIFT 특징을 온라인으로 등록하여 활용하고 있다. 제안한 알고리즘의 이러한 강인성은 위 세 가지 중첩상황에 대해서 기존 알고리즘들과의 비교실험을 통해서 보여준다.

키워드 : 얼굴추적, AAM(Active Appearance Model), SIFT(Scale Invariant Feature Transform), 중첩문제, 3차원 자세추정, 온라인 특징 등록

Robust AAM-based Face Tracking with Occlusion Using SIFT Features

Eom, Sungeun^{*} · Jang, Jun-Su^{**}

ABSTRACT

Face tracking is to estimate the motion of a non-rigid face together with a rigid head in 3D, and plays important roles in higher levels such as face/facial expression/emotion recognition. In this paper, we propose an AAM-based face tracking algorithm. AAM has been widely used to segment and track deformable objects, but there are still many difficulties. Particularly, it often tends to diverge or converge into local minima when a target object is self-occluded, partially or completely occluded. To address this problem, we utilize the scale invariant feature transform (SIFT). SIFT is an effective method for self and partial occlusion because it is able to find correspondence between feature points under partial loss. And it enables an AAM to continue to track without re-initialization in complete occlusions thanks to the good performance of global matching. We also register and use the SIFT features extracted from multi-view face images during tracking to effectively track a face across large pose changes. Our proposed algorithm is validated by comparing other algorithms under the above 3 kinds of occlusions.

Keywords : Face tracking, AAM(Active Appearance Model), SIFT(Scale Invariant Feature Transform), Occlusion Problem, 3D Pose Estimation, Online Feature Registration

1. 서 론

얼굴추적은 3차원 공간상에서 얼굴의 움직임을 추정하는 기술로, 머리(head)의 전체적인 움직임과 안면(face)의 부분적인 움직임으로 분리할 수 있다. 이러한 움직임의 정확한 추정은 얼굴·표정·감정인식과 같은 상위 분석단계의 중요한 기반기술에 해당하며, 자연히 이러한 인식을 토대로 하는 다양한 응용분야에서도 중요한 역할을 하게 된다 [1-6]. 또한 얼굴

* 본 연구를 위해서 실험을 도와주신 카네기멜론대학의 로봇연구소 동료들에게 깊은 감사를 드린다. 특히 비교실험을 위해서 AAM 코드를 제공해주신 Simon Lucey와 Takoe Kanade 교수님께 감사를 드린다. 이 논문은 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구이다(기초과학 분야 : KRF-2007-357-D00194).

† 정 회 원 : 카네기멜론대학 로봇연구소 Post Doc.

** 정 회 원 : LG전자 책임연구원

논문접수: 2010년 7월 2일

수정일: 2010년 8월 16일

심사완료: 2010년 8월 16일

은 움직임 그 자체의 정보와 더불어 그 사람의 내·외적인 의미를 내포하고 있기 때문에, 영상뿐 아니라 심리학 연구자들에게도 오래전부터 관심이 대상이 되고 있다[7-9].

얼굴추적은 앞서 언급했듯이 얼굴의 전역적인 움직임과 함께 부분적인 변형을 고려해야하기 때문에, 추적과정에서 변형 가능한 모델을 이용하는 접근방식이 필요하다. 그러한 접근방식 중 AAM(Active Appearance Model)이 가장 폭넓게 활용되고 있다[10, 11]. 이 모델의 기본적인 아이디어는 변형 가능한 대상을 형태와 텍스처(appearance) 두 정보로 충분히 모델링한 다음, 그 대상이 임의로 주어졌을 때 그 대상에 가장 근접한 모델 인스턴스(instance)를 생성해 주는 것이다. 즉 AAM은 얼굴변화에 따른 최적의 모델 인스턴스를 연속적으로 생성함으로써 얼굴추적에 바로 활용이 가능하다.

현재 AAM 모델은 추적의 정확도와 처리속도 측면에서 많은 향상을 보여주고 있다. 그러나 AAM을 실질적으로 활용하는데 여전히 여러 가지 해결해야할 문제들이 있다[1]. 특히 큰 자세(pose) 변화에 따른 자체중첩과 그 밖에 일반적으로 발생하는 다른 대상에 의한 부분적인 혹은 전체적인 중첩상황에 매우 취약한 단점이 있다. 이러한 상황에서 AAM 모델은 보통 국부해에 수렴하거나 발산을 하게 되며, 재초기화를 위해서 별도의 얼굴검출기가 필요로 한다.

