

# 메쉬 워핑에 의한 정면 영상으로부터의 3D 얼굴 모델링

김중식<sup>†</sup>, 김진모<sup>\*\*</sup>, 조형제<sup>\*\*\*</sup>

## 요 약

최근 컴퓨터 비전, 컴퓨터 그래픽스 분야의 급속한 발전과 하드웨어의 우수한 성능으로 인해 3차원 모델링 기술 또한 급속히 발전 되었다. 그리고 다양한 3D 콘텐츠들의 등장으로 3차원 모델링 기술의 수요가 더욱 많아지고 그 질 또한 높아지고 있다. 3D 얼굴 모델은 이러한 콘텐츠들에 폭 넓게 적용될 수 있어 활용성이 높다. 본 논문에서는 주어진 한 장의 2D 정면 얼굴 영상으로부터 3D 얼굴을 모델링하는 방법을 제안한다. 이를 위해 다수의 정점으로 구성된 3D 일반 얼굴 모델을 입력 얼굴 영상과의 대응관계를 통해 변형하여 3D 얼굴을 모델링한다. 우선 AAM을 이용하여 입력된 얼굴 영상에서 특징점을 추출한다. 그리고 추출된 특징점만으로 직관적인 모델을 변형하기 위해 메쉬 워핑을 적용한다. 여기에 얼굴의 깊이 정보를 유도하기 위하여 한 장의 정면 얼굴 영상을 활용한 명도 변화 기반의 깊이 값의 추출을 시도한다. 이러한 과정을 통해 경제적으로 부담되지 않으며 사용 환경에 대한 제한이 적어 보편적으로 활용될 수 있는 3D 얼굴 모델링 방법을 구현하고 그 타당성을 실험으로 보인다.

## 3D Face Modeling from a Frontal Face Image by Mesh-Warping

Jung-Sik Kim<sup>†</sup>, Jin-Mo Kim<sup>\*\*</sup>, Hyung-Je Cho<sup>\*\*\*</sup>

## ABSTRACT

Recently the 3D modeling techniques were developed rapidly due to rapid development of computer vision, computer graphics with the excellent performance of hardware. With the advent of a variety of 3D contents, 3D modeling technology becomes more in demand and it's quality is increased. 3D face models can be applied widely to such contents with high usability. In this paper, a 3D face modeling is attempted from a given single 2D frontal face image. To achieve the goal, we thereafter the feature points using AAM are extracted from the input frontal face image. With the extracted feature points we deform the 3D general model by 2-pass mesh warping, and also the depth extraction based on intensity values is attempted to. Throughout those processes, a universal 3D face modeling method with less expense and less restrictions to application environment was implemented and it's validity was shown through experiments.

**Key words:** Face Modeling(얼굴 모델링), Extract Feature Point(특징점 추출), Mesh Warping(메쉬 워핑), Brightness based deforming depth value(명도 기반 깊이 변형)

## 1. 서 론

3차원 모델링 기술은 컴퓨터 그래픽스와 컴퓨터

비전 분야에서 오랫동안 많은 연구가 이루어져 왔고, 최근 공학 기술의 급속한 발전과 하드웨어의 우수한 성능으로 인해 빠르게 발전되고 있다[1]. 3D 얼굴 모

※ 교신저자(Corresponding Author): 조형제, 주소: 서울특별시 중구 필동 3가 26번지 동국대학교 원홍관 E431호(100-715), 전화: 02) 2260-3344, E-mail: chohj@dongguk.edu

접수일: 2012년 8월 16일, 수정일: 2012년 10월 16일

완료일: 2012년 11월 8일

<sup>†</sup> 준회원, 동국대학교 멀티미디어학과  
(E-mail: musicpeople@dongguk.edu)

<sup>\*\*</sup> 정회원, 동국대학교 영상문화콘텐츠연구원  
(E-mail: mythsjin@dongguk.edu)

<sup>\*\*\*</sup> 정회원, 동국대학교 멀티미디어학과

델은 3차원 공간에서 얼굴 자세를 변형하고, 적절한 조명 효과를 적용하는 등 시뮬레이션을 할 수 있으므로 활용성이 높다. 또한 얼굴 자세, 표정 등을 자세히 표현할 수 있고 사실감이 뛰어나 애니메이션, 게임, 방송 등의 엔터테인먼트 분야에서 가상현실 및 증강현실, 얼굴 인식 분야에 이르기까지 다양한 분야에서 활발히 연구되어 왔으며, 현재는 기술 발전에 따라 더욱 섬세하며 개인의 얼굴 특징을 잘 표현할 수 있는 모델링 방법이 많이 연구되고 있다[2-11].

