

# 역전가능 메쉬워프 알고리즘에 의한 정면 얼굴 영상의 포즈 변형

## (Pose Transformation of a Frontal Face Image by Invertible Meshwarp Algorithm)

오 승 태<sup>†</sup> 전 병 환<sup>\*\*</sup>  
(Seung Taek Oh) (Byung Hwan Jun)

**요 약** 본 논문에서는 기하학적인 3차원 모델을 사용하지 않고 정면의 얼굴 영상 및 2차원 메쉬만으로 얼굴의 포즈 변형을 수행하는 영상기반 렌더링(Image Based Rendering: IBR) 기법을 제안한다. 3차원 기하학적 모델을 대신하기 위해, 먼저 표준 인물의 정면, 좌우 반측면, 좌우 측면의 얼굴 영상에 대한 표준 메쉬를 작성한다. 합성하고자 하는 임의의 인물에 대해서는 주어진 정면 얼굴 영상의 메쉬만을 작성하고, 그 밖의 메쉬는 표준 메쉬 집합을 근거로 자동 생성된다. 그런 다음, 메쉬 제어점들의 중첩 및 역전을 허용 하도록 개선한 역전가능 메쉬워프 알고리즘(invertible meshwarp algorithm)을 이용하여 얼굴의 입체적인 회전 변형을 수행한다. 또한, 눈이나 입의 개폐 변형도 동일한 워핑 알고리즘으로 구현한다.

얼굴 변형 성능을 평가하기 위해, 총 10명으로부터 머리를 수평으로 회전하면서 동영상을 취득한 후, 실제 영상과 이에 대응하는 변형된 영상에서 14개 주요 특징점들의 위치 오차를 측정하였다. 즉, 대응하는 실제 영상과 변형 영상마다 양 눈의 중간 위치인 기준점에서 각 특징점까지의 거리를 계산하여 평균 차이를 구하였다. 그 결과, 기준점에서 입의 중간 위치까지의 거리에 비해 약 7.0%의 평균 위치 오차만이 발생하였다.

**키워드** : 역전가능 메쉬워프 알고리즘, 얼굴 모델링, 얼굴 애니메이션, 영상 워핑, 영상기반 렌더링

**Abstract** In this paper, we propose a new technique of image based rendering (IBR) for the pose transformation of a face by using only a frontal face image and its mesh without a three dimensional model. To substitute the 3D geometric model, first, we make up a standard mesh set of a certain person for several face sides: front, left, right, half-left and half-right sides. For the given person, we compose only the frontal mesh of the frontal face image to be transformed. The other mesh is automatically generated based on the standard mesh set. And then, the frontal face image is geometrically transformed to give different view by using Invertible Meshwarp Algorithm, which is improved to tolerate the overlap or inversion of neighbor vertexes in the mesh. The same warping algorithm is used to generate the opening or closing effect of both eyes and a mouth.

To evaluate the transformation performance, we capture dynamic images from 10 persons rotating their heads horizontally. And we measure the location error of 14 main features between the corresponding original and transformed facial images. That is, the average difference is calculated between the distances from the center of both eyes to each feature point for the corresponding original and transformed images. As a result, the average error in feature location is about 7.0% of the distance from the center of both eyes to the center of a mouth.

**Key word** : invertible meshwarp algorithm, face modeling, face animation, image warping, image based rendering (IBR)

· 본 연구는 한국과학재단 지정 공주대학교 차원재원용 신소재 연구센터의 지원에 의한 것입니다.

† 비 회 원 : (주)라스텍  
ost@rastech.co.kr

\*\* 종신회원 : 공주대학교 정보통신공학부 교수  
bhjun@kongju.ac.kr

논문접수 : 2001년 10월 25일  
심사완료 : 2003년 1월 16일

1. 서론

컴퓨터 그래픽 기술의 발전과 컴퓨터의 성능 향상 그리고 게임 및 애니메이션 산업의 발달로 인해 가상현실의 구현에 대한 관심이 높아지고 있다. 그러나 물체나 장면을 직접 구현하는 전통적인 3차원 컴퓨터 그래픽 기법은 모델링 과정에서 많은 비용과 시간이 소요되며, 색상 및 질감에 대한 사실적 표현이 매우 어렵다는 단점을 가진다. 그림 1의 (a)는 전통적인 방법보다 다소 단순해진 3차원 모델에 영상을 텍스처로 사용함으로써 보다 사실적인 모델을 구축하는 기하기반 접근방법(geometry-based approach)이다. 이러한 방식의 얼굴 모델링은 주로 3차원 스캐너로 취득한 표면의 좌표 데이터(range data)를 토대로 폴리곤을 구축하며, 정적인 모델은 유용하지 않으므로 애니메이션을 지원하는 방식이어야 한다[1]. 이때, 일반 얼굴 폴리곤과 주어진 인물의 얼굴 영상을 적응적으로 정확하게 매핑하는 방법이 중요하다. 예로써, Ekman과 Friesen[2]이 개발한 Facial Action Coding System(FACS)을 토대로 근육기반 애니메이션(muscle-based animation)을 지원하는 물리적인 모델을 채택한 Lee 등[3]은, 3차원 스캐너로 취득한 얼굴 영상에서 적응을 위한 예지 특징을 추출하여 이용한다. 그러나 이를 위해서는 각 인물마다 3차원 영상을 취득해야 한다. 이와 달리, 근육과 같은 물리적 모델이 아니라 단순한 표면 폴리곤 모델을 사용하면서 표정을 포함한 애니메이션을 연출하는 방법들도 있다. Liu 등[4]은 두 개의 비디오 카메라로 취득한 첫 영상 쌍에 대해 눈, 코, 입의 주요 특징점 5개를 수작업으로 클릭하면 중립 얼굴 폴리곤을 각 영상에 맞게 재구성할 수 있는 방법을 제안하였다. 이를 위해서는 실제 인물의 행위가 요구된다. pighin 등[5]은 주요 포즈마다 대응하는 13개의 주요 특징점을 수작업으로 표시하여 각 표정에 따른 모델 적합(model fitting)을 수행하고 이를 보간하여 다양한 표정을 생성하는 방법을 제안하였다. 그러나 이를 위해서는 약속된 포즈와 표정을 위한 각 인물의 협조가 요구된다.

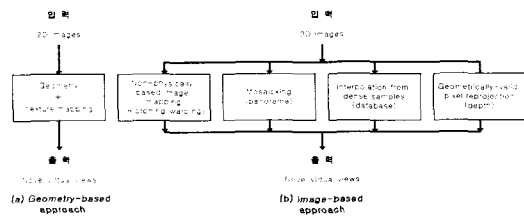


그림 1 기하기반 접근방법과 영상기반 접근방법의 비교

최근에는 3차원 모델 없이 미리 취득한 2차원 영상들로 새로운 가상 뷰(view)를 복원할 수 있는 영상기반 접근방법(image-based approach) 즉, 영상기반 렌더링(Image-Based Rendering; IBR)[6]이 활발하게 연구되고 있다. 3차원 모델을 사용하는 방법과 달리, 이 방식은 처리 속도가 물체나 장면의 복잡도에 무관하다는 장점을 갖는다.

