

3D 변형가능 형상 모델 기반 3D 얼굴 모델링

3D Face Modeling based on 3D Morphable Shape Model

장용석*, 정선태*, 김부균*, 조성원**

송실대학교 정보통신전자공학부*, 홍익대학교 지능정보처리 연구실**

Yong-Suk Jang(geniys@ssu.ac.kr)*, Sun-Tae Chung(cst@ssu.ac.kr)*,

Boo-Gyoun Kim(bgkim@e.ssu.ac.kr)*, Seongwon Cho(swcho@wow.hongik.ac.kr)**

요약

3D 얼굴 모델링은 3D 공간에서 얼굴을 자유롭게 회전 시켜 다양한 얼굴 자세를 표현하고 조명 효과도 적절하게 모델링 할 수 있으므로, 얼굴 자세, 조명, 표정 등의 표현에 있어서 2D 얼굴 모델링에 비해 보다 정교하며 사실감이 뛰어나 얼굴 인식, 게임, 아바타 등에서 많은 요구가 존재한다. 본 논문에서는 3D 변형 가능 형상 모델에 기반을 둔 3D 얼굴 모델링 방법을 제안한다. 제안된 3D 얼굴 모델링 방법은 먼저 3D 스캐너를 통하여 획득한 3D 얼굴 스캔 데이터를 이용하여 3D 얼굴 변형 가능 형상 모델을 구축한다. 다음, 3D 얼굴 모델링을 하고자 하는 얼굴의 2D 이미지 시퀀스로부터, 해당 얼굴의 특징점들을 검출하고 이들을 매칭하여, 매칭된 특징점들로부터 인수분해 기반 SfM 기법을 이용하여 해당 특징점의 3D 버텍스 좌표 값을 구한다. 이후, 구한 3D 버텍스들을 3D 변형 가능 형상 모델에 정합하여 해당 얼굴의 3D 형상 모델을 얻는다. 또한, 2D 얼굴 이미지 시퀀스들로부터 뷰 독립적인 2D 원통 좌표 텍스처 맵을 구하고 이를 이용하여 3D 형상 모델을 렌더링 함으로써, 최종적으로 3D 얼굴 모델을 완성한다. 제안된 3D 얼굴 모델링 방법에 의한 3D 얼굴 모델 생성 과정을 통해서, 본 논문에서 제안한 3D 얼굴 모델링 방법이 기존의 얼굴 모델링 방법들에 비해 상대적으로 빠르고 비교적 정교하게 수행됨을 볼 수 있었다.

■ 중심어 : | 3D 얼굴 모델링 | 3D 형상 모델 | 텍스처 합성 | SfM | 3D 복원 |

Abstract

Since 3D face can be rotated freely in 3D space and illumination effects can be modeled properly, 3D face modeling is more precise and realistic in face pose, illumination, and expression than 2D face modeling. Thus, 3D modeling is necessitated much in face recognition, game, avatar, and etc.

In this paper, we propose a 3D face modeling method based on 3D morphable shape modeling. The proposed 3D modeling method first constructs a 3D morphable shape model out of 3D face scan data obtained using a 3D scanner. Next, the proposed method extracts and matches feature points of the face from 2D image sequence containing a face to be modeled, and then estimates 3D vertex coordinates of the feature points using a factorization based SfM technique. Then, the proposed method obtains a 3D shape model of the face to be modeled by fitting the 3D vertices to the constructed 3D morphable shape model. Also, the proposed method makes a cylindrical texture map using 2D face image sequence. Finally, the proposed method builds a 3D face model by rendering the 3D face shape model with the cylindrical texture map. Through building processes of 3D face model by the proposed method, it is shown that the proposed method is relatively easy, fast and precise than the previous 3D face model methods.

■ keyword : | 3D Face Modeling | 3D Shape Modeling | Texture Blending | SfM | 3D Reconstruction |

* 본 연구는 송실대학교 교내연구비 지원으로 수행되었습니다.

접수번호 : #071126-003

접수일자 : 2007년 11월 26일

심사완료일 : 2007년 12월 31일

교신저자 : 정선태, e-mail : cst@ssu.ac.kr

1. 서론

3D 얼굴 모델링은 3D 공간에서 얼굴을 자유롭게 회전시켜 다양한 얼굴 자세를 표현하고 조명 효과도 적절하게 모델링할 수 있으므로, 얼굴 자세, 조명, 표정 등의 표현에 있어서 2D 얼굴 모델링에 비해 보다 정교하며 사실감이 뛰어나 얼굴 인식, 게임, 아바타 등에서 매우 필요로 하는 연구 분야로 컴퓨터 그래픽, 컴퓨터 비전 등의 분야에서 최근 20년 동안 활발히 연구되어 오고 있다[1-13].

얼굴은 볼륨(volume)이나 표면(surface)으로 표현이 가능하나 보통 다각형 표면으로 표현한다. 이는 실시간 처리 때문이다. 다각형 표면은 보통 삼각형 메쉬(triangular mesh)로 구성되며, 메쉬는 버텍스(vertex)로부터 구성된다. 따라서 얼굴 모델링은 얼굴 버텍스의 3D 좌표 값을 결정하는 것이 우선적으로 수행되어야 한다. 얼굴 메쉬가 구축되면, 얼굴 모델의 렌더링은 버텍스의 색상 정보를 이용하거나, 얼굴 모델 텍스처 맵을 생성하여 이를 이용하여 렌더링 한다.

사실적(realistic) 3D 얼굴 모델링 방법에는 3D 스캐너를 통해 스캔된 3D 데이터를 이용하는 방법[14], 3D 일반(generic) 모델을 변형하는 방법[2-4][7-9][11], 3D 변형 가능(morphable) 모델에 기반을 둔 방법[5][6][12], 3D 변형 가능 형상 모델에 기반을 둔 방법[10][13][14] 등이 있다.

3D 스캐너는 얼굴 표면의 각 샘플링 포인트에 대해 3D 좌표 값과 색상 값을 획득하여 준다. 이로부터 해당 인물의 3D 얼굴 모델링이 가능하다. 그러나 이 방법은 값비싼 3D 스캐너 장비를 필요로 하므로 현실적으로는 사용하기에는 비용 부담이 크다.

3D 일반(generic) 모델을 변형하여 3D 얼굴 모델링을 하는 방법은 일반 모델을 얻은 후에, 2D 이미지들에서 얼굴 표면 점들의 3D 좌표 값을 추출하고, 3D 일반 모델을 추출된 3D 좌표 값에 맞도록 변형시켜 해당 인물의 3D 얼굴 형상 모델을 얻고, 2D 이미지들로부터 텍스처를 얻어 이를 얻은 3D 형상 모델에 렌더링 하여 3D 얼굴 모델을 만든다. [7]과 [9]에서는 113개 버텍스와 168개 삼각형으로 구성된 캔디드(Candide) 일반 모델

을 이용하였으며, [8]에서는 얼굴의 정면에 대해 256개 버텍스 및 441개 삼각형을 갖는 워터스의 애니메이션 모델[3]을 이용하였다. 일반 모델을 이용하는 경우, 비교적 간단하고 처리 시간이 적어 빠르게 모델을 구성할 수 있다. 그러나 버텍스와 메쉬가 적어서 보다 정교한 얼굴 모델링이 불가능하다.

3D 변형가능 모델에 기반을 둔 3D 얼굴 모델링 방법에서는 먼저 3D 스캐너를 통해 얻어진 3D 스캔 얼굴 데이터들로 구성된 모델 얼굴 데이터베이스를 통계적 처리하여 3D 얼굴 형상 모델과 얼굴 텍스처 모델로 이루어지는 3D 변형 가능 얼굴 모델을 구축한다. 이후 입력 2D 얼굴 이미지에 정합이 잘 되도록 3D 변형 가능 모델을 조정하여 3D 얼굴 모델을 얻는다. 매우 정교하고 사실적인 3D 얼굴 모델링 방법이나, 시간이 많이 걸리는 단점이 있다.