기존의 3D 얼굴 모델링에 관한 연구들 중 3D 스캐너를 기반으로 모델링하는 방법[2]은 매우 정교하고 사실적인 모델을 얻을 수 있지만 보편적으로 사용하기에는 경제적인 부담이 크다. 일반 모델을 기반으로 모델링하는 방법[3-6]은 빠르게 모델을 구성할 수 있다는 장점이 있지만 정점(vertex) 수가 적기 때문에 정교한 모델을 구성하기 어렵다는 단점이 있다. 학습 모델을 기반으로 모델링하는 방법[7-11]은 비교적 정교한 모델링이 가능하지만 통계적인 모델을 구축하는데 있어서 3D 얼굴 데이터 수집이 어렵고, 모델을 구축하는데 있어서 수반되는 작업량이 많아 3D 얼굴을 복원하는데 많은 연산량을 필요로 하기 때문에 처리 시간도 길어지는 단점이 있다.

본 논문에서는 이러한 단점들을 대체 및 개선하여 직관적으로 활용하기 편하며 경제적 부담 또한 줄일 수 있는 3D 일반 모델 기반의 메쉬 워핑에 의한 3D 얼굴 모델링 방법을 제안한다. 모델 구축 단계에서는 통계적 모델 구축을 위한 데이터 수집의 어려움과 학습과정의 복잡성 등의 단점을 보완하기 위해 3D 일반 얼굴 모델로부터 정점(vertex) 수를 크게 세분화하여 활용한다. 그리고 입력 얼굴 영상과 3D 일반 얼굴과의 메쉬 워핑을 통한 3D 얼굴 모델링을 위한 첫 번째 단계로 입력 영상에서의 특징점을 추출한다. 이때 특징점 추출 방법은 조명에 영향을 많이 받는 색 기반 추출 방법이나 ACM(active contour model)[12], ASM(active shape model)[13]에 비해 비교적 안정적인 AAM(active appearance model)[14]을 활용한다. 다음은 추출된 특징점을 통해 3D 일반 모델을 변형하는 방법으로 많은 대응점들을 처리하기 때문에 발생하는 시간 지연과 RBF(radial basis function)의 기저함수와 적용 범위를 지정해야 하는 단점을 보완하기 위해 최소한의 대응점을 이용하여 두 단계 메쉬 워핑을 적용하여 변형을 시도한다. 또한

메쉬 워핑시 3D 일반 얼굴 모델의 UV 맵도 함께 워핑함으로써 별도의 텍스처 맵을 구성하지 않고 자연스럽게 텍스처를 입힌다. 마지막으로 3D 얼굴 모델의 깊이 정보를 계산하기 위하여 shape-from-shading에 착안하여 조명의 명도 변화에 따른 선형적 영역별 깊이 조절을 시도한다.

본 논문의 구성은 다음과 같다. 2장에서는 3D 얼굴 모델링과 관련된 기존의 연구들을 설명하고, 3장에서는 제안하는 3D 얼굴 모델링의 처리과정과 각 과정에 대한 배경이론, 방법에 대해 설명한다. 4장에서는 실제 구현된 프로그램을 통한 실험 결과에 대해 설명하고 마지막으로 5장에서 결론을 맺는다.

## 2. 관련연구

컴퓨터 비전과 그래픽스 분야에서 사람의 얼굴이라는 특정한 객체를 모델링하는 3D 얼굴 모델링은 비용적인 문제와 활용성을 개선할 수 있는 수동적 모델링 방법[15]이 많이 사용된다. 3D 얼굴 모델링을 하기 위해 필요한 것은 정점의 기하학적인 정보와, 텍스처 이미지 그리고 그에 대한 UV 정보이다. 따라서 개인화된 3D 얼굴 모델링을 하기 위해서 먼저 기하학적 정보를 가진 변형시킬 얼굴 모델과 텍스처 데이터가 갖추어져야 한다.

Crocombe와 Linney등은 3D 스캐너를 이용한 3D 얼굴 모델링 방법을 제안하였다[2]. 이 방법은 매우 정교한 모델링이 가능하지만 높은 비용과 복잡한 사용 방법으로 일반인이 사용하기에는 부담이 따른다.