그림 1의 (b)에는, Kang[6]이 분류한 네 가지 IBR 접근방법 즉, 비물리적 영상 매핑(non-physically based image mapping), 모자이킹(mosaicking), 방대한 샘플로부터의 보간(interpolation from dense samples), 기하학적으로 타당한 화소 재투영(geometrically-valid pixel reprojection)이 나타나 있다. 첫째, 비물리적 영상 매핑은 3차원 기하를 전혀 고려하지 않는 모핑(morphing)을 의미한다. 예로써, Vetter와 Poggio[7]는 먼저 여러 사람들의 얼굴 영상을 방대한 각 포즈별로 취득하여 학습 집합을 구성한 후, 새로운 얼굴을 합성하기 위한 참조 인물별 가중치를 계산하여, 이를 원하는 포즈의 모핑에 적용함으로써 새로운 뷰를 생성하는 방법을 제안하였다. 그러나 이 방법은 방대한 양의 참조 영상들을 요구할 뿐만 아니라 단지 참조 얼굴들의 합성만이 가능하다.

둘째, 모자이킹은 일반적으로 둘 이상의 영상을 조합하여 보다 큰 영상을 만드는 것을 의미하는데, 대표적으로 파노라마(panorama)가 있다. 예로써, 고정되지 않은 가변 위치에서 시각화할 수 있는 McMillan과 Bishop[8]의 'plenoptic modeling'은 실린더 형태의 파노라마 이미지를 생성해낸다. 한편, 상부와 하부의 영상을 포함시킬 수 없는 실린더 방식의 단점을 극복하기 위해, 다른 편향(tilt)과 팬(pan) 방향에서 통상적인 렌즈의 카메라로, 투영 중심에서 회전하면서 임의 형태의 물체를 찍거나 임의 위치에서 평탄한 장면을 찍은 영상들을 조합하여 구형 영상(spherical image)을 만드는 방법[9]도 발표되었다. 그러나 이와 같은 촬영 방법의 제한으로, 이 방법은 얼굴과 같은 물체보다는 주로 장면을 재구성하는데 적용된다.

셋째, 방대한 샘플로부터의 보간은 서로 다른 많은 카메라 시점에서 물체나 장면의 영상들을 찍어 룩업 테이블을 만들고, 이를 보간하여 주어진 임의의 시점에서의 영상을 합성하는 기법이다. 예로써, 'light field rendering'[10]과 'lumigraph'[11]는 모든 위치의 모든 방향에서의 빛의 흐름을 묘사하는 완전한 plenoptic function의 부분집합인 시점과 방향의 4D 파라미터화를 사용하는데, 이를 각각 light slab 혹은 lumigraph라 부른다. 이 방법

은 화소 대응을 필요로 하지 않는 반면, 각 영상에 대한 카메라 시점정보를 포함한 매우 방대한 영상 데이터를 요구한다. 따라서 효율적인 얼굴 모델링을 위해서는 적합하지 않다.

넷째, 기하학적으로 타당한 화소 재투영은 3차원 모델과 기하학적 카메라 정보를 필요로 하지 않고 상대적으로 적은 수의 영상을 사용하고 한다는 특성을 갖는다. 예로써, Seitz와 Dyer[12]은 두 영상에 대한 epipolar constraints 하에서 대응하는 스캔라인이 평행이 되도록 영상을 교정하는 전워프, 교정된 중간 영상 생성을 위한 선형 보간, 그리고 그 중간 영상을 역교정하는 후 워핑을 수행하여 뷰 모핑(view morphing)을 구현하였다. 이 방법으로는 두 참조 영상 사이의 중간 포즈 영상을 생성할 수 있다. 또한, Avidan과 Shashua[13]는 광류(optic-flow)와 세 영상간의 대응을 연결하는 3선형 텐서(trilinear tensor)를 사용하여 두 참조 영상으로 다양한 새로운 뷰를 생성하는 방법을 제안하였다. 그러나 텐서를 계산하기 위해서는, 두 참조 영상간의 집중적인 대응과 원하는 세 번째 영상과의 적어도 7개의 대응점을 명시할 수 있어야 한다.

한편, 기하기반 접근방법과 영상기반 접근방법을 혼합한 예로써, 사진측량적 모델링(photogrammetric modeling)과 모델기반 스테레오 알고리즘(model-based stereo algorithm)을 이용하여 매우 단순한 3차원 기본 모델과 기존 IBR 기법에 비해 훨씬 적은 수의 영상을 사용하는 하이브리드 접근방법(hybrid approach)[14]도 소개되었다. 그러나 이 방법은 얼굴과 같이 복잡한 대상이 아닌 건축물과 같이 비교적 단순한 형태의 물체 모델링에 적합하다.

본 논문에서는 게임이나 채팅과 같은 응용에서의 사실적인 3차원 아바타 얼굴을 구현할 수 있도록, 다수의 참조 영상을 사용하는 기존의 IBR과 달리, 단 한 장의 정면 얼굴 영상만을 가지고 얼굴 회전과 눈입 개폐에 따른 새로운 포즈를 생성할 수 있는 방법을 제안한다. 이를 위해, 특정 인물로부터 주요 포즈에 대한 표준 메쉬 집합을 작성하고, 회전 효과를 위한 메쉬간의 중첩이나 역전을 허용하도록 기존의 메쉬워프 알고리즘을 개선하였다.

## 2. 역전가능 메쉬워프 알고리즘

워핑(warping)이란 하나의 그래픽 객체를 임의의 형태로 변형시키는 작업으로, 컴퓨터비전, 의료 영상 분석, 원격 감지, 특수 효과, 시각화 분야 등에서 활발히 응용되고 있는 기법이다. 특히, 이미지 워핑(image warping)은 입력 영상에서의 임의의 화소를 특정 규칙에 따라 이

동시켜 원하는 목적 영상으로 변형시키는 것이다. 예로써, 크게 점 기반(point-based), 벡터/범위 기반(vector/field-based), 비정형 좌표 변환(free-form coordinate change), 그리고 메쉬 기반(mesh-based) 워핑 등이 있다[15].

### 2.1 두 단계 메쉬워프 알고리즘

본 논문에서는 단순히 얼굴의 모양을 변화시키는 것이 아니라 입체적인 회전 효과를 연출하고자 하기 때문에, 얼굴의 윤곽선을 감싸서 유지시킬 수 있는 Douglas Smythe가 제안한 두 단계 메쉬워프 알고리즘(two-pass meshwarp algorithm)[16]을 선택하였다. 첫 번째 단계에서는 영상의 매 행을 따라 처리하는 것으로 수직 스플라인 값을 제어하며, 두 번째 단계에서는 영상의 매 열을 따라 처리하는 것으로 수평 스플라인을 제어함으로써 결과 영상을 획득한다. 그림 2는 메쉬워프 알고리즘의 전체적인 흐름과 행·열 처리 단계 내의 세부적인 진행과정을 나타낸다.