이 문제에 대한 접근방식은 크게 두 가지로 분류할 수 있다. 첫째 현재 AAM을 위 중첩상황에 강인하게 추적이 가능하도록 모델 혹은 fitting 알고리즘을 개선하는 접근방식[12-15], 둘째 현 모델을 그대로 독립적으로 유지하면서 다른 기술을 활용하여 강인성을 향상시키는 방식이다[16]. 먼저 [12]에서는 중첩문제를 명시적으로 다루고 있지만, 좀 더 정확히 자세변화를 추정하기 위해서 기존 2D AAM의 최적화 함수에 3차원 형태정보를 제약조건으로 추가하여 사용하고 있다. 반면에 [13]에서는 3차원 형태정보 대신에 강인성 오차함수(robust error function)를 추가하여 그에 대한 새로운 fitting 알고리즘을 제안하였다. 여기에서 중요한 오차함수에 대한 선정문제는 [14]에서 다루어지고 있다. 3차원 강체모델을 이용하는 [15]에서는 보다 넓은 자세변화를 반영하기 위해서 다시점(multi-view) AAM 모델을 사용하고 있으며, 부분중첩에 대해서는 [13]과 유사한 fitting 알고리즘을 이용하고 있다.

한편 중첩에 대한 강인성을 향상시키기 위한 두 번째 접근방식인 [16]은 두 가지 모델, 즉 실린더 모델과 AAM 모델을 이용하여 자세변화에 강인한 얼굴추적 알고리즘을 제안하고 있다. 먼저 얼굴의 전체적인(rigid) 움직임을 실린더 모델을 이용하여 3차원 자세를 추정하고 있다[17]. 그 다음 그 자세정보를 이용하여 AAM을 초기화하고 나머지 지역적인(non-rigid) 움직임을 추정하고 된다. 이 알고리즘은 실린더 모델을 통해서 AAM이 추적할 수 있는 자세변화의 범위를 증가시킴으로써 자체중첩에 대한 강인성을 향상시키고 있다.

본 논문에서는 [16]과 유사한 두 번째 접근방식을 이용하

여 자세변화에 따른 자체중첩 뿐만 아니라, 다른 대상에 의한 부분적인 그리고 전체적인 중첩 상황에서도 강인성을 향상시키고자 한다. 이를 위해서 우리는 SIFT(Sale Invariant Feature Transform) [18] 특징을 이용한 추적 알고리즘을 제안한다. SIFT는 일부 특징점이 손실된다하더라도 남은 특징점을 이용하여 안정적인 추적이 가능하기 때문에 자체중첩과 부분중첩 상황에 효과적으로 활용할 수가 있다. 그리고 얼굴이 완전히 사라진 다음 다시 나타나는 경우에 보통 별도의 검출기를 이용하여 재초기화가 필요한데 제안한 알고리즘에서는 SIFT의 전역적인 매칭성능으로 별도의 재초기화 없이 연속적인 추적이 가능하다. 또한 추적과정에서 큰 자세변화에 따른 움직임을 효과적으로 추정하기 위해서 다시점 얼굴영상의 SIFT 특징을 온라인으로 등록하여 활용하고 있다[19]. 구체적인 알고리즘은 제 3장에서 기술하고 있으며, 그 이전에 제 2장에서는 AAM 모델의 정의를 간략히 소개한다. 마지막 장인 제 4장에서는 제안한 알고리즘의 중첩에 대한 강인성을 비교실험을 통해서 보여준다.

2. Active Appearance Model (AAM)

AAM은 얼굴과 같이 변형이 가능한 대상을 형태(shape)와 텍스처(appearance) 정보를 이용하여 선형적으로 표현한 모델이다[10, 11]. 일반적으로 형태와 텍스처 두 정보는 독립적으로 분리되어 사용된다. 먼저 형태는 다음과 같이 v 개의 2차원 좌표로 표현 된다:

$$s = (x_1, y_1, x_2, y_2, \dots, x_v, y_v)$$

위와 같이 표현된 형태는 다시 주성분 분석법(PCA: Principal Component Analysis)을 통해서 평균형태와 n 개의 형태벡터로 표현이 가능하다:

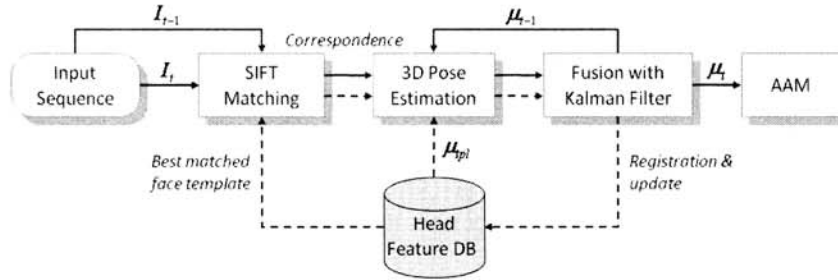
$$s = s_0 + \sum_{i=1}^n p_i s_i$$

여기에서 각 형태벡터의 가중치에 해당하는 p_i 값이 바로 형태 변화를 결정짓는 파라미터에 해당된다.