3D 변형 가능 형상 모델 기반 3D 얼굴 모델링 방법은 먼저 변형 가능 형상 모델을 구축한다. 이후, 입력 2D 이미지 시퀀스로부터 얼굴의 특징점들을 추출하고 매칭하여, 특징점들의 3D 버텍스 좌표 값을 구한다. 구축된 3D 변형 가능 형상 모델이 구한 3D 버텍스들에 잘 정합되도록 3D 변형 가능 형상 모델을 조정하여 3D 형상 얼굴 모델을 얻는다. 또한, 2D 얼굴 이미지 시퀀스들로부터 텍스처를 구하고 이를 이용하여 3D 형상 모델을 렌더링 함으로써, 최종적으로 3D 얼굴 모델을 완성한다. 현재까지 제안된 3D 변형 가능 형상 모델 기반 3D 얼굴 모델링 방법중 가장 나은 [13]에서는 3D 변형 가능 형상 모델을 194 개의 버텍스, 360 개의 삼각형으로 구성된 중립 얼굴(neutral face)과 사람마다의 얼굴 형상 변화를 고려하기 위한 65 개의 메트릭으로 구성된 변형 가능 형상 모델을 제안하였다. 이 형상 모델은 버텍스와 메쉬의 수가 적고, 또한 메트릭 생성 방법이 명확하지 않아 보다 사실적 3D 얼굴 모델링에는 미흡하다. [13]에서는 2D 이미지들에서 프레스시(Plessey) 코너 검출기를 사용하여 코너점을 검출하고 상관관계를 이용하여 코너 점들에 대해 대응점들을 구하였다. 이 코너점들의 3D 좌표 값을 구하는 데 있어, 카메라는 미리 보정(calibration)되어야 하고 2D 이미지 시퀀스의 각 얼굴에 대한 자세를 따로 구한 후에, 이들 정보를 역투

영에 사용하여 2D 코너점에 대한 3D 좌표 값을 구하였다. 그러나 프레스시 코너 검출기는 조명에 취약하며 이 방법에 의한 3D 좌표 값 결정은 카메라 보정을 필요로 한다.

본 논문에서는 3D 변형 가능 형상 모델에 기반한 3D 얼굴 모델링 방법을 제안한다. 본 논문의 방법은 [13]과 접근 방법이 있어 유사하나, 다음의 5가지 점에 있어서, [13]의 방법과는 차이가 있다. 1) 보다 사실적인 3D 얼굴 모델링을 위해 3D형상 모델을 3D 스캔 데이터베이스를 이용하여 구축하였다. 2) 조명에 좀 더 강한 AAM(Active Appearance Model) 을 이용하여 2D 얼굴 이미지들로부터 특징점을 보다 안정적으로 추출하였다. 3) 카메라 보정이 필요 없는 인수분해(factorization) 기반 SfM (Structure from Motion) 방법을 이용하여 AAM 에 의해 획득한 특징점의 3D 좌표 값을 얻었다. 4) 원통좌표계 텍스처 맵 구성에 필요한 각 2D 얼굴 이미지들의 얼굴의 자세를 따로 구하지 않고 인수분해 기반 SfM 으로부터 얻어지는 것을 사용하였다. 5) 3D 변형 가능 형상 모델을 구축하는 데 필요한 3D 조밀 대응점들을 신속하게 구하는 알고리즘을 제안하였다.

본 논문에서 제안한 3D 얼굴 모델링 방법은 3D 변형 가능 얼굴 모델링 방법[5][6][12]에 비해 손쉽고 빠르게 모델링이 가능하며, 3D 일반 모델기반 3D 얼굴 모델링 방법[2-4][7-9][11] 및 기존 3D 형상 모델 기반 3D 얼굴 모델링 방법[10][13][14] 에 비해서는 더 정교함을 볼 수 있다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문을 이해하는 데 필요한 기술적 배경을 기술한다. 3절에서 본 논문에서 제시한 3D 얼굴 모델링 방법의 각 단계와 단계별 구현 내용이 설명된다. 4절에서는 실험 및 분석 내용이 설명되며, 마지막으로 5절에서는 결론이 주어진다.

II. 배경이론

2.1 형상(shape) 및 프로크루스테스 분석(Procrustes analysis)

D.G. Kendal [15]에 의하면, 형상(shape)은 '물체로부터, 이동, 스케일링, 회전 등의 영향이 모두 걸러진 후에 남는 모든 기하학적 정보'를 말한다. 보통 물체의 특징점(landmark)은 동일 형상 클래스에 속하는 물체들의 대응점(동일한 형상 특성을 갖는 점들)을 말한다. 형상을 기술하는 방법의 하나는 바로 형상 위에 있는 유한한 개수의 특징점들로 기술하는 것이다. 따라서 2차원 또는 3차원 형상 S 위에 위치하는 n 개의 특징점들을 V_1, V_2, \dots, V_n 이라 할 때, 형상은 $S = [V_1^T, V_2^T, \dots, V_n^T]^T$ (이때, S 는 $2n \times 1$ 또는 $3n \times 1$ 행렬, 즉 $2n$ 또는 $3n$ 개의 요소를 갖는 열벡터), 또는 $S = [V_1, V_2, \dots, V_n]^T$ (이때, S 는 $n \times 2$ 또는 $n \times 3$ 행렬)로 표현할 수 있다. 한 물체 클래스의 형상들을 정렬(align) 하기 위해서는 이동, 스케일링, 회전 등을 통해 기준 좌표계에 대해 상호간 에러가 최소화하도록 하여야 하며 이렇게 에러가 최소화되도록 하면서 형상을 정렬하는 것을 프로크루스테스 분석을 수행한다고 한다. 프로크루스테스 분석 절차에 대한 상세한 내용과 이를 2D 또는 3D 얼굴 모델링에 이용한 예에 대해서는 [9][16]을 참조하라.

2.2 AAM(Active Appearance Model)

Active Appearance Model (AAM)은 물체의 형상(shape) 과 외양(appearance)의 통계적 모델을 구성하고, 이를 새 이미지에 정합하여, 물체를 찾아내는 알고리즘이다[17][18]. 즉, 어떤 물체 클래스의 통계적 모델을 미리 구성하고, 이 통계적 모델을 조정하여 구성한 해당 클래스 물체의 외양과 새 이미지에서의 물체의 외양과의 차이를 가장 적게 하는 최적화 과정을 통해, 새 이미지에서의 해당 물체를 찾아내는 알고리즘이다. 통계적 모델은 훈련 시에 모델 이미지 데이터베이스를 이용하여 구성된다. 코너 검출기를 이용하여 2D 얼굴 이미지 시퀀스에서 특징점들을 뽑는 방법은 2D 이미지 시퀀스의 이미지들이 조명과 자세가 많이 다른 경우에 안정적으로 동작하지 못한다. 반면, AAM은 조명, 자세에 비교적 강한 것으로 잘 알려져 있다. 따라서 본 논문에서는 2D 얼굴 이미지 시퀀스에서 얼굴 특징점들을 추출하는 데, AAM을 사용하였다. 그런데, 기존 AAM 방법으로는 텍선의 특징점들은 복잡한 얼굴 배경이나

자세에 따라 안정적 추출이 어렵다. 텍선의 특징점들을 보다 안정적으로 추출하기 위해 본 논문에서는 점진적 AAM 방법[18]을 사용하였다.

2.3 ICP(Iterative Closest Point)

Iterative Closest point (ICP) 는 2개의 포인트 클라우드(point cloud)를 매칭 하는데 사용하는 알고리즘이다 [19][20]. 대부분의 ICP 알고리즘들은 메쉬에 대해 적용된다. 따라서 보통 ICP 알고리즘은 먼저 포인트 클라우드로부터 메쉬들을 얻은 후에, 다음과 같이 수행된다. 두 메쉬간의 상대적 변환을 초기 추정하고, 이 변환을 이용하여 두 메쉬를 맞춘 후에 두 메쉬에서 서로 대응하는 버텍스 쌍들을 찾는다. 이후 찾은 버텍스 쌍들의 거리에 대한 에러를 최소화하도록 다시 변환을 교정해 나간다. 어느 정도의 문턱값에 도달하면, 이때의 점 쌍들이 바로 ICP에서 한 쪽 메쉬의 점에 대한 가장 가까운 점에 대한 정보이며, 이러한 가장 가까운 점을 얻게 하는 변환이 바로 이때의 변환이다. ICP 알고리즘이 Besl 과 McKay [19]에 의해 도입된 이후, ICP 알고리즘의 많은 변형[20]과 확장[21]들이 제안되었다. 본 논문에서는 스케일 및 어파인 변환을 지원하는 ICP 알고리즘을 이용하였다.

2.4 TPS(Thin Plate Spline)

TPS(Thin Plate Spline)은 비강체(non-rigid) 스플라인 매핑 함수 클래스로 공간상에 흩어져 있는 점들을 매우 미끈하게(smoothing) 연결하는 표면을 구하는 유용한 도구이다[12][22][21]. 본 논문에서는 샘플 얼굴(sample face)의 국소 형상을 참고 얼굴(reference face)의 버텍스들에 맞추는데(aligned) TPS를 사용하였다.

2.5 3D 변형 가능 얼굴 형상 모델

(3D morphable face shape model)

n 개의 버텍스들로 구성되는 3D 얼굴 형상 데이터는 다음과 같이 형상 벡터 S 로 쓸 수 있다.

$$S = (x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)^T$$

(여기서, x_i, y_i, z_i 는 i 번째 버텍스의 x, y, z 좌표를

나타낸다).