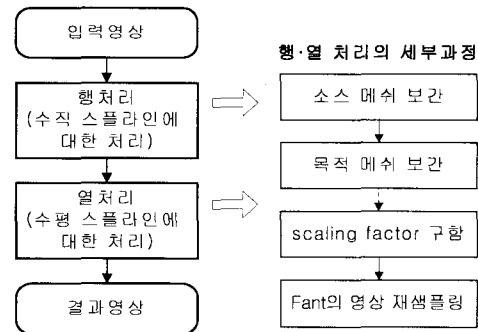


그림 2 두 단계 메쉬워프 알고리즘

### 2.2 메쉬워프 알고리즘의 개선

그러나 기존의 워핑 기법은 대부분 모핑 기법을 함께 사용하여 주로 같은 포즈를 취하는 두 객체 사이의 영상 변형에 사용되어져 왔다. 따라서 객체가 입체적으로 회전하면서 발생하는 메쉬 점이 중첩되거나 역전되는 현상을 고려할 필요가 없었다. 그러나, 본 논문에서는 2차원의 정면 영상 및 메쉬로 자연스러운 얼굴의 회전 변형을 구현하고자 하기 때문에, 얼굴 경계나 코의 능선에 위치하는 메쉬 점들이 한 곳으로 중첩되는 현상이 발생할 수 있으며, 회전 정도가 더 커지면 볼의 윤곽이 코에 가려지는 역전 현상도 발생할 수 있다. 그림 3은 메쉬의 인접, 중첩, 역전 현상의 예를 보여준다. (a)는 목적 메쉬에 이웃한 두 메쉬점 B와 C가 인접하는 현상을 보여주는데, 이는 기존의 메쉬워프 알고리즘으로도

처리할 수 있다. (b)는 중첩이 발생하여 B와 C가 동일한 메쉬점이 되는 경우이고, (c)는 역전 현상의 예로서, 메쉬점 순서 A-B-C가 A-C-B로 달라졌다.

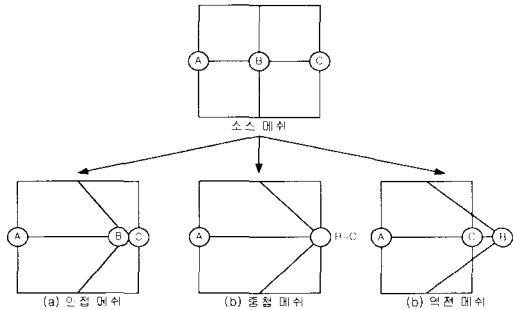


그림 3 메쉬의 인접, 중첩, 역전 현상

이와 같은 중첩 메쉬나 역전 메쉬가 발생할 경우에는, 메쉬 점들이 떨어져있다고 가정하고 보간 함수를 호출하는 기존의 메쉬워프 알고리즘으로는 각 행·열마다 제어점을 생성하기 위한 보간이나, 영상 재생플링 단계에서의 스케일 요소 계산을 위한 보간이 불가능하다.

본 논문에서는 이러한 문제를 해결할 수 있도록, 메쉬 제어점들의 중첩 및 역전을 허용하도록 개선한 "역전가능 메쉬워프 알고리즘(*invertible meshwarp algorithm*)"을 제안한다. 예로써, 그림 4는 소스 메쉬의 한 라인(행 또는 열) 상에서 연속된 제어점들이 회전에 의해 목적 메쉬의 라인 상에서 역전되는 경우의 처리과정을 보여준다. 소스 라인 상에서의 제어점의 순서 A-B-C-D가 목적 라인 상에서 A'-C'-B'-D'로 제어점 간의 역전 현상이 발생하게 되면 역전된 제어점 C'를 이전 제어점 B'에 중첩시킴으로써(B'=C') 역전 현상은 사라지게 된다. 이제 남은 중첩 현상을 처리하기 위해서는, A'부터 B'까지 처리한 후 중첩된 C'를 건너뛰고 D'으로 진행하면 된다. 즉, A'-B'-D'의 순서로 제어점을 처리하게 된다. 이는 전통적인 3차원 컴퓨터 그래픽 기법에서의

은면 또는 은선 제거에 해당한다.

### 3. 얼굴의 회전에 따른 포즈 변형

머리 회전은 고개를 상하로 끄덕이는 수직방향 회전(x축 회전), 고개를 좌우로 가로짓는 수평방향 회전(y축 회전), 고개를 좌우로 기울어놓듯이 기울이는 회전(z축 회전)으로 구분할 수 있는데, 본 논문에서는 점진적인 수평 회전에 대해서 자연스러운 얼굴의 포즈 변형을 시도하였다. 이때, 회전 각도  $-90^\circ \sim 90^\circ$ 를 구간  $-1.0 \sim 1.0$ 로 정규화한 회전 정도  $\delta$ 로 표현한다.

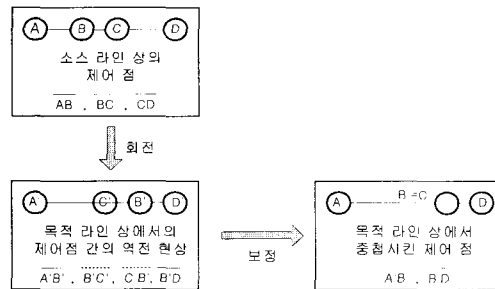
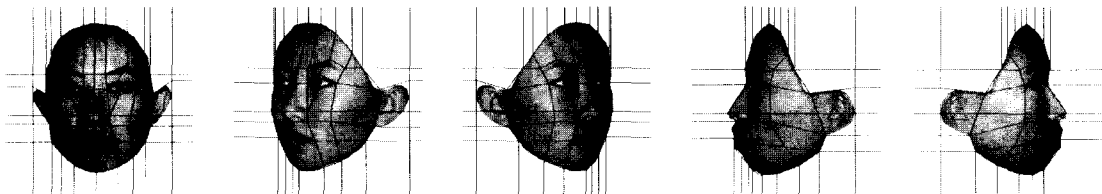


그림 4 역전 메쉬의 처리 과정

#### 3.1 표준 메쉬 집합

기하학적인 3차원 모델을 대신하기 위해서는, 그림 5와 같이, 표준 인물로부터 취득한 주요 포즈의 얼굴 영상에 대해 표준 메쉬 집합  $S = \{S_f, S_{fl}, S_r, S_{fr}, S_t\}$ 을 마련해야 한다. 먼저, 표준 인물의 각 포즈 영상에서 가급적 머리카락을 포함하지 않고 부드러운 윤곽선을 유지하도록 얼굴 영역을 잘라낸다. 메쉬는 얼굴의 윤곽선과 주요 구성요소인 눈, 눈썹, 코, 입을 감쌀 수 있도록  $13 \times 12$  크기로 설정한다. 이때, 수직 메쉬선의 경우는 얼굴 중심, 눈썹/코/입의 가장자리, 눈썹 상단, 눈동자 중심, 광대뼈 중앙, 귀와의 교차점 등을 통과하도록 설정하고, 수평 메쉬선의 경우는 이마 상단, 눈썹 상단, 눈의



(a) 정면 표준 메쉬( $S_f$ ) (b) 반좌측 표준 메쉬( $S_{fl}$ ) (c) 반우측 표준 메쉬( $S_{fr}$ ) (d) 좌측 표준 메쉬( $S_r$ ) (e) 우측 표준 메쉬( $S_t$ )

그림 5 표준 메쉬 집합 S

위 및 아래, 코끝과 코밑, 입의 위, 중간, 아래, 그리고 턱선 등을 통과하도록 설정한다. 특히, 눈과 입의 경우는 마름모형 메쉬에 내포시키고 있으며, 좌측과 우측의 얼굴 메쉬에서는 코에 인접한 볼 부분에 역전 메쉬를 적용하고 있다.