AAM 모델의 두 번째 구성요소인 텍스처는 앞서 정의된 평균형태 상에 존재하는 영상값(intensity)으로 표현된다. 즉 평균형태 s_0 의 좌표로 이루어진 mesh 위에 존재하는 영상이 텍스처에 해당된다. 이 텍스처 A 는 형태와 마찬가지로 평균값 A_0 와 주성분 벡터 A_i 들로 표현이 가능하며, 각 벡터들의 가중치(λ_i)에 의해 임의의 텍스처가 결정된다:

$$A(x) = A_0(x) + \sum_{i=1}^m \lambda_i A_i(x) \quad \forall x \in s_0$$

이와 같이 정의된 AAM 모델을 이용한 얼굴추적은 결국 매 프레임 얼굴의 움직임에 대응하는 형태와 텍스처 파라미



(그림 1) SIFT 특징을 이용한 AAM 기반 얼굴추적 알고리즘

터를 계산해주는 과정이다. 이는 다시 말해서 두 종류의 파라미터 값을 통해서 얻어진 AAM 모델과 실제영상과의 차를 최소화(fitting)하는 과정이다. 최소화 과정에서 그 오차값이 일정범위 내에 수렴하면 추적 성공으로 간주하고 다음 프레임으로 이동하게 되며, 다음 프레임에서의 초기위치는 보통 그 이전 프레임의 최종 수렴한 위치를 이용하게 된다.

3. SIFT 특징을 이용한 AAM 기반 얼굴추적

본 논문에서 제안하는 얼굴추적 알고리즘은 (그림 1)에서 보여주듯이 크게 두 부분으로 이루어져 있다: 1) 얼굴의 전체적인 움직임, 즉 머리(head)의 3차원 자세추정과 2) 이 자세를 토대로한 안면(face)의 지역적인 움직임 추정.

얼굴의 3차원 자세추정은 얼굴의 기하학적 모델을 이용하여, X, Y, Z 세 축에 대한 변이(translation)와 회전각을 계산하는 과정이다. 얼굴은 보통 간단한 평면(planar) 모델에서부터 타원(ellipsoidal), 실린더(cylinder), 그리고 좀 더 복잡한 모델로 표현될 수 있는데[20-23], 여기에서는 모델의 복잡도, 계산량, 정확도 등을 고려하여 머리 정면과 측면만을 포함하는 반(half) 실린더 모델을 이용하고 있다.

실린더 모델을 이용하여 얼굴자세를 추정하기 위해서는 먼저 기준이 되는 얼굴템플릿이 필요하다. 기준 템플릿이 초기영상으로 고정되면 조명과 자세변화가 큰 경우 추정이 어려워지기 때문에 고정 템플릿 대신 일정기준에 따라 템플릿을 갱신/등록하는 방식이 바람직하다.

제안한 알고리즘에서는 일정한 얼굴 각도변화에 따라 온라인으로 여러 개의 템플릿을 등록하고 갱신하는 방식을 취하고 있으며, 자세 추정시 현재 등록된 템플릿 중 가장 잘 매칭이 되는 템플릿을 활용하고 있다. 특히 중첩(occlusion)에 대한 강인성을 향상시키기 위해서 SIFT 특징점을 이용하고 있다[18].

기본적인 흐름은 다음과 같다. 먼저 연속적인 두 프레임에서 SIFT 특징점을 추출한 다음 서로 대응관계를 파악한다 (그림 1의 실선). 이 대응관계는 연속적인 두 프레임에서 뿐만 아니라, 얼굴 특징 데이터베이스(head feature DB)에 등록된 영상과 현재 프레임에서도 계산이 된다 (그림 1의 점선). 이렇게 얻어진 두 대응관계로부터 각각 3차원 자세를 추정하게 되며, 그 후 칼만 필터(Kalman filter)를 이용해서 현재 프레임의 얼굴자세를 추정하게 된다. 끝으로 이 자세

정보는 AAM 모델에 인가되어 최종적인 얼굴의 움직임을 추정하게 된다. 각 단계에 대한 구체적인 설명은 다음 각 절에서 기술한다.

3.1 온라인 얼굴특징 등록

얼굴 특징 데이터베이스는 얼굴자세를 추정하는 과정에서 실린더 모델이 X, Y축 방향으로 일정한 크기만큼 회전할 경우 그 각각의 자세에 대한 특징을 저장하게 된다. 특징으로는 거짓-긍정률(false positive rate)이 낮고 넓은 영역에 대한 매칭효과가 좋은 SIFT를 사용하고 있으며, X, Y축에 대해 각각 10도 간격으로 저장하고 있다. Z축(in-plane)에 대한 회전은 SIFT에서 자체적으로 반영이 되기 때문에 별도로 데이터베이스를 형성할 필요가 없다. 이러한 온라인 다시점 접근방식을 이용하면 초기 정면얼굴의 특징에 국한하지 않고 여러 각도의 얼굴영상에 대한 특징들을 활용할 수가 있다. (그림 2)는 얼굴자세를 추정하는 과정에서 온라인으로 등록된 얼굴영상과 그 특징점을 보여주고 있다.