이제, m 개의 3D 얼굴 스캔 데이터에 대하여 PCA (Principal Component Analysis)하여, 평균 형상 벡터 \bar{S} , 형상 주성분 모드 벡터 S_1, \dots, S_{m-1} 을 구하면, 임의의 3D 얼굴 형상 벡터 S 에 대해, 다음과 같이 3D 변형 가능 얼굴 형상 모델로 표현이 가능하다 [5][6][10][12-14].

$$S = \bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i \quad (1)$$

2.6 3D 조밀 대응 (Dense Correspondence)

(1)과 같이 PCA를 이용한 통계적 모델 구성을 할 때 가장 중요한 문제는 각기 다른 형상들 사이에서 포인트 대응점들을 결정해야 한다는 것이다. 왜냐 하면, 형상 벡터들에 대해 동일 인덱스의 점들은 물리적으로 동일한 점들이어야 PCA하는 것이 의미가 있기 때문이다.

얼굴 모델 구축의 경우에서 한 얼굴의 3D 스캔 데이터의 한 버텍스에 대해 다른 얼굴의 3D 스캔 데이터에서 대응되는 점(신체적으로 동일한 점; 예, 모든 3D 스캔 데이터에서 동일한 코끝 지점의 버텍스)을 찾아내는 것을 대응점을 찾는 과정이라 하며, PCA와 같은 데이터들의 통계 처리에 있어서 매우 중요한 작업이다.

매우 많은 3D 스캔 데이터의 버텍스들 사이에 대해 대응점들을 구하는 것을 3D 조밀 대응(3D dense correspondence) 문제를 푸는 것이라 한다.

3D 조밀 대응 문제를 풀기 위해 다음과 같은 방법들이 제안되었다[5][6][12][14][23][24]. 3차원 공간의 표면들 끼리에서의 대응점들을 찾기 위해 수동으로 한 표면에서의 해당 점을 찍고, 다른 표면에서 대응되는 점을 수동으로 찾는 방법이 있다. 그러나 이러한 방법은 3D 스캔 데이터처럼 데이터양이 매우 많은 경우에는 시간이 많이 걸리며, 정확한 작업이 이루어지기 힘들다. 3D 조밀 대응점들을 찾는 방법에는 2D 원주 좌표계 텍스처 맵에서 광류(optical flow)를 이용하여 대응점을 찾는 방법[5][6]이 제안되어 있으나, 이는 각 3D 스캔 데이터간의 조명 변화가 없는 경우에만 잘 적용되는 방법 이어서, 이를 개선한 범선 탐색 방법[14], TPS(Thin Plate Spline) 기반 방법[12]이 제안되었다. 그러나 범선

벡터 탐색 방법은 기존 얼굴 생성과 법선 벡터 탐색이 복잡하다. 또한 [12]에서 제안한 단순 TPS 기반 방법은 기존 3D 얼굴 데이터와 입력 3D 얼굴 데이터들 간에 TPS 할 때마다 대응점을 발견하지 못하는 구멍(hole) 영역이 매번 다르게 생겨, 몇 십 개의 3D 얼굴 데이터 간의 공통 대응점의 수는 급속히 줄어들어 PCA를 수행하기 위한 3D 얼굴 모델링 데이터로 사용하기에는 대응점의 수가 적게 되는 단점이 발생한다. 또한, 1차 오더 타원체로 맞추어진 균일 영역 파라미터화를 통해 대응점 문제를 해결한 SPHARM(Spherical Harmonics) 방법[23]과 최소 기술 길이(MDL; minimum description length)에 기반을 둔 비용 함수를 도입하여 대응점 문제를 최적화 문제로 푸는 방법[24]이 제안되어 있다. 그러나 MDL과 SPHARM은 genus 0의 닫힌 표면에 대해서만 적용될 수 있으며, 계산량이 매우 크다. 3D 스캔 데이터로 얻어진 얼굴 형상 데이터는 머리카락 부분에 대한 데이터가 없기 때문에 닫힌 표면이 아니어서, MDL과 SPHARM을 적용하기 어렵다.

본 논문에서는 3D TPS 및 최단거리 주변 탐색(Nearest Neighbor Search) 방법을 이용한 개선된 빠른 3D 얼굴 조밀 대응점 탐색 알고리즘을 제안하였다.

2.7 Structure from Motion

2D 이미지 시퀀스로부터, 장면의 구조(structure)와 카메라 움직임(motion), 자세(pose)를 구하는 과정을 보통 SfM(Structure from Motion)이라 하며, 컴퓨터 비전 분야의 주요 문제로서 오래전부터 연구되어 오고 있다[25-30]. SfM에서는 2D 이미지들의 특징으로부터 3D 물체의 모델을 구하고, 카메라 자세(카메라 움직임)를 구한다. 2D 특징(특징점, 코너, 에지, 윤곽 커브)의 위치는 ① 3D 공간에서의 좌표 ② 카메라와 장면의 상대적 3D 움직임(motion) ③ 카메라 내부 파라미터 등에 의존한다. SfM에서는 이들 3가지에 대해 아무런 정보를 갖지 않는다고 가정하며, 여러 프레임의 2D 이미지들에 걸친 얻어진 2D 특징 값으로부터, 이들 값들을 구하고자 한다.

SfM 방법에는 2장의 2D 이미지 프레임으로부터 3D 정보(장면의 구조 및 카메라 움직임)를 추출하는 에피

플라(epipolar) 기하학 기반 SfM[25], 3장의 2D 이미지 프레임으로부터 3D 정보를 추출하는 트리포칼(trifocal) 텐서 기반 SfM[25], 다수의 2D 이미지 프레임으로부터 인수분해(factorization)를 이용하여 3D 정보를 추출하는 SfM 방법[26-28], 비선형 동적 시스템 기반의 비선형 칼만 필터를 이용한 SfM 방법[29], 얼굴 등과 같은 비강체 물체(non-rigid body)에 대한 인수분해 기반 SfM 방법[30] 등이 제안되어 있다. 에피플라 기하학에 의한 SfM은 완전한 특징과 큰 베이스라인을 갖는 이미지들에 대해서는 잘 동작하지만, 잡음에 민감하다. 본 논문에서는 카메라로부터 얼굴 표면들의 깊이 값의 차이가 크지 않으므로 근사적으로 약원근투영(weak perspective)[25]을 적용할 수 있다는 가정 하에 얼굴과 같은 비강체 물체에 대해 개발된 인수분해 기반 SfM 방법[30]을 이용하여, 2D 이미지들의 특징점들에 대한 3D 좌표 값과 카메라 자세들을 구하였다.

III. 3D 변형가능 형상모델 기반 3D 얼굴모델링

3.1 개요

본 논문에서 제안한 3D 변형 가능 형상 모델 기반 3D 얼굴 모델링은 크게 1) 변형 가능 형상 모델 구성 2) 3D 얼굴 모델링 등의 2개의 과정으로 이루어진다.

1) 변형가능 형상 모델 (Morphable Shape Model) 구성
여러 명의 3D 얼굴 스캔 데이터들로부터, PCA를 통해 식 (1)로 표현되는 3D 변형 가능 모델을 구성한다. 이 과정에서 가장 중요한 작업은 3D 얼굴 스캔 데이터들 간의 조밀 대응점을 구하는 것이다.

2) 3D 얼굴 모델링

2D 얼굴 이미지 시퀀스들로부터 해당 얼굴의 3D 얼굴 모델을 구하는 과정으로 다음의 절차로 이루어진다.

- ① 2D 얼굴 이미지 시퀀스에서 특징점 추출, 매칭 및 매칭된 2D 특징점들로부터 해당 특징점의 3D 벡터 좌표 값 복원
- ② 3D 특징점 벡터스들로 구성되는 포인트 세트를

- 변형 가능 형상 모델에 정합
- ③ 2D 얼굴 이미지들로부터, 3D 얼굴 모델에 입힐 텍스처 맵 생성
- ④ 정합된 3D 형상 모델에 텍스처를 입혀 3D 얼굴 모델 완성

3.2 3D 얼굴 스캔 데이터베이스

본 논문에서는 Cyberware[14]사의 모델 3030 컬러 3D 스캔헤드를 이용하여 20대에서 50대까지 남녀 40명의 인원의 머리를 3D 스캔하였다. 스캔헤드의 스캔 데이터는 형상에 대해서는 512개의 균등하게 분할된 각도 ϕ 와 균등하게 분할된 수직 길이 스텝 h 에서 샘플링된 점의 x, y, z 좌표 값과 각 점의 R, G, B 값을 채널당 8비트로 제공한다. 모든 얼굴은 화장 없이, 액세서리 없이, 그리고 얼굴의 털이 없이 스캔되었다. 스캔된 얼굴은 대략 70,000개 이상의 버텍스와 같은 수의 컬러 값으로 표현된다.

다음 [그림 1]은 3D 스캔된 얼굴 데이터베이스의 일부이다. [그림 1]의 3D 얼굴은 스캔한 후에, 본 논문의 얼굴 모델링을 위해 전처리(귀 및 머리카락 부분 제거)된 영상이다.