또한, 정면 표준 메쉬  $S_f$ 와 기타 표준 메쉬  $S_i$ 의 위치 차이  $\Delta_i$ 를 미리 계산해 놓는다. 이는 대상 인물의 참조 메쉬를 생성하는데 사용된다.

3.2 회전 변형 과정

그림 6은 본 논문에서 제안하는 얼굴 회전 변형의 전체 흐름도 이고, 그림 7은 주어진 인물의 정면 영상을 가지고 원하는 포즈로 변형하는 과정을 도식적으로 나타낸 것이다.

먼저, 단계 1에서는 변형에 사용되는 인물의 정면 얼굴 영상을 CCD 카메라, 디지털 카메라, 스캐너 등으로 입력받아 얼굴 영역을 오려내어 정면 얼굴 영상  $I_f$ 를 마련한다. 단계 2에서는 이에 대응하는 정면 참조 메쉬  $M_f$ 를 작성한다. 단계 3에서는, 앞서 언급한 정면 표준 메쉬  $S_f$ 와 기타 표준 메쉬  $S_i$ 의 위치 차이  $\Delta_i$ 를 변형하고자 하는 대상 인물의 정면 참조 메쉬  $M_f$ 에 더함으로써, 나머지 참조 메쉬들  $M_{lh}$ ,  $M_l$ ,  $M_{lr}$ ,  $M_r$ 이 자동 생성된다. 단계 4에서는 원하는 회전 정도  $\delta$ 에 가장 근접한 두 참조 메쉬를 선형 보간하여 목적 메쉬를 생성한다. 끝으로, 단계 5에서는 정면 영상  $I_f$ 와 정면 메쉬  $M_f$ 를 소스로 사용하여 제안한 역전가능 메쉬워프 알고리즘에 의해 회전 변형된 결과 영상을 산출한다.

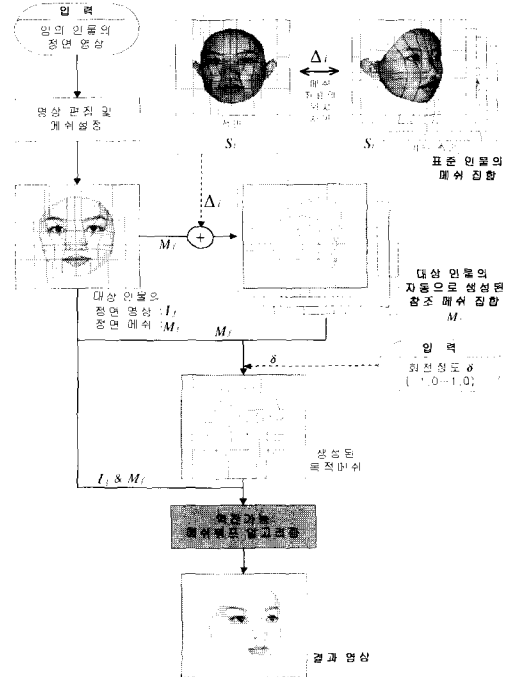


그림 7 얼굴 회전 변형의 과정

4. 눈과 입의 개폐 변형

이 절에서는, 얼굴의 회전 변형에 적용한 동일한 역전가능 메쉬워프 알고리즘이 표정 변형을 포함한 얼굴 애니메이션에 적용할 수 있음을 보이기 위해, 열림 정도에 따른 눈과 입의 개폐 변형에 적용하였다. 이때, 눈과 입이 완전히 열려 있는 경우를 열림 정도 1.0으로 정의하고 완전히 닫힌 상태를 열림 정도 0.0으로 정의하여, 눈과 입의 열림 정도  $\delta_{eye}$ 와  $\delta_{mouth}$ 를 0.0~1.0의 정규화된 값으로 표현한다. 그림 8은 실제 인물의 눈과 입의 대표적인 상태를 보여준다.

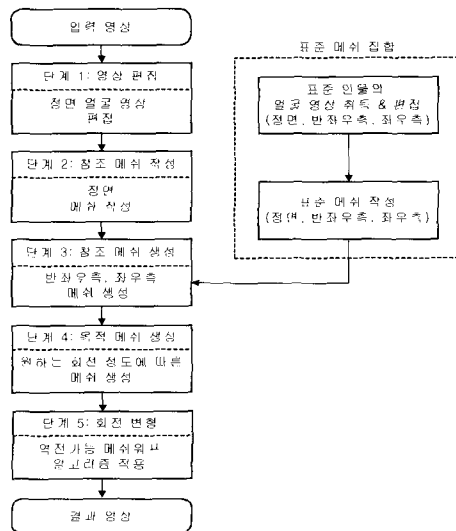


그림 6 얼굴 회전 변형의 전체 흐름도

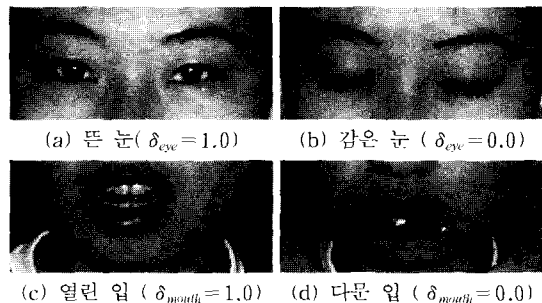


그림 8 눈과 입의 대표적인 상태

4.1 눈의 개폐 변형

실제 인물이 눈을 감을 때 눈꺼풀을 당겨 눈을 덮는 것과 유사하게, 눈의 개폐 효과는 상단에 있는 눈꺼풀 메쉬 점들의 상하 이동으로 연출할 수 있다. 눈꺼풀이 자연스럽게 눈을 덮기 위해서는 그림 9와 같이 최소 7×4의 메쉬가 필요하며, 그림 10은 실제 눈 영상에 메쉬를 설정한 모습이다.