3D 일반 모델 기반 3D 얼굴 모델링 방법은 표준이 되는 하나의 일반 모델을 변형하여 모델링하는 방법이다. 본 논문에서 사용한 Candide 모델[3] 또는 Warter's 모델[4]이 일반적으로 사용되며 직접 디자인한 모델[5,6]을 사용할 수도 있다. 2D 이미지들에서 얼굴의 특징점을 추출하고 shape-from-X에 기반한 알고리즘들을 사용하여 3D 좌표를 추출한다. 추출된 3D 좌표 값에 맞도록 RBF 보간을 통해 일반 모델을 변형시켜 형상 모델을 얻는다. 2D 이미지들로부터 얻은 텍스처를 이 형상 모델에 맵핑하여 최종적인 3D 얼굴 모델을 완성한다. 3D 일반 모델을 사용하는 방법은 비교적 처리 과정이 간단하고 처리 시간이 빠르다는 장점이 있지만 정점과 폴리곤 수가 적어 정교하고 세밀한 모델링이 어렵다는 단점이 있다.

Blanz와 그의 동료들은 3D 변형가능 모델 기반 3D 얼굴 모델링 방법을 제안하였다[7]. 이 방법은 얼굴 인식을 위해 개발된 방법으로 3D 스캐너를 통해 구축한 3D 얼굴 데이터베이스를 주성분 분석(PCA, principal component analysis)을 사용하여 3D 변형가능 얼굴 모델을 구축한다. 구축된 얼굴 모델을 입력된 2D 얼굴 영상에 투영했을 때 정합이 잘 되도록 변형가능한 모델을 조정하여 개인화된 3D 얼굴 모델을 얻는다. 매우 정교하고 사실적인 방법이지만, 3D 변형가능 모델을 구축할 때의 데이터 수집이 어렵고, 처리해야할 데이터도 많아 시간이 많이 걸리는 단점이 있다.

3D 변형가능 형상 모델 기반 3D 얼굴 모델링 방법 [11]은 변형가능 모델과 같은 방법으로 형상 모델을 구축하고 입력 이미지들로부터 얼굴의 특징점을 추출한다. 추출된 특징점을 바탕으로 SfM(structure-from-motion) [16]등의 방법을 통해 3D 정점 좌표 값을 계산한다. 구축된 3D 변형가능 형상 모델이 계산된 3D 정점들에 잘 정합되도록 형상 모델을 조정하여 개인화된 3D 형상 모델을 얻는다. 또한, 입력받은 2D 얼굴 이미지들로부터 텍스처를 구하고, 이를 3D 형상 모델에 맵핑함으로써, 최종적인 3D 얼굴 모델을 완성한다. 이 방법은 3D 형상 모델만을 사용함으로써 처리 시간을 많이 단축할 수 있는 장점이 있다. 하지만 이 방법 역시 변형가능 형상 모델을 구축하는데 데이터 수집의 어려움이 있고, 3D 정점 계산과정이 복잡하다는 단점이 있다.

모델링 방법마다 세부 알고리즘은 다르지만 일반적으로 다음과 같은 처리과정을 거친다. 첫 번째로 입력 영상에서 얼굴의 특징점을 추출하고, 두 번째로 3D 정점의 좌표 값을 계산한다. 세 번째로 계산된 좌표 값으로 얼굴 모델을 변형하고, 네 번째로 텍스처 맵을 구성하고 텍스처 맵핑을 하게 된다.

입력된 얼굴 영상의 특징점을 추출하는 방법으로는 색 기반 방법, ACM[12], ASM[13], AAM[14]등이 있다. 색 기반의 얼굴 특징 추출 방법은 조명 환경에 큰 영향을 받고, ACM이나 ASM 또한 조명과 자세에 영향을 많이 받고 학습 환경과 입력 환경의 차이가 심하면 추출 성능이 떨어지는 단점이 있다. 반면 본 논문에서 사용한 AAM은 텍스처 모델과 형상 모델을 모두 참고하기 때문에 좀 더 나은 추출 성능을 기대할 수 있다. 깊이 값을 구하는 방법으로는 두 장