그림 1. 3D 얼굴 스캔 데이터 예

3.3 3D 얼굴 조밀 대응점 구하기 (3D face dense correspondence)

3D 스캔으로 얻어진 40명의 얼굴 스캔 데이터들로부터 변형 가능 형상 모델을 구성하기 위해서는 3D 얼굴 조밀 대응점 문제를 해결하여야 한다.

본 논문에서 제안한 조밀 대응점 구하기 알고리즘은 대략적으로 다음과 같이 이루어진다. 먼저, 3D 스캔 데이터 얼굴 중에서 기준 얼굴을 선택하고, 샘플 얼굴(나머지 얼굴중 하나)을 프로크루스테스 분석과 3D TPS 알고리즘을 이용하여 전체 형상 (자세, 크기, 중심점 위

치) 과 국소 형상(국소적 부분의 형상) 이 기준 얼굴과 가급적으로 유사하도록 정렬 한 후에, 기준 얼굴의 버텍스들과 가장 가까운 해당 버텍스를 샘플 얼굴에서 찾는다. 만일 기준 얼굴 버텍스와 대응하는 샘플 얼굴 버텍스의 거리가 기준 값보다 크지 않으면, 샘플 얼굴의 버텍스를 기준 얼굴의 해당 버텍스에 대응하는 대응 버텍스로 결정한다.

초기부터 ICP 를 이용하지 않은 이유는 ICP 알고리즘은 계산량이 많아 대상 포인트 세트가 커지면 시간이 많이 소요되며, 또한 처음에 2개의 포인트 세트(기준 얼굴 데이터와 샘플 얼굴 데이터)가 잘 정렬되어 있지 않으면 ICP 알고리즘 동작의 안전성이 보장되지 않기 때문이다. 따라서, 먼저 기준 얼굴 데이터와 샘플 얼굴 데이터들을 정렬할 필요가 있으며, 본 논문에서는 전체 형상 정렬에 프로크루스테스 분석을, 국소 형상 정렬에 3D TPS를 이용하였다. 제안 3D 대응점 구하기 알고리즘을 좀 더 자세히 설명하면, 다음과 같다.

1) 제어점(control point) 및 기준 얼굴 선택

프로크루스테스 분석 [9]과 TPS 알고리즘 [12][21][22]은 제어점을 이용하여 수행된다. 본 논문에서는 눈, 코, 입, 귀 등에서 23개의 점을 제어점으로 선택하였다. 또한 버텍스의 개수가 가장 적은 3D 얼굴 데이터를 갖는 얼굴을 기준 얼굴로 선택 했다. 아래 그림 2에서 점은 프로크루스테스와 TPS 알고리즘에서 사용된 제어점들의 위치를 나타낸다.



그림 2. 사용된 제어점 위치

2) 프로크루스테스 분석을 이용한 얼굴 전체 형상 정규화

프로크루스테스 분석[9]을 이용하여 기준 얼굴 이의 나머지 얼굴들을 스케일링하고 이동하고, 회전하여 기준 얼굴과 같은 크기를 가지고 같은 방향을 바라보도록

기준 얼굴 외 나머지 39명의 얼굴들의 전체 형상이 비슷하도록 정규화 한다. 3D 스캐너를 통해 스캔한 데이터는 각기 다른 방향과 다른 크기를 가지고 있을 수 있다. [그림 3]은 이러한 경우의 예를 보여 준다. [그림 3]의 왼쪽 그림에서 보면 두 얼굴(색깔이 다르게 표시)의 크가 각기 다른 방향을 가리키고 있는 것을 알 수 있다. 또한, 위쪽에 놓인 얼굴 부분이 아래에 놓인 다른 얼굴을 가리는 것을 명확히 보여 주기 위해, [그림 3]의 오른쪽 그림은 왼쪽의 그림에서 녹색 색깔 얼굴(후백 출력물에서는 진한 농도의 얼굴)에 투명도를 주어 아래에 놓인 얼굴 부분이 비쳐보이게 한 그림이다.

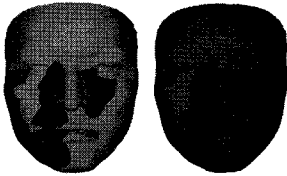


그림 3. 크기와 방향이 다른 2개의 3D 얼굴 스캔 데이터

[그림 3]의 경우처럼, 기준 얼굴과 다른 크기와 방향을 갖는 3D 얼굴 스캔 데이터가 같은 크기를 가지고 같은 방향을 바라보도록 프로크루스테스 분석을 수행한다. 이 수행 결과를 보여 주는 경우가 [그림 4]이다.

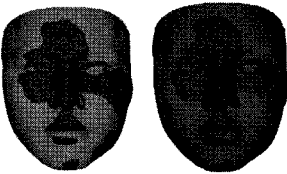


그림 4. 프로크루스테스 분석후 3D 스캔 데이터 비교

3) 3D TPS 를 이용한 얼굴 형상 국소 형상 정렬 프로크루스테스 알고리즘 수행 후 샘플 얼굴은 기준 얼굴과 형상의 전체 형태는 비슷하지만, 국소적 형태(예, 눈 부분, 입 부분, 코 부분 등)에서는 여전히 많은 차이가 있다([그림 4]참조). 물론 이렇게 전체적 형태를 맞추는 정렬 후에, 샘플 얼굴의 버텍스 중 기준 얼굴의 각 버텍스에서 가장 가까운 점이나, 각 버텍스의 법선 벡터와 교차하는 점들을 대응점으로 처리할 수 있다.

이렇게 하는 경우에, 국소적 형태가 맞지 않았기 때문에 대응점으로 판명한 점들은 에러가 크게 된다. 따라서 가급적 샘플 얼굴을 기준 얼굴과 맞추어야 한다.

본 논문에서는 제어점을 이용한 3D TPS 을 통해, 샘플 얼굴을 기준 얼굴과 가급적이면 국소적으로도 가깝게 되도록 맞추었다.

[그림 5]는 3D TPS를 이용하여 샘플 얼굴이 기준 얼굴과 국소적으로도 정렬되도록 한 결과의 두 얼굴(기준 얼굴과 샘플 얼굴) 형상의 비교를 보여 준다.

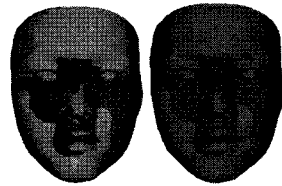


그림 5. TPS가 적용된 후 3D 얼굴 스캔 데이터 비교

[그림 4]와 비교해보면, 두 얼굴이 국소적으로도 비교적 잘 맞추어져 있음을 알 수 있다.

4) 가장 가까운 거리 기준 대응점 탐색

기준 얼굴의 각 버텍스에 대해, 기준 얼굴에 맞추어진 샘플 얼굴에서 가장 가까운 점을 찾는다. 만일 둘 사이의 거리가 기준 값보다 크면 해당 버텍스는 대응점이 아니라고 판단하고 버린다.

5) 원래 얼굴 데이터에서 대응점 확정

4)에서 얻은 가장 가까운 버텍스는 프로크루스테스 과정과 TPS를 이용하여 변형되어 있기 때문에 원래 샘플 데이터의 점은 아니다. 따라서 이들 버텍스와 대응되는 스캔 데이터에서의 버텍스를 최종 대응점의 버텍스로 확정한다.

실험을 통해 81,134개부터 598,121개의 버텍스들을 가진 40명의 3D 스캔 데이터들로부터 대략 4만개의 공통 대응점(대응 버텍스)으로 구성된 40명의 3D 스캔 데이터를 구축할 수 있었다. [그림 6]은 원래의 3D 스캔 데이터를 세이딩 한 것과 대략 4만개의 대응점만을 가

지고 세이딩 한 2개의 3D 얼굴들을 비교하여 보여 준 것이다. 전체적인 얼굴 형상과 대략적인 곡소 형상에 있어서, 원래의 3D 스캔 데이터의 얼굴과 공통 대응점만의 데이터의 얼굴이 비슷함을 알 수 있다.

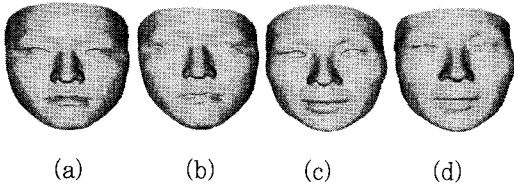


그림 6. (a,c) : 원래 3D 스캔 데이터
(b,d) : 3D 공통 대응점만의 3D 데이터
(a는 기준얼굴, b는 샘플얼굴)

3.4 PCA를 통한 3D 변형 가능 형상 모델 구성

40명의 3D 스캔 데이터에 대해, 대응점들을 구해서 정렬시킨 대략 4만개의 버텍스를 갖는 얼굴 형상 벡터 F_1, F_2, \dots, F_m ($m = 40$) 에 대해 먼저, 평균 형상 $\bar{S} = (F_1 + F_2 + \dots + F_m) / m$ 을 구한다.