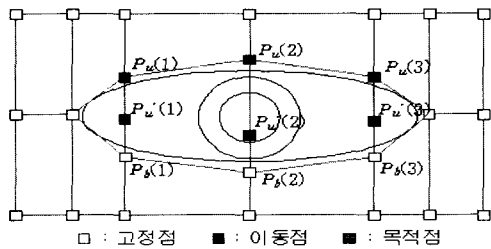
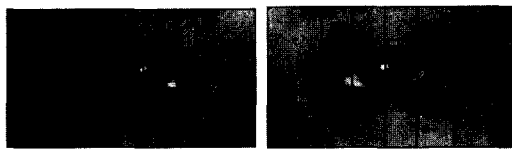


그림 9 눈 메쉬의 구조



(a) 좌측 눈                      (b) 우측 눈

그림 10 눈 영상의 소스 메쉬 설정

눈 영상에 설정한 소스 메쉬의 메쉬점을  $P(i) = (x(i), y(i))$ 로 나타내면, 목적 메쉬의 각 메쉬점  $P'(i) = (x'(i), y'(i))$ 는 눈의 열림 정도  $\delta_{eye}$ 에 따라 식 (1)과 같이 계산한다. 이때, 첨자 u와 b는 해당 메쉬점이 각각 눈의 위와 아래 경계 상에 위치한다는 것을 의미한다.

$$\begin{aligned}
 x_u'(i) &= x_u(i) \\
 y_u'(i) &= y_u(i) + (y_b(i) - y_u(i)) \times (1 - \delta_{eye}) \\
 0.0 \leq \delta_{eye} \leq 1.0, \quad 1 \leq i \leq 3
 \end{aligned}
 \tag{1}$$

4.2 입의 개폐 변형

실제 인물이 입을 열고 닫을 때 움직이는 윗입술, 아랫입술, 그리고 턱에서 발생한다[11]. 본 논문에서는 얼굴 윤곽을 변형시키는 턱의 움직임을 제외하고, 아랫입술과 윗입술의 경계에 해당하는 메쉬 점들의 상하 이동으로 입의 개폐 효과를 연출한다. 아랫입술과 윗입술의 자연스러운 이동을 위해서는 그림 11과 같이 최소 9×6 크기의 메쉬가 필요하며, 그림 12는 실제 입 영상에 메쉬를 설정한 모습이다.

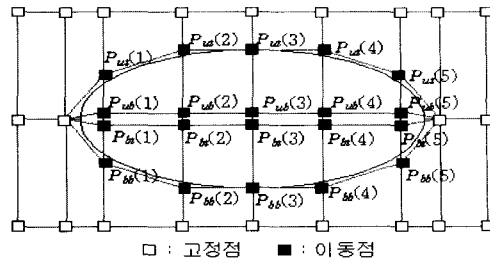


그림 11 입 메쉬의 구조



그림 12 입 영상의 소스 메쉬 설정

입의 벌림은 눈과 달리 얼굴의 다른 부분에 영향을 미칠 수 있으므로, 최대로 벌어지는 높이  $D_{max}$ 를 코밑과 턱선 사이에서 상수로 정해야 한다. 각 이동점의 이동 거리는 입술의 중앙에서 가장 크고 가장자리로 갈수록 줄어들기 때문에, 식 (2)와 같이, 입술의 중앙 이동점으로부터의 수평거리에 반비례하도록 각 이동점  $P(i)$ 의 최대 수직 이동 거리  $d_{max}(i)$ 를 계산한다. 이때, 각 입술의 위 메쉬점과 아래 메쉬점을 동일한 길이로 이동시킴으로써 입술의 두께를 일정하게 유지하도록 한다.

$$d_{max}(i) = D_{max} \times \frac{L_m - |x(3) - x(i)|}{L_m}
 \tag{2}$$

$$L_m = \max \{ |x(3) - x(1)|, |x(3) - x(5)| \}$$

그런 다음, 식 (3)과 같이, 입의 열림 정도  $\delta_{mouth}$ 에 비례하도록 각 이동점의 수직 이동 거리  $d(i)$ 를 구한 후, 식 (4)와 같이, 윗입술의 이동거리  $d_u(i)$ 와 아랫입술의 이동거리  $d_b(i)$ 를 계산한다. 이때,  $d(i) = d_u(i) + d_b(i)$ 이다.

$$d(i) = \delta_{mouth} \times d_{max}(i), \quad 0.0 \leq \delta_{mouth} \leq 1.0
 \tag{3}$$

$$d_u(i) = \alpha d(i)
 \tag{4}$$

$$d_b(i) = (1 - \alpha)d(i) \quad 0 \leq \alpha \leq 0.5$$

따라서, 소스 메쉬 상의 각 이동 메쉬점  $P(i) = (x(i), y(i))$ 의 이동할 목표 위치 즉, 목적 메쉬 상의 좌표  $P'(i) = (x'(i), y'(i))$ 는 다음의 식 (5)와 같다. 이때, 첨자  $u$ 와  $b$ 는 각각 해당 메쉬점이 윗입술의 위 경계와

아래 경계에 위치함을 의미하고, 첨자 bt와 bb는 각각 아랫입술의 위 경계와 아래 경계에 위치함을 의미한다.

$$\begin{aligned}
 y_{at}'(i) &= y_{at}(i) - d_a(i), & y_{ab}'(i) &= y_{ab}(i) - d_a(i) \\
 y_{bt}'(i) &= y_{bt}(i) + d_b(i), & y_{bb}'(i) &= y_{bb}(i) + d_b(i)
 \end{aligned}
 \quad 1 \leq i \leq 5 \quad (5)$$

**5. 실험 및 분석**

성능 분석 실험을 위해, 총 10명의 대상자를 회전 의자에 앉혀 회전시키면서 CCD 카메라 혹은 디지털 캠코더로 동영상상을 취득하였다.

**5.1 눈과 입의 개폐 변형**

그림 13은 눈과 입의 열림 정도에 따른 변형 결과를 보여준다. 눈의 경우에는 열림 정도  $\delta_{eye}$ 에 따라 고른 변형 결과를 보여주고 있으나, 입의 경우에는 열림 정도  $\delta_{mouth}$ 가 증가하면서 치아나 혀와 같은 입안의 모습이 나타나지 않아 사실감이 떨어진다. 이를 위해서는, 입안 영상과의 합성이 필요하다.

**5.2 반측면 메쉬의 역할**

본 논문에서는 두 참조 메쉬를 선형 보간하여 중간 목적 메쉬를 생성하는 방식을 채택하고 있기 때문에, 만일 참조 메쉬의 형태가 크게 차이가 난다면 형태의 왜곡이 심하게 발생할 수 있다. 이러한 왜곡 현상을 방지하기 위해서는, 정면과 측면 메쉬 이외에 반측면 메쉬와 같은 중간 메쉬를 추가하고 회전 정도에 따라 선택적으로 사용할 수 있어야 한다.

그림 14는 회전 정도가  $\delta = \pm 0.6$ 일 때 반측면 메쉬의 유무에 따른 변형 결과를 비교하여 보여주고 있다. 반측면이 있는 경우에는 전체적으로 자연스러운 점진적인 수평회전 변형을 보이고 있으나, 반측면이 없는 경우에는 변형 결과 영상의 왜곡이 심함을 알 수 있다.