(그림 2) 온라인으로 등록된 Y축 방향의 얼굴영상과 SIFT 특징

3.2 3차원 얼굴자세 추정

얼굴자세 추정은 두 가지 경로를 통해서 이루어진다. 먼저 현재 프레임과 바로 이전 프레임에서 추출된 SIFT 특징을 이용해서 서로간의 대응관계를 파악한다. 여기에서 각 특징점은 128차원으로 이루어져 있다. 두 번째 경로는 현재 프레임과 얼굴 데이터베이스의 특징을 이용한 대응관계 파악이다. 데이터베이스 내에서 선정된 특징은 현재 프레임과 가장 잘 매칭이 되는 특징에 해당된다.

이렇게 얻어진 특징점들간의 2차원 대응관계는 얼굴모델인 실린더 상의 3차원 대응관계로 표현할 수 있으며 그로부터 실린더의 3차원 움직임을 추정할 수가 있다. 그림 3은 그 과정을 보여주고 있다. 먼저 시간 $t-1$ 시점에서 실린더의 포즈(μ_{t-1})가 주어졌을 때 그 프레임상의 2차원 좌표(p_{t-1})는 실린더 상의 3차원 좌표 $X_{t-1} = [x_{t-1}, y_{t-1}, z_{t-1}]^T$

로 변환할 수 있다. 그리고 시간 t 시점으로 움직임 $\Delta\mu = [\Delta\theta_x, \Delta\theta_y, \Delta\theta_z, \Delta x, \Delta y, \Delta z]^T$ 가 발생했을 때 X_{t-1} 은 다음과 같은 관계식을 통해 X_t 로 이동하게 된다[24].

$$\begin{bmatrix} x_t \\ y_t \\ z_t \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -\Delta\theta_z & \Delta\theta_y & \Delta t_x \\ \Delta\theta_z & 1 & -\Delta\theta_x & \Delta t_y \\ -\Delta\theta_y & \Delta\theta_x & 1 & \Delta t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \\ 1 \end{bmatrix}$$

이 X_t 는 다시 투시 투영(perspective projection)을 통해 다음과 같이 영상상 2차원 좌표로 투영된다:

$$p'_t = \frac{\begin{bmatrix} x_{t-1} - y_{t-1}\Delta\theta_z + z_{t-1}\Delta\theta_y + \Delta t_x \\ x_{t-1}\Delta\theta_z + y_{t-1} - z_{t-1}\Delta\theta_x + \Delta t_y \end{bmatrix}}{f} \cdot \begin{bmatrix} -x_{t-1}\Delta\theta_y + y_{t-1}\Delta\theta_x + z_{t-1} + \Delta t_z \end{bmatrix}$$

여기에서 f 는 카메라의 초점거리로서, 실린더 움직임 추정에는 크게 중요하지 않다 [25]. 결국 위 두 식으로부터 시간 t 시점으로 투영된 p'_t (expected location)과 p_t (observed location)간의 거리를 최소화하는 움직임벡터 $\Delta\mu$ 가 바로 그 시점에서의 실린더 움직임에 해당한다.

3.3 최종 얼굴움직임 추정

위 방식으로 얻어진 각각의 얼굴자세는 칼만필터를 통해서 최종적으로 추정된다. 칼만필터를 이용해 이 두 데이터를 통합(fusion)하기 위해서는 해당 필터의 두 방정식, 즉 상태변이 방정식(state transition equation)과 관측 방정식(observation equation)이 필요로 하며, 그 이후는 칼만필터의 기본 과정을 따른다.

먼저 상태변이 방정식을 다음과 같이 정의하였다:

$$\mu_t = \mu_{t-1} + u_{t-1} + \alpha_{t-1}$$

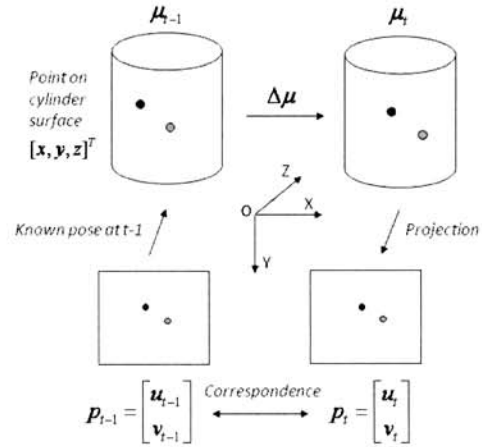
여기에서 u 는 첫 번째 경로 즉 연속적인 프레임사이의 자세변화에 해당하는 제어입력이며, α 는 프로세스 잡음(process noise)을 나타낸다. 이 방정식은 얼굴자세 μ 의 변이를 기술하고 있다. 보통 상태변이 방정식은 연속적인(smooth) 변이를 가정하고 있지만, 여기에서는 얼굴 움직임의 특성상 그 가정을 고려하지 않았다. 마지막 항에 추가된 잡음 α 은 정규분포 $N(0, Q_t)$ 으로 가정하며, covariance matrix Q_t 는 대각행렬로 각 원소가 p'_t (expected location)과 p_t (observed location)간의 오차에 의해서 결정된다.