이제, $A_1 = F_1 - \bar{S}$, $A_2 = F_2 - \bar{S}, \dots, A_m = F_m - \bar{S}$ 으로 구성된 데이터 행렬 $A = (A_1, A_2, \dots, A_m)$ 에 대해 공분산 $C = \frac{1}{m} AA^T$ 의 주 성분 모드 S_1, S_2, \dots, S_{m-1} 를 구한다.

이제 평균 형상과 주 성분 모드를 이용하여 3D 샘플 얼굴 S 는 (1) 같이 표현될 수 있다.

3.5 2D 얼굴 이미지 시퀀스에서 특징점 추출,

매칭 및 3D 좌표값 복원

이미지 특징점 매칭은 여러 장의 이미지 시퀀스에서 지정 특징점들에 대해 여러 장에 걸쳐 대응점들을 찾아 내는 것을 말한다. 2D 이미지들에서 대응점들을 구하는 방법에는 광류 또는 특징 벡터 기반 코너 매칭 등이 있다. 광류 방법이나 특징 벡터 기반 코너 매칭 방법은 기본적으로 이미지 시퀀스들이 조명의 변화가 별로 없는 경우에는 비교적 잘 동작이 되나, 보통의 이미지 시퀀스들이 조명 변화가 있으므로, 잘 동작하지 않는다.

본 논문에서는 점진적 AAM[18]을 이용하여 이미지 시퀀스에서 특징점을 구하고 이들의 대응점들을 구하였다. 이후, 대응 특징점 위치 정보로부터 얼굴과 같은 비강체 물체에 적용되는 인수분해 기반 SfM [30]을 이용하여 특징점들의 3D 버텍스의 좌표값들을 복원하였다.

2D 얼굴의 AAM 에 사용한 특징점들은 63개이고, 위치는 다음 [그림 7]과 같다.

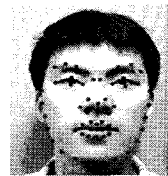


그림 7. AAM에서 사용한 얼굴의 특징점들의 위치

[그림 7]의 특징점들 가운데 눈, 코, 입 등 부분의 특징점은 얼굴 자세와 관계없이 2D 이미지 시퀀스에서 항상 같은 해부학적 점을 표시한다. 그러나 턱 부분의 일부 윤곽점들은 자세에 따라 2D 이미지 시퀀스의 각 이미지에서 같은 해부학적 점이 아닐 가능성이 있다. 따라서 본 논문은 2D 얼굴 이미지 시퀀스들에 대해 [그림 7]에서 나타난 특징점들에 대해 2D AAM 모델을 구축하고, 사용되는 2D 이미지 시퀀스 각각에서 [그림 7]에 나타난 특징점들의 위치를 추출하고 SfM 에서도 이들 특징점들에 대한 3D 버텍스의 좌표값을 구하였으나, 3D 변형 가능 형상 모델과의 정합시에 턱의 15개의 버텍스는 식(2)와 같이 거리에 따라 가중치를 매우 낮추거나 가중치를 0 으로 하여 사용하지 않았다.

[그림 8]의 그림은 2D AAM에서 구한 2D 특징점들로부터 구한 특징점의 3D 좌표 값을 가지고 메쉬를 만들고, 이 메쉬에 2D 얼굴 이미지를 렌더링 하여 보여 준다. [그림 8]과 [그림 9]를 살펴보면, 본 논문에서 2D AAM 에 의해 특징점들을 추출하고, 직교 투영에 대한 인수분해 기반 SfM 을 적용하여 얻은 3D 데이터가 비교적 잘 복원되었음을 보여 준다(4.3절 참조).

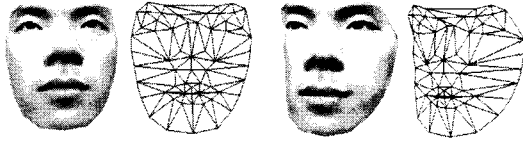


그림 8. 복원된 3D 좌표의 메쉬

아래 [그림 9]는 복원된 3D 벡터로 구성된 메쉬를 해당 자세의 2D 이미지에 투영한 사진을 보여 준다.



그림 9. 복원된 3D 메쉬를 2D 이미지에 투영한 사진

3.6 3D 얼굴 형상 모델 정합

2D AAM과 인수분해 기반 SfM을 이용하여 복원한 얼굴 특징점들의 3D 벡터들이 주어 졌을 때, 이를 3D 변형 가능 형상 모델에 정합하는 과정은 다음과 같다.

먼저, 구한 3D 특징점 벡터들 가운데, 눈, 코, 입 등을 표현하는 특징점 벡터 세트를 $\{Q_1, \dots, Q_l\}$ 라 하고 [그림 10], 그 밖의 다른 3D 특징점 벡터 세트를 $\{P_1, \dots, P_n\}$ 으로 표기하기로 하자.



그림 10. 특징점 벡터 세트 $\{Q_1, \dots, Q_l\}$ 의 2D 이미지에서의 위치

본 논문에서는 3D 얼굴 모델 정합을 [13]에서와 같이 다음 (2)으로 주어지는 비용을 최소화하는 얼굴 자세 T 와 변형 가능 형상 모델의 계수 $\alpha_i (i=1, \dots, m-1)$ 를 구하는 최적화 문제로 해결하였다.

$$\sum_{i=1}^n w_i^2 d^2(P_i, S) + \sum_{j=1}^l d^2(Q_j, V_j) \quad (2)$$

$$w_i^2 = \frac{1}{1 + d^2(P_i, G_i)}$$

여기서, $d(P_i, S)$ 는 특징점 벡터 P_i 와 얼굴 모델 S 와의 거리를 나타내는 것으로, 이는 얼굴 모델 S 에서 벡터 P_i 와 가장 가까운 점을 찾고, 이 가까운 점과 P_i 사이의 거리로 정의된다.

V_i 는 얼굴 모델 S 에서, 특징점 벡터 세트 $\{Q_1, \dots, Q_l\}$ 에 대응하는 벡터 $\{V_1, \dots, V_l\}$ 를 말하며, $d(Q_i, V_i)$ 는 3D 복원된 특징점 벡터 Q_i 와 얼굴 모델의 V_i 와의 유클리드 거리로 정의된다.

또한, G_i 는 얼굴 모델 S 에서 벡터 P_i 에 가장 가까운 벡터이다. 수식 (2)의 기본 아이디어는 눈, 코, 입 등을 표현하는 벡터 [그림 10]는 보다 정확히 정합되어야 하므로, 가중치를 다른 벡터들의 정합 가중치보다 높게 한 것이다.

이제, 구해진 3D 특징점 벡터가 정합할 얼굴 3D 모델은 변형 가능 형상 모델에 대해 얼굴 자세(회전(R), 이동(t), 스케일링(s))가 적용된 얼굴 모델 S 이다. 즉,

$$S = T(\bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i). \quad (\text{여기서, } T = \begin{pmatrix} sR & t \\ 0^T & 1 \end{pmatrix})$$

3D vector p 에 대해, $T(p)$ 는 $T(p) = sRp + t$ 를 의미하며, n 개의 3D vector로 구성된, 3×1 vector S 에 대해, $T(S)$ 는 각 3D 벡터 p 에 대해, $T(p)$ 를 적용한 것 같은 의미이다.)

수식 (2)를 최소화 하는, 얼굴 자세 T 와 변형 가능 형상 모델의 계수 $\alpha_i (i=1, \dots, m-1)$ 가 해결되면, 정

합된 최종 얼굴 형상 모델은 $S = T(\bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i)$ 로 표현된다.

$$\text{이제, } P_Q = [P_1^T, P_2^T, \dots, P_n^T, Q_1^T, \dots, Q_l^T]^T,$$

$$G_V = [G_1^T, G_2^T, \dots, G_n^T, V_1^T, \dots, V_l^T]^T,$$

$W = \text{diag}(w_1, \dots, w_n, 1, \dots, 1)$ 라 할 때, 수식 (2)는 다음과 같이 수식 (3)으로 표현되며,

$$\|W(P_Q - G_V)\|^2 \quad (3)$$

3D 얼굴 모델 정합은 (3)를 최소화 얼굴 자세 T 와 변형 가능 형상 모델의 계수 $\alpha_i (i=1, \dots, m-1)$ 를 구하

는 것과 같다.

본 논문에서는 수식 (2)을 최소화하는 얼굴 자세 T 와 변형 가능 형상 모델의 계수 $\alpha_i (i=1, \dots, m-1)$ 는 다음의 단계로 구한다.