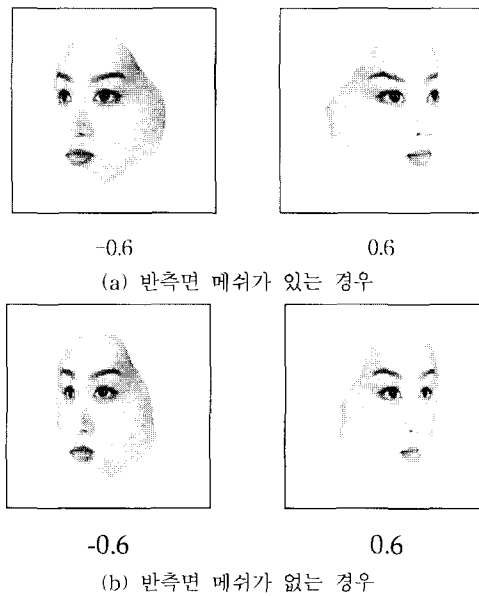


그림 14 반측면 메쉬의 유무에 따른 회전 변형 결과

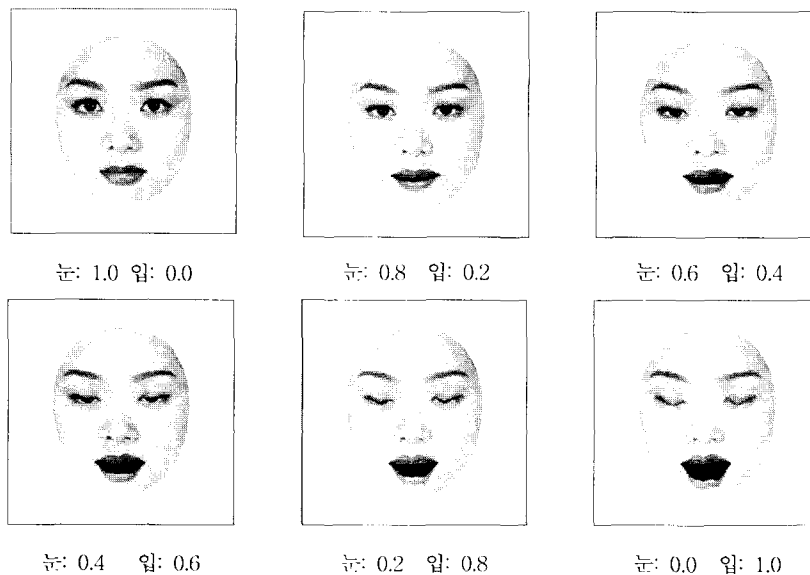
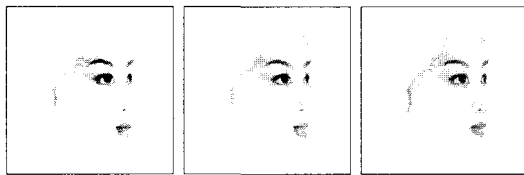


그림 13 눈과 입의 열림 정도에 따른 변형 결과

**5.3 인접, 중첩, 역전 메쉬의 성능 비교**

그림 15는 인접, 중첩, 역전 메쉬에 따라 회전 정도  $\delta=0.8$ 일 때의 회전 변형한 결과를 보여준다. 기존의 메쉬워프 알고리즘은 인접 메쉬만을 허용하므로, (a)와 같이, 코에 가려져야 될 볼 부분이 코끝보다 더 튀어나오는 매우 부자연스러운 결과를 초래한다. 중첩 메쉬까지만 허용하면서 제한한 역전가능 메쉬워프 알고리즘을 적용한 경우에는, (b)와 같이, 볼의 경계가 코끝과 중첩되면서 여전히 뽀족해지는 현상이 발생한다. 이와 달리, 역전 메쉬까지 허용한 경우에는, (c)와 같이, 보다 자연스러운 결과 영상을 얻을 수 있다.



(a) 인접 메쉬 (b) 중첩 메쉬 (c) 역전 메쉬

그림 15 메쉬의 인접, 중첩, 역전 허용에 따른 회전 변형 결과( $\delta=0.8$ )

실제로 변형된 결과 영상과 이에 대응하는 실제 영상과의 차이를 정량적으로 비교하기 위해, 그림 16에 나타난 바와 같이, 코의 능선과 얼굴 경계에 위치하는 특징점들 간의 거리 두 곳을 측정한다. 먼저, 변형된 결과 영상에서 회전 방향에 대해 반대쪽에 위치하는 눈의 중심  $S_E$ 를 지나는 수직선과 코끝을 지나는 수평선이 교차하는 점을 기준점  $S$ 로 정한다. 이때, 이 수평선이 코의 능선과 교차하는 점(코끝)을  $Q_{11}$ 이라 하고, 얼굴의 윤곽선과 교차하는 점을  $Q_{12}$ 라고 하자. 이와 유사하게,  $S_E$ 와  $S$ 의 중간 지점을 지나는 또 하나의 수평선을 곱고 코의 능선과 얼굴 경계와 교차하는 점을 각각  $Q_{21}$ 과  $Q_{22}$ 이라 하자. 또한,  $S$ 와  $S_E$  사이의 거리를 기준 거리  $D$ 로,  $Q_{11}$ 과  $Q_{12}$  사이의 거리를  $d_1$ 로, 그리고  $Q_{21}$ 과

$Q_{22}$  사이의 거리를  $d_2$ 로 정한다. 한편, 이에 대응하는 실제 영상에서의 거리를 각각  $D'$ ,  $d_1'$ ,  $d_2'$ 이라 하자. 그러면, 코 부분에서 변형 영상과 실제 영상 사이의 정규화된 오차는 식 (6)과 같이 정의할 수 있다.

$$E_Q = \left( \left| \frac{d_1}{D} - \frac{d_1'}{D'} \right| + \left| \frac{d_2}{D} - \frac{d_2'}{D'} \right| \right) / 2 \quad (6)$$

여기서,  $d_1, d_2$ 와  $d_1', d_2'$ 의 차를 직접 계산하지 않고 각각 기준거리  $D$ 과  $D'$ 에 대해 정규화된 값의 차를 계산함으로써, 변형 영상과 실제 영상에서 얼굴의 크기가 일치하지 않는 경우에도 오차를 제대로 계산할 수 있다.

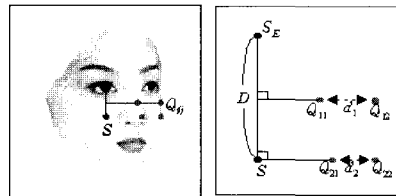


그림 16 코 주변의 주요 특징점과 측정 거리

총 10명으로부터  $\pm 0.5 \leq \delta \leq \pm 1.0$ 일 때 취득한 영상에 인접가능, 중첩가능, 역전가능 메쉬를 각각 적용하여 변형 영상을 120장씩 취득하고, 대응하는 실제 영상과의 평균 거리 오차  $E_Q$ 를 표 2와 같이 구하였다. 회전 정도가 작은 경우에는 변형의 성능이 비슷하지만, 회전 정도가 큰 경우( $\pm 0.8 \sim \pm 1.0$ )에는 역전가능 메쉬를 사용할 때 거리 오차가 현격하게 감소함을 관찰할 수 있다. 즉, 회전 정도가 증가할수록 제안한 역전가능 메쉬워프 알고리즘이 보다 우수한 변형을 수행하는 것으로 나타났다.