의 이미지를 사용하는 방법[6], 일련의 이미지를 이용하여 SfM을 사용한 방법[11]등이 있다. 이 방법들은 여러 장의 입력 영상이 필요하기 때문에 환경에 자유롭지 못하고 입력 영상간의 대응점을 지정해 주어야 하는 단점이 있다. 반면 본 논문에서는 한 장의 입력 사진을 사용하여 명도 변화에 기반한 변형 방법을 사용하기 때문에 입력환경에 좀 더 자유롭고 처리 과정도 줄어든다. 모델 변형 방법으로는 RBF 보간을 주로 사용한다. 이 방법은 기저함수와 적용 영역을 별도로 지정해 주어야한다는 단점이 있다. 하지만 본 논문에서는 두 단계 메쉬 워핑 방법을 사용함으로써 추출된 특징점 기반으로 직관적인 메쉬를 자동적으로 구성할 수 있고, 영역기반의 변형으로 주위 정점들이 좀 더 자연스럽게 변형될 수 있다. 텍스처 맵핑을 하기 위해 입력 영상을 기반으로 텍스처 맵을 구성하게 되는데 이 과정에서 소요되는 시간이 얼굴 모델링에서 큰 비중을 차지 한다. 이와는 달리 본 논문에서는 이러한 처리과정을 얼굴 변형 시에 일반 모델의 UV좌표도 두 단계 메쉬 워핑을 통해 변형함으로써 별도의 텍스처 맵을 구성할 필요가 없다.

### 3. 3D 얼굴 모델링 방법

본 논문에서 제안하는 3D 얼굴 모델링 방법은 다음과 같다. 정교한 3D 일반 얼굴 모델을 바탕으로 입력 얼굴 영상의 분석과 대응관계 유도를 통해 3D 얼굴을 모델링한다. 이때 얼굴 영상 분석은 AAM 기반에서 얼굴의 특징점을 추출하고 3D 일반 얼굴 모델과의 대응관계를 위해 메쉬 워핑과 명도 변화에 기반한 깊이 값 추출후 텍스처 맵핑을 수행한다. 그림 1은 제안하는 3D 얼굴 모델링 방법의 흐름을 나타낸 것이다.

#### 3.1 Candide-3 Face Model & MPEG-4 FDP

본 논문에서는 3D 일반 얼굴 모델로 Candide-3 모델[3]을 사용한다. Candide-3 얼굴 모델은 MPEG-4에서 정의한 얼굴 특징점을 기반으로 모델링되었기 때문에 비교적 적은 수의 정점으로 얼굴 모델링이 가능하지만 적은 수의 정점으로 인하여 정교하고 세밀한 얼굴 모델링은 불가능하다. 그래서 본 논문에서는 Candide-3 얼굴 모델을 기반으로 정점들을 세분화하여 2,304개의 정점과 4,416개의 삼각형 폴리곤

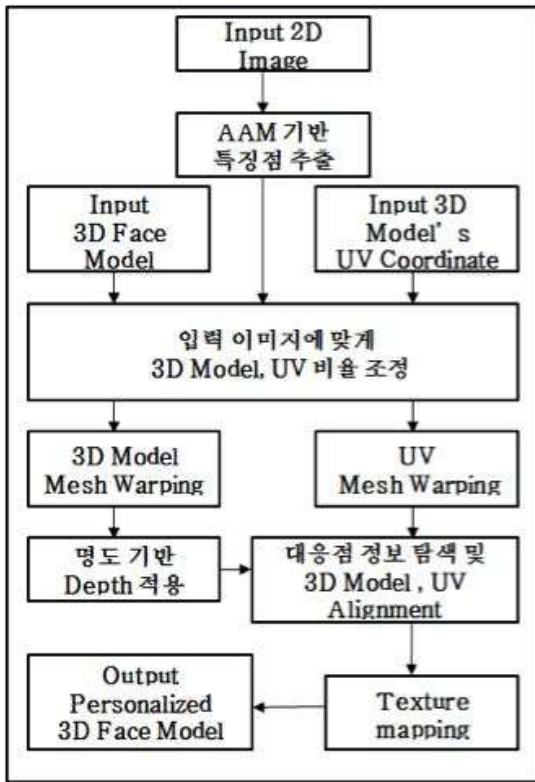


그림 1. 제안하는 3D 얼굴 모델링의 시스템 구성

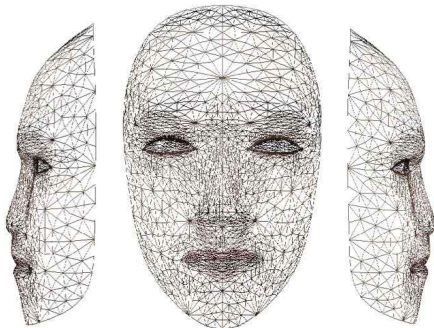


그림 2. Candide-3 모델을 세분화한 3D 얼굴 모델

(triangle polygon)으로 구성된 얼굴 모델을 사용한다. 그림 2는 본 논문에서 사용된 3D 얼굴 모델을

보인 것이다.