단계 0)

처음, $\alpha_1 = \alpha_2 = \dots = \alpha_{m-1} = 0$ 로 취한다. 또한,

또한, $\alpha_i = 0 (i=1, \dots, m-1)$ 를 $\alpha_i^0 (i=1, \dots, m-1)$

라 하고, $k=0$ 으로 취한다.

단계 1)

$\alpha_i^k (i=1, \dots, m-1)$ 가 고정된 상태에서 가장 가까운 점세트를 구하고, $\|P_Q - T(\hat{S} + \sum_{i=1}^{m-1} \alpha_i^k \hat{S}_i)\|^2$

를 최소화 하는 T 를 구한다. (여기서, \hat{S} 와 $\hat{S}_i (i=1, \dots, n)$ 는 \bar{S} 와 $S_i (i=1, \dots, m-1)$ 의 성분 가운데, 벡터 세트 $\{P_1, \dots, P_n, Q_1, \dots, Q_l\}$ 의 성분들과 같은 위치에 있는 점들만을 취하여 만든 벡터이다.

이때의 T 를 T^k 로 표기하기로 한다.

단계 2)

단계 1)에서 구한 T^k 를 사용한 $S (S = T^k \bar{S} + \sum_{i=1}^{m-1} \alpha_i^k T^k S_i)$ 상에서, 특징점 벡터 세트 $\{P_1, \dots, P_n, Q_1, \dots, Q_l\}$ 에 가장 가까운 점 세트 $\{G_1^k, \dots, G_n^k, V_1^k, \dots, V_l^k\}$ 을 구하고, 이를 $T^k \hat{S}$ 라 한다. 즉, $T^k \hat{S} = [(G_1^k)^T, \dots, (G_n^k)^T, (V_1^k)^T, \dots, (V_l^k)^T]^T$

단계 3)

$\|P_Q - (T^k \hat{S} + \sum_{i=1}^{m-1} \alpha_i T^k \hat{S}_i)\|^2$ 를 최소화하는 $\alpha_i (i=1, \dots, m-1)$ 를 구한다. 여기서 구한 $\alpha_i (i=1, \dots, m-1)$ 를 $\alpha_i^{k+1} (i=1, \dots, m-1)$ 라 하자.

단계 4)

$\|P_Q - (T^k \hat{S} + \sum_{i=1}^{m-1} \alpha_i^{k+1} T^k \hat{S}_i)\|^2$ 가 충분히 작거나(정해진 문턱값보다 작거나) 각 반복에서 거리의 값이 더 이상 줄지 않는 경우에는 종료한다.

만일 그렇지 않으면, k 를 하나 증가시키고 단계 1)로 가서 반복한다.

참고로, $S = T(\bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i) = T\bar{S} + \sum_{i=1}^{m-1} \alpha_i T S_i$ 이며, 결국 $G_i (i=1, \dots, n)$ 및 $V_j (j=1, \dots, l)$ 는 고정된 T 에 대해 $\alpha_i (i=1, \dots, m-1)$ 의 선형 함수 이고, 비용 함수 $\|P_Q - G_i\|^2$ 는 볼록함수(convex function)이므로, 위 반복 알고리즘은 수렴한다. 즉, 해가 존재한다.

상기 정합 알고리즘에 의해, [그림 7]의 2D 얼굴 이미지들에 대해 SfM으로 구한 3D 특징 벡터들을 3D 변형 가능 형상 모델에 정합하여 얻어진 3D 얼굴 형상과 스캔 데이터 형상 모델과의 비교는 [그림 15]에서 볼 수 있다.

3.7 텍스처 맵(texture map) 생성

본 논문에서는 [13]에서의 접근 방법과 유사하게 다음과 같이 2D 이미지 프레임들을 합성(blending)하여 뷰 독립적 텍스처 맵을 생성 하였다.

뷰 독립적 텍스처 맵을 구한 이유는 획득한 2D 얼굴 이미지의 조명 환경에 의존적이지 않고 정확한 색상의 3D 얼굴 모델을 얻기 위해서 이다.

k 개의 2D 얼굴 이미지 프레임에 대해 각 얼굴 이미지의 머리 자세, 카메라 각도를 구한 후에 각 이미지에 대해서 카메라 각도와 각 얼굴의 각 메쉬와의 각도를 구한다[그림 11]. [그림 11]는 시각 효과를 위해, 원래 복원된 메쉬를 800 개 정도로 줄여 투영한 것이다.



그림 11. 복원된 메쉬를 2D 이미지에 투영한 이미지

각 메쉬의 가중치는 카메라 시선 단위 벡터와 3D 얼굴 매쉬의 단위 법선 벡터의 내적에 기반을 두어 각 이미지 프레임 i ($i=1, \dots, k$)에 대한 각 메쉬의 합성 가중치를 정한다. 메쉬가 보이지 않으면 가중치는 0으로 하고 전체 가중치의 합은 1로 정규화 한다. 각 이미지 프레임 i ($i=1, \dots, k$)의 각 메쉬의 가중치를 원통 좌표 맵의 메쉬(원통 좌표로 펼친 메쉬)에 대한 가중치 맵 W_i ($i=1, \dots, k$)로 사용한다. 또한, 각 2D 이미지 프레임 i ($i=1, \dots, k$)을 텍스처 맵으로 하여 원통 좌표로 매핑된 메쉬를 렌더링하여 원통 좌표계 텍스처 맵 C_i ($i=1, \dots, k$)를 생성한다. C_i 와 W_i ($i=1, \dots, k$)를 원통좌표 텍스처 맵과 가중치 맵이라 하자. 또한 C 를 최종 합성된 원통 좌표계 텍스처 맵이라 하자. 각 픽셀 (u, v) 에 대해, 최종 합성되는 텍스처 맵의 색상은 다음과 같이 결정된다.

$$C(u, v) = \sum_{i=1}^k W_i(u, v) C_i(u, v) \quad (4)$$

다음 [그림 12]는 최종적으로 구해진 합성된 원통 좌표계의 텍스처 맵 $C(u, v)$ 의 예를 보여준다.

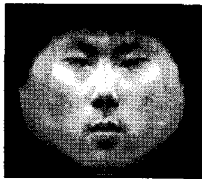


그림 12. 원통좌표계의 텍스처 맵

3.8 3D 얼굴 모델 완성

3.6절의 절차에 의해 구해진 3D 얼굴 형상 모델에 3.7절에서 구한 합성된 뷰 독립 원통좌표계의 텍스처 맵을 렌더링 하여 최종적으로 3D 얼굴 모델을 구한다.

다음 [그림 14]는 [그림 11]의 복원된 3D 얼굴 형상 모델에, [그림 13]의 합성된 원통좌표계의 텍스처 맵을 렌더링 하여 완성된 3D 얼굴 모델과 해당하는 2D 얼굴 이미지들을 비교하여 보여준다.

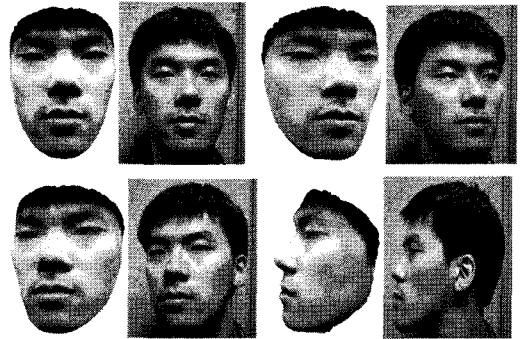


그림 13. 3D 얼굴 모델 과 2D 얼굴 이미지 비교

[그림 13]에서 보면 본 논문에서 제안한 3D 얼굴 모델링 방법에 의해 제작된 3D 얼굴 모델이 비교적 2D 얼굴 이미지와 유사하게 만들어졌음을 볼 수 있다. 복원된 3D 얼굴에서 코 부분의 윤곽이 2D 얼굴 이미지에 비해 몽실하게 보이는 것은 복원된 3D 얼굴은 뷰 독립적인 텍스처링을 하였기 때문에 2D 이미지들의 텍스처들이 합성되면서 조명에 영향을 덜 받기 때문에 뷰에 따라 조명의 영향으로 또렷이 드러나는 코의 윤곽 등이 희미하게 되어 버렸기 때문이다. 마찬가지로, 복원된 3D 얼굴 모델은 뷰 독립적 텍스처를 사용했기 때문에 해당 2D 얼굴 이미지와는 조명 효과에 따른 색상의 차이가 존재한다.