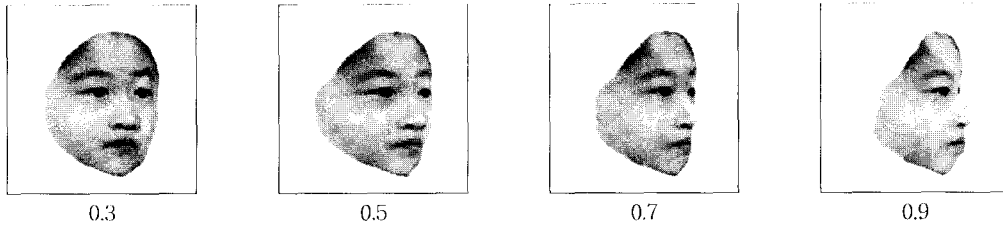
**5.4 얼굴의 회전 변형 성능**

다양한 연령과 성별에 따라, 제한한 역전가능 메쉬워프 알고리즘을 적용한 점진적인 회전 변형 결과를 그림 17~그림 21에서 볼 수 있다.

표 1 실제 및 변형 영상에 대한 코 부분에서의 위치 오차

회전 정도 $\delta$	위치 오차 $E_Q$			회전 정도 $\delta$	위치 오차 $E_Q$		
	인접	중첩	역전		인접	중첩	역전
0.5	0.15	0.15	0.17	-0.5	0.18	0.19	0.12
0.6	0.15	0.13	0.17	-0.6	0.26	0.18	0.17
0.7	0.16	0.10	0.12	0.7	0.37	0.20	0.18
0.8	0.23	0.18	0.09	-0.8	0.47	0.27	0.14
0.9	0.23	0.30	0.05	-0.9	0.47	0.26	0.09
1.0	0.20	0.12	0.01	1.0	0.45	0.24	0.04





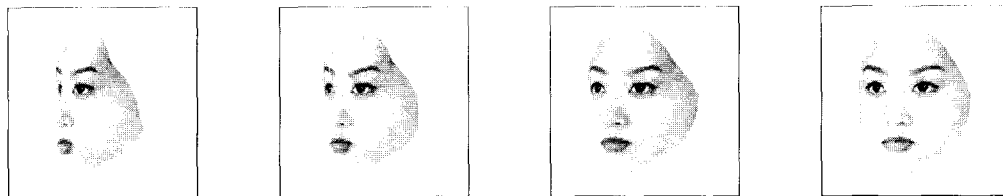
0.3

0.5

0.7

0.9

그림 17 제안한 방법에 의한 회전 변형 예: 어린이



-0.9

-0.7

-0.5

-0.3

그림 18 제안한 방법에 의한 회전 변형 예: 성인 여자



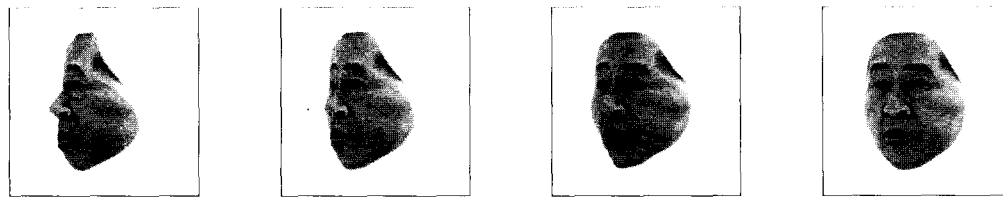
0.3

0.5

0.7

0.9

그림 19 제안한 방법에 의한 회전 변형 예: 성인 남자



-0.9

-0.7

-0.5

-0.3

그림 20 제안한 방법에 의한 회전 변형 예: 남성 노인



0.3

0.5

0.7

0.9

그림 21 제안한 방법에 의한 회전 변형 예: 여성 노인

일반적으로, 회전 정도가 증가할수록 얼굴 형태의 왜곡이 심해지고 부자연스러워지는 것을 관찰할 수 있다. 이는 원론적으로 회전 각도가 커질수록 2차원 정면 얼굴에서 표현되지 않는 깊이 정보가 차지하는 비율이 점점 높아지는데, 표준 메쉬 집합을 토대로 복원한 측면과 반 측면의 참조 메쉬들이 원래의 입체정보를 정확히 반영하기 어렵기 때문에 발생한다. 특히, 회전 정도가 매우 클 때( $\delta \approx \pm 1.0$ ) 얼굴 내부에 있던 코의 능선과 같은 부분이 얼굴의 경계가 되면서 윤곽이 매우 단순하고 어색해지는 현상은, 얼굴 중앙의 수직 메쉬 라인 상에 메쉬점들을 증가시키거나 혹은 윤곽선을 스플라인 보간하는 방법에 의해 해결할 수 있다.

인물의 실제 얼굴을 어느 정도 사실적으로 변형할 수 있는지를 정량적으로 평가하기 위해서, 제안한 워핑 기법을 이용하여 변형한 영상과 이에 대응하는 실제 인물의 영상에서, 그림 22와 같이, 얼굴의 구성요소인 눈, 코, 입에 설정한 총 14개의 주요 특징점들 간의 정규화된 위치 오차를 측정한다.

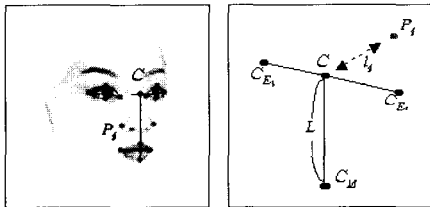


그림 22 얼굴의 주요 특징점과 측정 거리

먼저, 변형된 결과 영상에 대해 주요 특징점들을 다음과 같이 설정한다. 양 눈과 입의 좌우 끝점을 설정하고 양 끝점을 잇는 선분의 중간을 지나는 수직선이 윤곽선과 교차하는 상하 끝점을 설정한다. 여기에 코의 좌우 끝점을 포함시키면 총 14개의 특징점이 된다. 이때,  $C_E$ 과  $C_E$ 은 좌측과 우측 눈의 좌우 끝점을 잇는 선분의 중간 위치이고,  $C_M$ 은 입의 좌우 끝점을 잇는 선분의 중간 위치이다. 그리고  $C_M$ 을 지나는 수직선과  $C_E$ 과  $C_E$ 을 잇는 직선의 교차점을 기준 위치  $C$ 라 하고,  $C$ 에서  $C_M$ 까지의 거리를 기준 거리  $L$ 이라 하며,  $C$ 에서 각 특징점  $P_i$ 까지의 거리를  $l_i$ 라 하자. 한편, 이에 대응하는 실제 영상에서의 거리를 각각  $L'$ 과  $l'_i$ 이라 하자.