### 3.2 AAM 기반 얼굴 특징점 추출

다음은 입력 얼굴 영상에서 얼굴의 특징 정보를 추출하는 과정으로 본 논문에서는 AAM을 활용한다. AAM은 영상의 특징을 추출하기 위한 통계적인 모델링 방법으로 객체의 형상과 텍스처 정보인 통계적 모델을 입력된 영상에 검출하는 알고리즘이다. AAM은 [14]에서 처음 제안되었고, 이후 AAM 개선 및 확장에 대해 많은 연구가 이루어졌다. AAM은 모델링 단계와 검출(fitting) 단계로 이루어진다. 본 논문에서의 AAM 기반 얼굴 특징점 추출은 Cootes 등이 제안한 방법 [14]의 내용을 기반으로 구현하였다.

#### 3.2.1 Shape Model

형상 모델(shape model)은 학습 영상에 대해 사용자가 직접 특징점을 지정하여 구축하는 모델이다. 얼굴 형상 모델  $S$ 는 평균 형상  $\bar{X}$  와 고유 값들 중 큰 값에 대응되는  $n$  개의 형상 고유벡터  $\phi_i (i=1, \dots, n)$  를 선형 결합하여 식(1)와 같이 나타낼 수 있다.

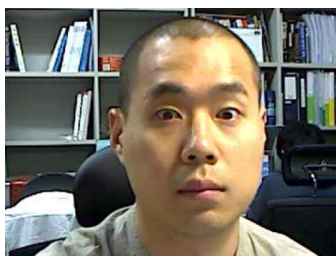
$$S = \bar{X} + \sum_{i=1}^n \phi_i b_i = \bar{X} + P_s b_s \tag{1}$$

$$(P_s = [\phi_1, \dots, \phi_n], b_s = [b_1, \dots, b_n]^T)$$

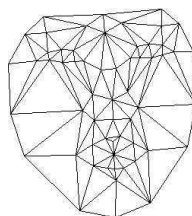
#### 3.2.2 Texture Model

텍스처 모델(texture model)은 추출하고자 하는 특징 영역에 대한 영상에서의 픽셀 명도 값을 말한다. 그림 3과 같이 평균형상 내부 공간의 화소 집합을 이용하여 모델링하는 과정은 형상 차이에서 발생하는 텍스처 변화율을 줄여준다.

정규화 과정을 거친 텍스처 데이터에 대하여 형상 모델과 동일한 PCA 과정을 통해 선형적인 모델을



(a) 입력 이미지



(b) 평균 형상



(c) 형상 정규화 이미지

그림 3. 학습 데이터들의 정규화

식(2)와 같이 구한다.

$$g = \bar{g} + \sum_{i=1}^k T_i c_i = \bar{g} + P_g b_g \quad (2)$$

$$(P_g = [T_1, \dots, T_k], b_g = (c_1, \dots, c_k)^T)$$

정규화 된 평균 그레이 레벨 벡터  $\bar{g}$ 와  $k$ 개의 텍스처 고유벡터  $T$ , 텍스처 파라미터  $c$ 의 선형 결합으로 텍스처 모델  $g$ 를 구할 수 있다.

### 3.2.3 Combined Model

형상 파라미터  $b_s$ 와 텍스처 파라미터  $b_g$ 를 조절하면 다양한 얼굴의 모양과 생김새를 표현할 수 있다. 이 두 파라미터를 하나의 열벡터로 결합하여 형상과 텍스처 파라미터가 합쳐진 형태의 외모 파라미터  $c$ 를 조절하면 형상 모델  $S'$ 와 텍스처 모델  $g'$ 를 동시에 제어할 수 있다.