IV. 실험 및 결과 분석

4.1 실험 환경

본 논문에서는 고정된 1개의 웹캠을 바라보고 사람이 천천히 얼굴을 좌우상하로 움직이며 촬영한 연속된 2D 얼굴 이미지 시퀀스 366장을 사용하여 3D 얼굴 모델링을 수행하였다. 3.2절의 [그림 7][그림 9] 등은 획득한 2D 얼굴 이미지 시퀀스의 일부 프레임들을 보여 준다. 웹캠의 촬영 속도는 초당 10~12 장 이었다. 획득한 얼굴 이미지 시퀀스 중 80장 프레임에 대해 AAM을 이용하여 얼굴의 63개의 특징점들을 자동으로 추적하였다. 자동으로 추적된 특징점들로부터 비강체 물체에 적용되는 인수분해 기반 SfM [30]을 이용하여 3D 버텍스 좌

표값을 복원하였다. 특징점들의 버텍스를 미리 구축한 3D 변형 가능 형상 모델에 정합하여 3D 형상 모델을 얻은 후에, 이 형상 모델을 촬영된 2D 얼굴 이미지 시퀀스 중 5장을 선택하여 투영하고 투영된 5개의 2D 메쉬 정보로부터 뷰 독립 텍스처 맵을 구한 후, 이를 이용하여 형상 모델에 렌더링하여 3D 얼굴 모델을 구하였다.

변형 가능 형상 모델은 남녀 40명을 이용하여 모델을 구성하였다. 구성한 3D 변형 가능 모델은 본 논문이 제시한 3D 조밀 대응으로 구한 공통 대응 버텍스 38,257개의 버텍스로 구성된다.

본 논문의 실험에서는 상기 실험 환경에서 제3절에 설명한 3D 모델링 절차들을 구현하고 구현의 각 단계에 대해 구현의 정확성 및 수행 속도 등을 분석하였다.

4.2 2D AAM에 의한 특징점 추출 분석

2D AAM은 먼저 수동으로 얼굴 특징 벡터 위치들을 추출하고 이를 이용하여 AAM 모델 (형상 모델 및 텍스처 모델)을 구성한다. 이후, 입력 2D 이미지를 AAM 모델에 정합하여 얼굴 특징 벡터 위치를 자동으로 추출한다. 추출된 위치의 정확도는 AAM 모델을 얼마나 잘 구축하였는가에 의존한다. 또한, 턱부분 등의 외곽 부분은 특히 배경이 복잡한 경우에는 정확히 추출이 어려운 경우가 적지 않다. 따라서 본 논문에서는 [18]에서 제안한 점진적 AAM 기법을 사용하여 턱 윤곽 부분의 특징점 추출도 안정적으로 이루어지도록 하였다. 다음 [그림 14]는 본 논문에 이용한 점진적 AAM 기법에 의한 특징점 추출 결과들을 보여 주며, 이는 본 논문에서 제안한 방법에 의한 점진적 AAM에 의해 특징점이 안정적으로 추출된다는 것을 확인하여 준다.

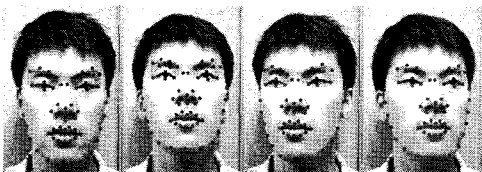


그림 14. AAM에 의한 특징점 추출 결과

4.3 SfM에 의한 3D 버텍스 좌표값 복원 분석

2D AAM에 의해 추출된 얼굴 특징점 위치 정보를 가지고 비강체 물체에 적용되는 인수분해 기반 SfM [30]을 통하여 해당 특징점 3D 버텍스의 좌표를 구하였다. SfM 결과로 얻어진 3D 버텍스 좌표값의 정확성 분석을 위해, 3D 스캔 얼굴 데이터 및 SfM 버텍스 좌표값들이 한 번의 길이가 2 (-1 ~ +1) 이 되는 정육면체 내에 중앙을 중심으로 위치하도록 정렬한 후 동일 특징점 사이의 차이(에러)의 평균을 구하였다. [표 1]의 결과는 [그림 7]의 얼굴에 대해 63개 특징점 [그림 7]과 17개 특징점 [그림 10]들 각각에 스캔 데이터 3D 좌표와 SfM으로 복원된 3D 좌표값의 차이를 구하고 이들을 평균한 값이다. [표 1]에서 프레임 수는 SfM에서 복원할 때 사용한 프레임수를 말한다. 프레임 수가 증가될수록 잘 복원이 되며, 얼굴 텍선의 특징점들보다는 얼굴 내부 특징점들이 잘 복원되어 얼굴 내부의 특징점들만의 에러가 텍선의 특징점들을 포함한 전체 얼굴 특징점들의 평균 에러가 더 작음을 알 수 있으며, 어느 정도 잘 복원되었음을 보여 준다.

표 1. SfM 결과 비교

프레임 수	63개 점	17개 점
8	0.2888	0.0642
35	0.0759	0.0164
80	0.0073	0.0021

본 논문의 실험에서는 80장의 이미지 프레임과 63개의 특징점을 사용하였다.

4.4 변형 가능 형상 모델과의 정합 분석

복원된 특징점들의 3D 버텍스를 미리 구성한 3D 변형 가능 형상 모델에 정합한 결과, 얼마나 잘 정합되었는가를 보여 주는 실험 결과가 [표 2]에 나타나 있다.

[표 2]의 데이터는 전체 버텍스 38,257개 각각에 대해, 3D 스캔 데이터와 처음 평균 형상 모델, 정합된 결과로 나오는 최종 3D 형상 모델간의 거리(에러)를 구하고 평균한 값이다. [그림 15]는 평균 형상 모델(a), 최초 정합 형상 모델(b), 최종 정합 형상 모델(c), 3D 스캔 데이터

에서 얻어진 형상 모델(d), 850개의 버텍스만의 변형 가능 형상 모델을 정합한 형상 모델(e)들을 비교하여 보여준다.

표 2. 변형 가능 형상 모델과의 정합 분석

	평균 형상모델	최종 정합 형상 모델
평균 예러	0.0106	0.0055

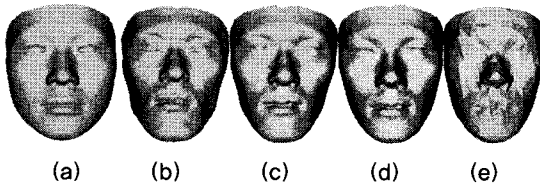


그림 15. 형상 모델 비교

[표 2]에서 보면 처음 평균 형상 모델에 대해서부터 정합을 시작하여 14번의 반복 수행으로 얻어진 최종 정합 모델에서의 예러는 많이 줄었으며, [그림 16]에서 보면 최종 형상 모델(c)이 3D 스캔 데이터 형상 모델(d)에 매우 유사함을 알 수 있다.

기존 3D 일반 모델 기반 3D 얼굴 모델링 방법이나 기존 3D 형상 모델 기반 3D 얼굴 모델링 방법에서는 사용하는 버텍스 수가 적기 때문에, 얼굴 굴곡 표현의 정확성이 떨어진다. [그림 15](e)는 버텍스 개수가 850개인 경우에 최종적으로 정합하여 얻어진 형상 모델을 보여 주며, 버텍스 개수가 적을 경우에 정합에서 얻어지는 3D 형상 모델이 표현 성능이 떨어짐을 알 수 있다.

4.5 수행 시간 분석

기존 3D 일반 모델 기반 3D 얼굴 모델링 방법 [2-4][7-9][11] 이나 기존 3D 형상 모델 기반 3D 얼굴 모델링 방법[10][13][14] 등의 연구에서는 수행 시간에 대한 실험 결과가 보고되어 있지 않다. 다만, 변형 가능 모델 기반 3D 모델링 방법에는 정합시에 펜티엄4 2GHz 프로세서에서 4.5분 정도 걸렸다고 보고되어 있다[6]. 본 논문의 실험 환경에서 펜티엄 4 코어 2 듀오 2GHz 를 사용하였으나, CPU 하나만을 이용하는 프로그래밍을 하였기 때문에 [6]의 컴퓨팅 환경과 동일하다

고 간주된다. 다음 [표 3]은 본 논문에서 제안한 방법과 기존 변형 가능 모델 기반 3D 모델링 방법 [5][6]과 수행 속도를 비교한 실험 결과이다. 본 논문의 방법이 빠르며, 단순히 정합 시간만을 비교할 때는 수행 속도 차이가 크을 볼 수 있었다. 기존 3D 일반 모델 기반 3D 얼굴 모델링 방법[2-4][7-9][11]이나 기존 3D 형상 모델 기반 3D 얼굴 모델링 방법은 본 논문보다 적은 버텍스를 사용하기 때문에 수행속도는 더 빠를 것으로 사료된다.

표 3. 수행 속도 비교

항목	변형 가능 모델 기반 3D 모델링 방법[5,6]	제안방법
2D 특징점 추출	필요 없음	11.2초 (AAM 기반, 80 프레임 이용시)
특징점의 3D 버텍스좌표 계산	필요 없음	12초 (비강체 SIM [30] 이용)
정합	270초 (변형가능 모델에 기반)	19초 (변형가능 형상 모델에 정합)
텍스처맵 브렌딩	필요 없음	110초 (5 프레임 이용시)
렌더링	모름	0.01 초 미만
총 소요 시간	270+α(>0) 초	142.21 초

V. 결론

본 논문에서는 3D 변형 가능 형상 모델에 기반을 둔 3D 얼굴 모델링 방법을 제안하였다.