그러면 변형 영상과 실제 영상 사이에 대응하는 특징점들의 평균 위치 오차  $E_p$ 는 다음 식 (7)과 같이 정의할 수 있다.

$$E_p = \left( \sum_{i=1}^n \left| \frac{l_i}{L} - \frac{l'_i}{L'} \right| \right) / n, \quad n: \text{특징점의 수} \quad (7)$$

여기서,  $l_i$ 와  $l'_i$ 의 차를 직접 계산하지 않고 각각 기준거리  $L$ 과  $L'$ 으로 나누어 정규화된 값의 차를 계산하는 것은, 변형 영상과 실제 영상에서 얼굴의 크기가 일치하지 않는 경우에도 위치 오차를 제대로 계산하기 위함이다.

표 2는 총 10명으로부터 수평 회전 정도에 따라 변형한 영상을 200장 취득하여 실제 영상과 비교한 결과이다. 여기서도 회전 정도가 증가할수록 얼굴 형태의 왜곡이 심해지는 것을 확인할 수 있다. 이때, 전체 평균 위치 오차는 약 7%로 나타났다.

표 2 실제 및 변형 영상에 대한 주요 얼굴 특징점의 위치 오차

회전 정도 $\delta$	위치 오차 $E_p$	회전 정도 $\delta$	위치 오차 $E_p$
0.1	0.03	-0.1	0.05
0.2	0.04	-0.2	0.06
0.3	0.05	-0.3	0.06
0.4	0.05	-0.4	0.07
0.5	0.06	-0.5	0.07
0.6	0.06	-0.6	0.07
0.7	0.06	-0.7	0.08
0.8	0.07	-0.8	0.09
0.9	0.06	-0.9	0.08
1.0	0.09	-1.0	0.11
전체 평균 위치 오차		0.07	

## 6. 결론

본 논문에서는 3차원 기하학적 모델을 사용하지 않고 정면의 영상과 메쉬만으로 입체적인 회전에 따른 포즈 변형을 구현할 수 있는 새로운 영상기반 렌더링 기법을 구현하였다. 이를 위해, 메쉬의 중첩이나 역전을 허용하는 역전가능 메쉬위프 알고리즘을 제안하여 얼굴의 회전 변형과 눈과 입의 개폐 변형에 적용하였다.

제안한 방법의 타당성을 보이기 위하여 눈과 입의 개폐 효과, 표준 메쉬 집합에서의 반측면 메쉬의 역할, 인접/중첩/역전 메쉬에 따른 변형 성능, 그리고 다양한 연령과 성별에 대한 변형 성능 등에 대해 살펴보았다. 결론적으로, 다소의 왜곡현상이 있었지만 정면 영상 1장만을 가지고 회전에 따른 자연스러운 포즈 변형을 효과적으로 구현할 수 있음을 확인할 수 있었다. 또한, 제안한 방법은 모델 구축에 소요되는 비용과 시간을 획기적으로

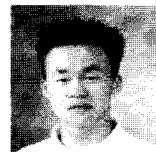
줄일 수 있고, 변형을 위한 처리속도도 크게 향상시킬 수 있다.

향후에는, 회전 정도가 커짐에 따라 심해지는 형태 왜곡을 완화시키기 위해 보다 정교한 메쉬 설정이나 보간에 대한 연구가 필요하고, 보다 다양한 회전 포즈 변형을 위한 연구가 필요하다. 또한, 치아나 혀로 구성되는 입안과 같이 정면 영상에서 명확하게 보이지 않는 부분의 처리를 위한 영상 합성(image composition)에 대한 연구도 필요하며, 눈과 입의 메쉬를 얼굴 메쉬와 통합 운영하면서 다양한 표정을 연출할 수 있는 방법에 대한 연구도 필요하다.

참 고 문 헌

[1] F. I. Parke and K. Waters, Computer Facial Animation, A. K. Peters, Ltd., 1996.  
 [2] P. Ekman and W. V. Friesen, Manual for the Facial Action Coding System, Consulting Psychologist Press, Inc., 1978.  
 [3] Y. Lee, D. Terzopoulos, and K. Waters, "Realistic Modeling for Facial Animation", Proc. of SIGGRAPH'95, pp.55-62, Aug. 1995.  
 [4] Z. Liu, Z. Zhang, C. Jacobs, and M. Cohen, "Rapid Modeling of Animated Faces from Video Images", Proc. of ACM Multimedia 2000, pp.475-476, 2000.  
 [5] F. Pighin, J. Auslander, D. Lischinski, and D. Salesin, "Realistic Facial Animation Using Image-Based 3D Morphing", Technical Report UW-CSE-97-01-03, 1997.  
 [6] S. B. Kang, "A Survey of Image-based Rendering Techniques", Videometrics VI(SPIE Int'l Symp. on Electronic Imaging: Science and Technology), Vol. 3641, pp.2-16, Jan. 1999.  
 [7] T. Vetter and T. Poggio, "Linear Object Classes and Image Synthesis from a Single Example Image", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp.733-742, Jul. 1997.  
 [8] L. McMillan and G. Bishop, "Plenoptic Modeling: An Image-Based Rendering System", Proc. of SIGGRAPH 95, pp.39-46, Aug. 1995.  
 [9] R. Szelicki and H. -Y. Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps", Proc. of SIGGRAPH'97, pp.251-258, Aug. 1997.  
 [10] M. Levoy and P. Harahan, "Light Field Rendering", Proc. of SIGGRAPH'96, pp.31-42, Aug. 1996.  
 [11] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The Lumigraph", Proc. of SIGGRAPH'96,

pp.43-54, Aug. 1996.  
 [12] S. M. Seitz and C. R. Dyer, "View Morphing", Proc. of SIGGRAPH'96, pp.21-30, Aug. 1996.  
 [13] S. Avidan and A. Shashua, "Novel View Synthesis in Tensor Space", Proc. of Int'l Conf. on Computer Vision and Pattern Recognition, pp.1034-1040, Jun. 1997.  
 [14] P. E. Debevec, C. J. Taylor, and J. Malik, "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach", Proc. of SIGGRAPH'96, pp.11-20, Aug. 1996.  
 [15] Jonas Gomes, Lucia Darsa, Bruno Costa, and Luiz Velho, Waring and Morphing of Graphical Objects, Morgan Kaufmann Pub., 1998.  
 [16] Randy Crane, Simplified Approach to Image Processing, Prentice-Hall PTR, 1997.



오 승 택  
 2000년 공주대학교 전자계산학과(이학사). 2002년 공주대학교 대학원 전자계산학과(이학석사). 2002년~현재 (주)라스텍 연구원. 관심분야는 컴퓨터비전, 가상현실



김 병 환  
 1989년 연세대학교 전자공학과(공학사). 1991년 연세대학교 대학원 전자공학과(공학석사). 1996년 연세대학교 대학원 전자공학과(공학박사). 1991년~1992년 (주)YC&C(현, 주SKC&C) 연구원. 2000년~2001년 (주)모리아테크놀로지 연구소장. 1997년~현재 공주대학교 정보통신공학부 컴퓨터전공 조교수. 관심분야는 컴퓨터비전, 가상현실