한편 두 번째 관측 방정식은 다음과 같이 정의하였다.

$$h_t = \mu_t + \beta_t$$

여기에서 h 는 얼굴특징 데이터베이스를 이용해서 계산된 얼굴자세에 해당하며, β_t 는 정규분포 $N(0, R_t)$ 특성을 갖는 관측 잡음에 해당한다. R_t 는 상태변이 방정식의 Q_t 와 동일한 방식으로 계산된다.

이렇게 칼만필터를 통해서 얻어진 μ_t 는 얼굴의 전역적인 움직임을 나타내기 때문에, 이 자세를 토대로 매 프레임 AAM를 초기화할 수 있다. AAM 초기화는 앞서 (그림 3)에서 설명했던 과정을 따른다. 즉 시간 $t-1$ 시점에서 AAM의 2차원 형태좌표를 동일시간의 실린더 위치에 투영한 다음, 위에서 얻어진 움직임벡터($\Delta\mu$)만큼 이동하고 다시 2차원 영상으로 투영시켜주면 된다. 그 이후의 지역적인 움직임을 위한 AAM fitting 과정은 표준적인 방법을 이용하고 있다[11].



(그림 3) 2차원 특징점간의 대응관계를 이용한 실린더 움직임 추정

4. 실험결과

본 논문에서 제안한 알고리즘의 중첩에 대한 강인성을 평가하기 위해서, 먼저 총 10명의 피실험자를 대상으로 실험 데이터를 만들었다. 각 피실험자에게는 중첩이 발생하는 다음과 같은 세 가지 시나리오를 요구하였다:

- 1) 자체중첩: X, Y, Z 각 축을 중심으로 $-70^\circ \sim 70^\circ$ 범위 내에서 일정한 속도의 움직임,
- 2) 부분중첩: 손이나 다른 물체에 의한 부분적인 중첩 혹은 시야에서 부분적으로 벗어남,
- 3) 완전중첩: 물체에 의한 전체적인 중첩 혹은 시야에서 완전히 벗어난 후 복귀.

그리고 실제 실험에 적용할 데이터와 AAM 모델을 생성(training)할 데이터를 분리하기 위해서, 피실험자당 각 2회씩 데이터를 생성하였다. 첫 번째 데이터를 실제 비교실험에 활용하였고, 두 번째 데이터를 AAM 모델을 생성하는데 활용하였다.

영상은 초당 30 프레임에서 320x240 크기로 총 2000 프레임내외로 구성이 되어있다. 특히 머리(실린더)의 3차원 움

직임과 안면(AAM)의 형태좌표에 대한 기준값을 용이하게 추출하기 위해서, 4개의 마커가 있는 머리밴드를 피실험자에 착용시킨 다음 12개의 MX-40 카메라가 장착된 Vicon 모션캡션 장비를 이용하였다[26, 27].

이렇게 생성된 실험 데이터를 토대로 본 실험에서는 세 종류의 얼굴추적 알고리즘을 비교실험 하였다. 먼저 실린더 정보를 전혀 사용하지 않고 로컬정보만을 이용하는 표준적인 AAM 알고리즘[11], 둘째 Xiao의 실린더 모델과 AAM 모델을 이용한 알고리즘[16], 그리고 마지막으로 본 논문에서 제안한 알고리즘을 적용하였다. 여기에서 공통적으로 사용하는 AAM 알고리즘은 서로 동일하게 적용하기 위해서 두 번째 비교 대상인 [16]에서 사용한 2D+3D AAM 알고리즘을 사용하였다[12]. 그리고 AAM 모델 생성은 각 실험 데이터별로 다양한 얼굴움직임이 반영된 60개 영상을 이용하였다.

알고리즘의 성능은 추적성공률을 이용하여 평가하였다. 추적성공률은 해당 중첩상황의 발생횟수에 대해 실제 추적에 성공한 비율로 계산하였다. 여기에서 중첩의 발생횟수는 중첩이 발생하고 있는 모든 프레임의 수가 아닌, 상황 즉 이벤트의 횟수이다. 먼저 각 중첩 상황별 알고리즘의 강인성을 비교평가하기 위해서 각각의 추적 성공률을 계산하였고, 그 다음 모든 중첩 상황에 대한 평균 추적 성공률을 계산하였다.