제안된 방법은 먼저 3D 변형 가능 형상 모델을 구성하고, 2D 얼굴 이미지 시퀀스로부터, 해당 얼굴의 특징점들의 3D 버텍스 좌표 값을 복원하고 이 3D 버텍스들을 3D 변형 가능 형상 모델에 정합하여 3D 얼굴 형상 모델을 구한다. 이후, 2D 얼굴 이미지 시퀀스들로부터 뷰 독립적인 2D 원통 좌표 텍스처를 생성한 후에, 최종적으로 생성된 텍스처 맵을 이용하여, 구한 3D 형상 모델을 렌더링 함으로써 3D 얼굴 모델을 완성한다.

본 논문의 방법은 3D 변형 가능 얼굴 모델링 방법 [5][11]에 비해 손쉽고 빠르게 모델링이 가능하며, 3D 일반 모델기반 3D 얼굴 모델링 방법[2-4][6-8][10] 및

기존 3D 형상 모델 기반 3D 얼굴 모델링 방법 [9][12][13] 에 비해서는 좀 더 정교함을 볼 수 있었다.

향후, 1) 3D 좌표 값 복원에 있어서 보다 정확한 원근 투영 기반 SfM 기법을 적용하고, 2) 3D 얼굴 형상 모델 정합을 보다 안정적이고 정교하게 할 수 있는 모델 기반 번들 교정(model-based bundle adjustment)[12] 방법을 적용하여 본 논문에서 제안한 3D 얼굴 모델링 방법을 더욱 정교하고 강인하게 할 예정이다.

본 논문의 방법은 3D 얼굴 아바타 제작이나, 3D 얼굴 캐릭터 제작에 도움이 될 것으로 기대된다.

참고 문헌

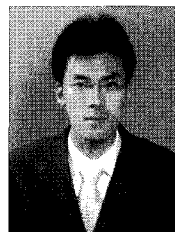
- [1] W. Zhao and R. Chellappa, *Face Processing: Advanced Modeling and Methods*, Elsevier, 2005.
- [2] Y. Lee, D. Terzopoulos, and K. Waters, "Realistic Modeling for Facial Animation," Proc. SIGGRAPH, Los Angeles, pp.55-61, 1995(8).
- [3] F. I. Parke and K. Waters, "Appendix 1: Three-dimensional muscle model facial animation," Computer Facial Animation, A.K. Peters, 1996(9).
- [4] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. H. Salesin, "Synthesizing realistic facial expressions from photographs," In Computer Graphics, Annual Conference Series, SIGGRAPH, pp.75-84, 1998(7).
- [5] V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," Proc. of the SIGGRAPH'99, Los Angeles, USA, pp.187-194, 1999(8).
- [6] V. Blanz and T. Vetter, "Face Recognition Based on Fitting a 3D Morphable Model," IEEE Pattern Anal. Mach. Intell, Vol.25, No.9, pp. 1063 - 1074, 2003(9).
- [7] J. Ahlberg, "CANDIDE-3 -- an updated parameterized face," Report No. LiTH-ISY -R-2326, Dept. of Electrical Engineering, Linköping University, Sweden, 2001.
- [8] R. L. Hsu and A. K. Jain, "Face Modeling for Recognition," Proc. Int'l Conf. Image Processing(ICIP), Vol.2, pp.693-696, 2001.
- [9] A. Ansari and M. Abdel-Mottaleb, "3-D Face Modeling Using Two Views and a Generic Face Model with Application to 3-D Face Recognition," IEEE Conf. on Advanced Video and Signal Based Surveillance, pp.203-222, 2003.
- [10] Y. Hu, D. Jiang, S. Yan, L. Zhang, and H. zhang, "Automatic 3D reconstruction for face recognition," Proc. 6th IEEE Int'l Conf. on Automatic Face and Gesture Recognition, pp.843-848, 2004.
- [11] Z. Mandun, M. Linna, X. Y. Zeng, and Y. S. Wang, "Image-Based 3D Face Modeling," Proc. of Int'l Conf. on Computer Graphics, Imaging and Visualization, pp.165-168, 200(7).
- [12] H. Guo, J. Jiang, and L. Zhang, "Building a 3D morphable face model by using thin plate splines for face reconstruction," LNCS Vol. 3338, pp.258-267, 2004.
- [13] Z. Zhang, Z. Liu, D. Adler, M. F. Cohen, E. Hanson, and Y. Shan, "Robust and Rapid Generation of Animated Faces from Video Images: A Model-Based Modeling Approach," International Journal of Computer Vision, Vol.58, No.2, pp.93-119, 2004(6).
- [14] T. Russ, C. Boehnen, and T. Peters, "3D Face Recognition Using 3D Alignment for PCA," IEEE Conf. on Computer Vision and Pattern Recognition, Vol.2, pp.1391-1398, 2006.
- [15] <http://www.cyberware.com/>
- [16] M. B. Stegmann, B. Mikkel, Gomez, and D. David : A Brief Introduction to Statistical Shape

- Analysis Technical University of Denmark, Lyngby, 2002.
- [17] T. F. Cootes, D. J. Edwards, and S. J. Taylor, "Active Appearance Models," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.23, No.6, pp.681-685, 2001(6).
- [18] 정선태, "점진적 AAM을 이용한 강인한 얼굴 윤곽 검출", *한국콘텐츠학회논문지*, pp.11-20, 2007(2).
- [19] P. Besl and McKay, N. "A Method for Registration of 3-D Shapes," *Trans. PAMI*, Vol.14, No.2, 1992.
- [20] S. Rusinkewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," *Third International Conference on 3D Digital Imaging and Modeling*, pp.145-152, 2001(6).
- [21] B. Brown and S. Rusinkewicz, "Non-Rigid Range-Scan Alignment Using Thin-Plate Splines," *Symposium on 3D Data Processing, Visualization and Transmission*, Vol.6, No.9, pp.759-765, 2004(9).
- [22] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.11, No.6, pp.567-585, 1989(6).
- [23] C. Brechbuhler, G. Gerig, and O. Kubler, O.: "Parameterization of Closed Surfaces for 3-D Shape Description," *Comp. Vision and Image Understanding*, Vol.61, Issue.2, pp.154-170. 1995.
- [24] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. "A minimum description length approach to statistical shape modeling," *IEEE Transactions on Medical Imaging*, Vol.5, Issue.5, pp.525-537, 2002(5).
- [25] R. Hartley and A. Zisserman, *Multiple Geometry in Computation Vision*, 2nd ed. Cambridge University Press, 2003.
- [26] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *International Journal of Computer Vision*, Vol.9, No.2, pp.137-154, 1992.
- [27] C. J. Poelman and T. Kanade, "A Paraperspective Factorization Method for Shape and Motion Recovery," *IEEE PAMI*, Vol.19, No.3, 1997(3).
- [28] S. Mahamud and M. Hebert, "Iterative projective reconstruction from multiple views," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol.2, pp.430-437, 2000.
- [29] T. Jebara, A. Azarbayejani, and A. Pentland, "3D structure from 2D motion," *IEEE Signal Processing Magazine*, Vol.16, Issue.3, pp.66-84, 1999(5).
- [30] L. Torresani, A. Hertzmann, and C. Bregler, "Learning Non-Rigid 3D Shape from 2D Motion," *Proc. Of Neural Information Processing Systems*, 2003.

저 자 소 개

장 용 석(Yong-Suk Jang)

준회원



• 2006년 2월 ~ 현재 : 숭실대학교 대학원 석사
<관심분야> : 생체 및 얼굴인식, 임베디드 컴퓨팅

정 선 태(Sun-Tae Chung)

정회원



• 1991년 ~ 현재 : 숭실대학교 정보통신전자공학부 교수
<관심분야> : 생체인식, 컴퓨터 비전, 영상 감시, 임베디드 컴퓨팅

김 부 균(Boo-Gyoun Kim)

정회원



- 1979년 2월 : 서울대학교 전자공학과 학사
- 1989년 8월 : Dept. of E.E., U.S.C., 공학박사
- 1981년 3월 ~ 현재 : 숭실대학교 정보통신전자공학부 교수

<관심분야> : 위성 배열 안테나, SiP, 광통신 및 광네트워크용 소자

조 성 원(Seongwon Cho)

정회원



- 1982년 2월 : 서울대 전기공학과 학사
- 1992년 2월 : Purdue University 전기공학과 박사
- 현재 : 홍익대 전자전기공학부 교수

<관심분야> : 인공지능, 생체인식, 지능형 시스템, 패턴인식