$$S' = \bar{S} + Q_s c \quad (3)$$

$$g' = \bar{g} + Q_g c \quad (4)$$

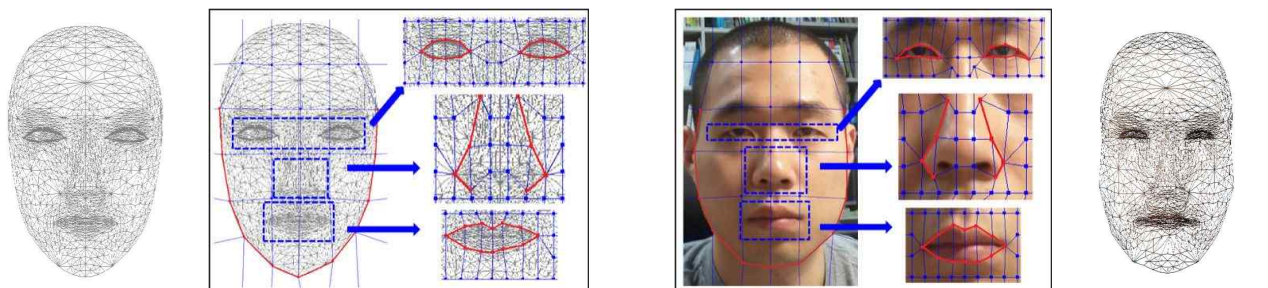
### 3.2.4 AAM 검출

검출(Fitting)은 학습된 형상 모델과 텍스처 모델에 기반하여 입력된 영상에서의 특징을 검출하는 과정이다. 식 (3)과 (4)로 모델링된 얼굴 이미지의 텍스처와 입력된 검출 대상 이미지를 평균 형상  $\bar{X}$ 로 워핑한 텍스처와의 에러의 제곱을 최소화하도록 하는 외모 파라미터  $c$ 를 찾고, 이때의  $c$ 를 가지고 식 (3)으로 표현되는 형상 모델을 찾고자하는 얼굴 특징점들로 판정한다. 모델과 입력 데이터의 에러 제곱을 최소화하는 외모 파라미터  $c$ 를 찾는 방법은 최적화의 문제이다.

### 3.3 메쉬 워핑을 통한 3D 얼굴 및 UV 변형

AAM을 통해 얼굴의 특징 정보가 추출되면 이를 통해 3D 일반 얼굴 모델링과 UV 정보를 추출된 특징점을 바탕으로 변형한다. 이를 위해 대응 관계가 명확하고 처리 시간을 단축시킬 수 있는 두 단계 메쉬 워핑을 사용한다[17]. 메쉬 워핑은 소스영상과 목적영상에 서로 대응되는 제어점을 지정하여 메쉬를 구성하고 소스영상의 화소 값이 목적영상의 메쉬 영역에 맞게 맵핑하는 것이다.

두 단계 메쉬 워핑을 하기 위해서는 소스영상의 제어점과 목적영상의 제어점을 기반으로 소스 메쉬와 목적 메쉬를 구성한다. 목적 메쉬는 입력 영상에서 AAM을 통해 추출된 얼굴 특징점들로 구성하고, 소스 메쉬는 특징점에 해당하는 3D 일반 모델 정점의(x,y) 좌표로 구성된다. 자연스러운 워핑을 위해 얼굴형태와 얼굴 구성요소(눈, 코, 입)를 그림 4와 같이 독립적으로 워핑한다. 얼굴형 워핑 시 메쉬의 구성은 5×5 메쉬로 얼굴 외곽의 특징점들을 포함하도록 구성하고, 얼굴 구성요소 워핑 시에는 눈, 코, 입의 특징 점을 포함하는 17×17의 메쉬로 구성한다. 메쉬가 구성되면 워핑은 이 메쉬를 기준으로 두 단계의 처리 과정을 거쳐 수행된다. 첫 번째 단계는 수직 스플라인에 대한 처리로서 소스 메쉬 제어점의 x좌표가 목적 메쉬의 x좌표로 이동한다. 제어점들의 x 좌표 값들에 대해 보간 함수를 적용시키고 Fant의 리샘플링 방법을 사용하여 재추출 함으로써 새로운 값들로 행 내의 각 화소들이 보간된다. 보간법은 간단한 선형 보간법에서부터 복잡한 스플라인 함수까지 다양하다. 두 번째 단계는 수평 스플라인에 대한 처리로서, x좌표에 대해 이루어졌던 처리가 y좌표에 대해 반복되며 출력영상의 각 열은 중간 영상으로부터 재 추출 된다. 각 단계에서 변형되는 해당 위치를 계