추적 성공여부는 AAM의 형태좌표와 기준형태좌표(ground-truth)간의 거리차의 합이 일정한 임계치를 넘어가는 경우 추적실패로 간주하였다. 본 실험에서 사용한 임계치는 시각적으로 큰 차이를 보이는 280으로 설정하였다; AAM의 형태모델은 총 70개의 좌표로 이루어져있기 때문에 좌표당 평균적으로 4픽셀 오차에 해당한다. 이 평가 기준으로 부터 발생할 수 있는 추적실패는 두 가지 경우가 있다. 즉 1) 얼굴영역 밖으로 발산하여 더 이상 수렴하지 못하는 경우와, 2) 비록 이후에 다시 수렴하지만 그 이전에 위 오차가 임계치를 넘어 실패하는 경우가 있을 수 있다. 두 번째 경우에는 이후 중첩상황에 대해 연속적으로 실험이 가능하지만 첫 번째 경우는 어렵기 때문에, 각 중첩상황 이전에

〈표 1〉 중첩 상황별 추적 성공률 비교평가

중첩 상황	발생 횟수	AAM [11]	Xiao+AAM [16]	제안한 알고리즘
자체	150	85 (56.7%)	110 (73.3%)	126 (84.0%)
부분	33	23 (69.7%)	26 (78.8%)	27 (81.8%)
완전	40	27 (67.5%)	24 (60%)	31 (77.5%)
총	223	135 (60.5%)	160 (71.8%)	184 (82.5%)

항상 기준 형태좌표를 이용하여 재 초기화를 해주었다.

총 10명의 실험 데이터에 대한 각 추적 알고리즘의 추적 성공률은 다음 <표 1>에서 보여주고 있다.

먼저 자체중첩은 포즈변화에 따른 자체적인 중첩현상이기 때문에, [16]에서 언급을 했듯이 실린더의 전역적인 정보를 이용한 방식이 표준적인 AAM 보다 향상된 결과를 보여주었다. 제안한 알고리즘과 두 번째 알고리즘의 성능차이는 추적 가능한 포즈변화의 크기에서 비롯되었다. 즉 (그림 4)의 세 번째 열에서 보여주듯이 제안한 알고리즘은 [16]에서 언급한 추적 가능한 포즈변화의 크기(X, Y, Z축에 대해 각각 약 $\pm 20^\circ$, $\pm 40^\circ$, $\pm 10^\circ$)를 벗어나는 경우에도 그 일부 얼굴의 SIFT 특징으로 추적가능성이 높아졌다.

부분중첩 실험에서는 포즈변화에 따른 영향을 최소화하기 위해서 대부분 정면에서 크게 벗어나지 않았고, 얼굴의 움직임도 최소화시켰다. 이러한 조건에서 세 알고리즘의 성능은 <표 1>에서 보여주듯이 여전히 실린더와 SIFT 특징을 이용한 방식에서 보다 안정적인 결과를 보여주었다. 그러나 알고리즘의 상대적인 성능차이는 자체중첩에 비해 감소하였다. (그림 5)의 2, 3, 4열 결과를 비교해 보면 AAM의 형태가 손의 의해 어느 정도 일그러지고 있지만 심각하게 벗어나지는 않고 있다. 즉, 위 실험 조건의 부분중첩에서는 얼굴의 전역적인 자세 정보의 기여도가 상대적으로 감소함을 알 수가 있다.