(a) 일반 모델 (b) 일반 모델의 특징점 기반 소스 메쉬 구성 (c) 입력 영상의 특징점 기반 목적 메쉬 구성 (d) 변형된 모델

그림 4. 3D 일반모델과 입력영상의 특징점 기반 메쉬 구성



산 할 때, 3D 모델 정점 값들의 위치이면 변형되는 대응 좌표를 버퍼에 저장하게 되고 그 좌표로 3D 얼굴 모델의 정점 값들을 정렬하게 된다. 결과적으로, AAM 으로 추출된 소수의 특징 점 좌표에 기반하여 총 2304개의 3D 모델 정점을 부드럽게 변형시키게 된다. 그림 4는 3D 일반 모델과 입력 영상의 특징점으로 메쉬를 구성하는 것을 개념적으로 나타낸 것이다.

### 3.4 명도 변화 기반의 깊이 값 추출

마지막으로 3D 얼굴의 깊이 정보 유도를 위해 한 장의 정면 이미지 내에서의 명도 값 변화를 통한 깊이 값(depth value)을 추출한다. 이 방법은 고정된 조명 조건하에 수행되어야 하는 단점이 있지만 일련의 이미지를 입력받아야 하고 복잡한 처리 과정을 거치는 shape-from-X 알고리즘들[15] 보다 간단하고 처리 속도가 빠르다는 장점이 있다.

본 논문에서는 [18]의 조명 추정과 그림자 생성 기법을 토대로 수작업으로 조명 환경을 결정한다. 사진을 촬영할 때 조명 위치를 정해진 곳에 배치하고 가상 조명을 이를 기반으로 유사한 위치에 배치하도록 한다. 조명은 그림자로 깊이 값의 크기를 구분하기 쉽도록 조명을 얼굴 보다 위쪽에 배치한다. 그림 5와 같이 렌더링한 2D 영상의 각 샘플 포인트에서의 명도 분포가 입력 영상의 샘플 포인트에서의 얼굴 명도 분포와 유사한 값을 갖도록 사용자가 3D 공간상의 조명을 조절한다. 이는 깊이 값을 조절하기 전에 인접한 샘플 포인트의 밝기 차이가 보편적인 인간의 얼굴 영역에서 들어간 부분과 나온 부분의 명도 값 차이를 나타낼 수 있도록 그 값을 임계 값을 두어 제한함으로써 깊이 값을 변형할 때 오류확률을 적게 하는 작업을 하는 것이다.

두 영상의 얼굴 영역에서 유사한 밝기 분포를 갖

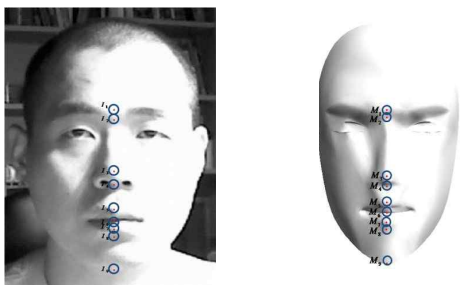


그림 5. 입력 영상과 얼굴 모델의 샘플 포인트

게 되면 미리 지정한 특징점들의 상대적 밝기를 비교하여 얼굴 모델의 특징 정점들의 깊이 값을 조절한다. 그림 5는 입력영상을 그레이 스케일 변환한 영상과 Phong 셰이딩으로 렌더링한 얼굴 모델을 나타낸다.

렌더링된 얼굴 모델 영상에서 조명의 영향을 많이 받는 영역(코, 이마, 입, 턱 등)에 대한 각 샘플 포인트들의 명도 값( $M_{n(n=1...9)}$ )은 주위 일정영역의 평균 명도 값을 취하고, 마찬가지로 입력 영상에서 해당 영역에 대한 평균 명도 값을  $I_{n(n=1...9)}$ 이라 할 때, 식 (5)를 적용하여 각 샘플 포인트에 해당하는 정점의 깊이 값  $v_n(z)$ 을 구한다.

$$v_n(z) = |A_n B_{n+1} - A_{n+1} B_n| / A_{n+1} \quad (5)$$

$$(A_n = |I_n - I_{n+1}|, B_n = |M_n - M_{n+1}|)$$

$A$ 는 이미지내의 인접 샘플 포인트간의 밝기 값의 차이,  $B$ 는 렌더링된 모델의 인접 샘플 포인트간의 밝기 값의 차이를 나타낸다. 해당 정점의 깊이 값이 구해지면 해당 정점과 연결된 인접 정점들을 추출하여 인접 샘플 포인트까지의 총 step인 totalstep 으로 각 정점들을 구분한다. 각 step에 식(6)을 통해 구한 가중치  $\alpha$ 로 정점들의 깊이 값을 조정한다.