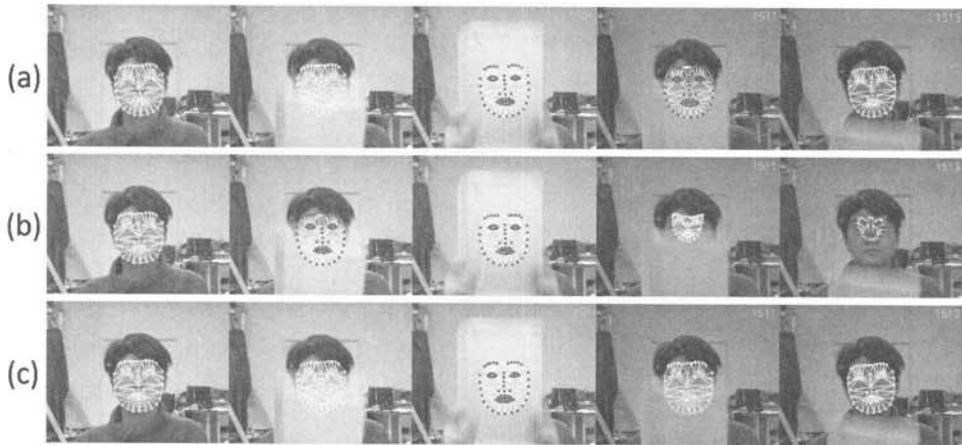
한편 다른 물체로 얼굴을 완전히 가리거나 혹은 영상 밖



(그림 4) 자체중첩에 대한 추적결과, (a) AAM [11], (b) Xiao+AAM [16], (c) 제안한 알고리즘



(그림 5) 부분중첩에 대한 추적결과, (a) AAM [11], (b) Xiao+AAM [16], (c) 제안한 알고리즘



(그림 6) 완전중첩에 대한 추적결과, (a) AAM [11], (b) Xiao+AAM [16], (c) 제안한 알고리즘

으로 잠시 얼굴이 완전히 벗어나는 경우에는, 기대와는 달리 실린더정보를 이용한 두 번째 알고리즘이 오히려 표준적인 AAM 보다 성능이 떨어졌음을 보여주고 있다. 그 이유는 (그림 6)(b)에서 보여주듯이 두 번째 알고리즘은 얼굴이 완전히 사라진 다음 다시 나타나는 과정에 보이는 일부 얼굴영상을 마치 다른 형태의 얼굴로 간주하여 추적에 실패하는 경우가 발생하기 때문이다. 제안한 알고리즘에서는 얼굴이 사라지는 경우 SIFT 검출과 매칭이 실패하여 그 위치에서 추적이 중지된 다음, 다시 나타나면 SIFT의 전역적인 매칭특성으로 돌아오는 시점과 위치에 관계없이 재추적의 가능성이 높아진다. 단 복구되는 시점은 얼굴이 얼굴데이터베이스에 저장되어있는 얼굴특징(SIFT) 중 하나에 근접하는 시점에 해당한다. 본 실험에서는 (그림 6)에서와 같이 완전중첩 상황에서 그 이전과 이후의 얼굴 위치를 동일하게 유지할 하였는데, 그로인해 표준적인 AAM과 두 번째 알고리즘에서 우연히 추적에 성공하는 경우가 종종 발생하였다. 따라서 완전중첩 동안에 얼굴의 움직임이 있을 경우, 기존 두 알고리즘과의 성능차이는 좀 더 커질 것으로 예상된다.

제안한 알고리즘은 이와 같이 일반적으로 자주 발생하는 자체·부분·완전 중첩상황에서 향상된 추적성공률을 보여주었다. 정리하자면 지역적인 정보만을 활용하는 표준적인

AAM 보다 일관되게 안정적인 성능을 보여주었고, 비록 동일한 접근방식이지만 SIFT의 매칭 특성을 바탕으로 기존의 알고리즘 [16] 보다 향상된 결과를 보여주었다. 끝으로, 본 실험에서는 위 세 가지 중첩상황에 대한 비교평가에 중점을 두었기 때문에, 피 실험자 개인의 얼굴 특징은 여기에서 고려하지 않았다; 예를 들어, 안경착용 여부와 머리카락에 의한 얼굴가림 현상. 이러한 개별적인 특징에 대한 평가는 앞으로 연구를 계속 진행할 계획이다.

5. 결 론

본 논문에서는 AAM 모델을 이용한 얼굴추적 알고리즘을 제안하였다. 제안한 알고리즘은 AAM 모델의 중첩에 대한 강인성을 향상시키기 위해서 SIFT 특징의 온라인 등록과 전역적인 매칭을 이용하였다. 이 과정은 얼굴의 3차원 자세를 추정하여 AAM의 국부해 수렴과 발산을 최소화시키고, 또한 추적이 실패하여 재초기화가 필요한 경우에도 별도의 검출과정 없이 자체적으로 연속적인 추적을 가능하게 한다. 이러한 특성으로, 제안한 알고리즘은 자체·부분·완전 중첩상황에서 기존 알고리즘보다 향상된 추적성공률을

보여주었다.

참 고 문 헌

- [1] X. Gao, Y. Su, X. Li, and D. Tao, "A review of active appearance models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.40, No.2, Dec. 2010.
- [2] A. Samal and P. Iyengar, "Automatic Recognition and Analysis of Human Faces and Facial Expression: A Survey," *Pattern Recognition*, vol. 25(1), pp. 65-77, 1992.
- [3] B. Fasel and J. Luetten, "Automatic Facial Expression Analysis: A Survey," *Pattern Recognition*, Vol.36, pp. 259-275, 2003.
- [4] Y. Du, X. Lin, "Mapping emotional status to facial expressions," *Proceedings of 16th International Conference on Pattern Recognition*, pp.524-527, August 2002.
- [5] X. Li, C. Chang, S. Chang, "Face Alive Icons," *Journal of Visual Languages and Computing*, Vol.18, No.4, pp.440-453, 2007.
- [6] C. H. Lee, J. Wetzel, C. Y. Jang, Y. T. Shen, T. H. Chen, T. Selker, "Attention Meter: A Vision-based Input Toolkit for Interaction Designers," *Conference on Human Factors in Computing Systems (CHI)*, pp.1007-1012, Montreal, Quebec, Canada, 2006.
- [7] P. Ekman, "Facial expressions of emotion: an old controversy and new findings," *Philosophical Transactions: Biological Sciences*, Vol.335, No.1273, pp.63-69, 1992.
- [8] P. Ekman, W. Friesen, and J. Hager, "Facial Action Coding System," *Tech. Report, Research Nexus, Network Research Information*, Salt Lake City, UT, 2002.
- [9] J. Cohn, T. Kanade, T. Moriyama, Z. Ambadar, J. Xiao, J. Gao, and H. Imamura, "A Comparative Study of Alternative Faces Coding Algorithms," *Tech. Report CMU-RI-TR-02-06*, Robotics Institute, Carnegie Mellon University, Nov. 2001.
- [10] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.23, No.6, pp.681-685, Jun. 2001.
- [11] I. Matthews and S. Baker, "Active appearance models revisited," *Int. J. Comput. Vis.*, Vol.60, No.2, pp.135-164, 2004.
- [12] J. Xiao, S. Baker, I. Matthews, and T. Kanade, "Real-time combined 2D+3D active appearance models," *CVPR*, 2004.
- [13] R. Gross, I. Matthews, and S. Baker, "Constructing and fitting active appearance models with occlusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. Workshops*, Vol.5, pp.72, 2004.
- [14] B. Theobald, I. Matthews, and S. Baker, "Evaluating error functions for robust active appearance model." *Proc. International Conference on Automatic Face and Gesture Recognition*, pp.149-154, 2006.
- [15] P. Mittraipyanuruk, G. N. EdSouza, A. C. Kak, "Accurate 3D Tracking of Rigid Objects with Occlusion Using Active Appearance Models," *Proc. of the IEEE Workshop on Motion and Video Computing*, pp.90-95, 2005.
- [16] J. Sung, T. Kanade, D. Kim, "Pose robust face tracking by combining active appearance models and cylinder head models," *Int. J. Comput. Vis.*, Vol.80, No.2, pp.260-274, 2008.
- [17] J. Xiao, T. Kanade and J. Cohn, "Robust full-motion recovery of head by dynamic templates and re-registration techniques," *Proc. International Conference on Automatic Face and Gesture Recognition*, pp.156-162, 2002.
- [18] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, Vol.2, No.60, pp.91-110, 2004.
- [19] J. Jang and T. Kanade, "Robust 3D Head Tracking by Online Feature Registration," *Proc. International Conference on Automatic Face and Gesture Recognition*, 2008.
- [20] M. Black and Y. Yacoob, "Recognizing facial expressions in image sequences using local parameterized models of image motion," *IJCV*, Vol.25, No.1, pp.23-48, 1997.
- [21] S. Basu, I. Essa and A. Pentland, "Motion regularization for model-based head tracking," in *ICPR*, pp. 611-616, 1996.
- [22] M. La Cascia, S. Sclaroff and V. Athitsos, "Fast, reliable head tracking under varying illumination: An approach based on robust registration of texture-mapped 3D models," *IEEE Trans. PAMI*, 2000.
- [23] L. Lu, X.-T. Dai, G. Hager, "A particle filter without dynamics for robust 3D face tracking," in *CVPRW*, pp.70, 2004.
- [24] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical introduction to robotic manipulation*, CRC Press, 1994.
- [25] G. Aggarwal, A. Veeraraghavan, and R. Chellappa, "3D facial pose tracking in uncalibrated videos," in *PRMI*, pp. 515-520, 2005.
- [26] CMU Graphics Lab Motion Capture, <http://mocap.cs.cmu.edu>.
- [27] Vicon, <http://www.vicon.com>



엄 성 은

e-mail : seom@cs.cmu.edu

1999년 광운대학교 전기공학과(학사)

2001년 광주과학기술원 기전공학과(석사)

2007년 광주과학기술원 기전공학과(박사)

2007년~2010년 카네기멜론대학 로봇연구소

Post Doc.

2010년~현 재 카네기멜론대학 로봇연구소 Research Programmer
관심분야: 영상처리, 컴퓨터비전, 패턴인식, 의료영상분석



장 준 수

e-mail : junsu.jang@lge.com

1999년 한국과학기술원 전기및전자공학과
(학사)

2001년 한국과학기술원 전자전산학과(석사)

2006년 한국과학기술원 전자전산학과(박사)

2006년~2009년 카네기멜론대학 로봇연구소
Post Doc.

2010년~현 재 LG전자 책임연구원

관심분야: 패턴인식, 물체검출 및 추적, 진화 알고리즘, 로봇 위
치인식