$$\alpha = v_{first}(z) + v_{end}(z) / totalstep \quad (6)$$

## 4 실험 및 분석

제안한 일반 모델 기반의 3D 얼굴 모델링을 4단계로 나누어 실험을 진행하였다. 첫째로, AAM 학습 및 얼굴 특징 추출하였고 둘째로, 3D 모델을 메쉬 워핑으로 3D 일반 얼굴 모델을 변형하는 실험을 하였다. 셋째로, 명도 기반 깊이 값 추정 방법으로 3D 일반 얼굴 모델의 깊이 값을 변형하고 변형된 좌표의 정확성을 검증하였다. 넷째로, 텍스처 맵핑을 통해 3D 얼굴 모델을 완성시켰다.

본 연구는 고정된 조명 환경에서 이루어졌고, 실험 PC 사양은 Intel Core2 Duo CPU E8500, DDR2 3GB memory, Geforce GT240 VGA, @info PC camera를 이용하였다. 또한 Visual C++ 2010, MFC, OpenCV, DirectX9.0 라이브러리를 이용하여 구현하였다.

### 4.1 AAM 모델링 & 검출

AAM 모델링에서 얼굴 특징점의 개수를 그림 6과

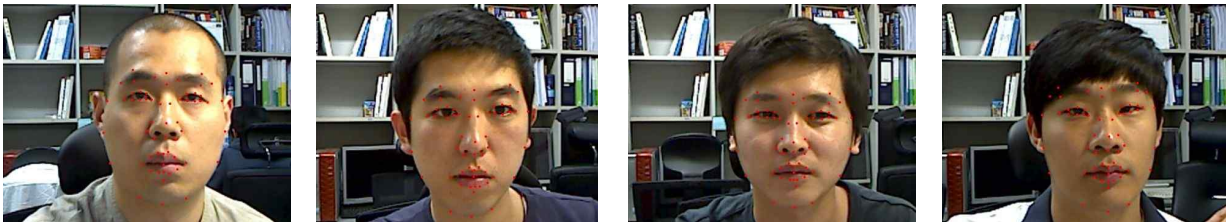


그림 6. AAM 학습 영상들과 특징점의 위치

같이 총 59개로 지정하였다. 학습 이미지들은 고정된 환경에서 캠으로 취득한 다섯 사람에 대한 자세, 조명, 표정이 다른 150장의 영상을 사용하였다. 각 데이터들의 해상도는 640×480이며, JPEG 포맷으로 되어 있다.

그림 7은 AAM 검출 알고리즘으로 검출된 결과를 보여준다. 대부분의 특징점이 정확하게 검출된 것을 볼 수 있다.



그림 7. AAM으로 검출된 특징점

#### 4.2 두 단계 메쉬 워핑을 이용한 3D 얼굴 모델 변형

AAM으로 추출된 얼굴 특징점으로 메쉬 워핑을 통해 3D 얼굴 모델을 변형시키려면 먼저 입력 얼굴 이미지의 해상도와 3D 얼굴 모델의 비율을 상대적으로 맞추어 주어야 한다. 목적 메쉬로 입력 영상에 대해 AAM을 사용하여 추출한 특징점으로 구성된 메

쉬가 사용되고, 소스 메쉬로는 추출된 특징점에 대응하는 3D 얼굴 모델의 정점들로 구성된 메쉬를 사용한다. 그림 8은 메쉬 워핑을 통해 변형된 얼굴 모델, UV 맵을 보여준다.

#### 4.3 명도 변화 기반의 깊이 값 변형

입력 이미지를 그레이 스케일 변환한 이미지에서의 샘플 포인트에 대한 명도 값과, 일반 모델에 Phong 셰이딩을 적용하여 계산된 샘플 포인트에 해당하는 정점의 명도 값을 이용하여 일반 모델의 깊이 값을 변형한다. 그림 9는 렌더링된 3D 일반 모델의 각 샘플 포인트에 대해 식(5)를 적용하여 깊이 값을 조절한 영상이다.

그림 10은 각 실험 영상에 대한 실제 측면 사진을 취득하여 가장 높은 영역인 코 끝의 높이에 대한 각 샘플 포인트들의 위치를 상대적인 비율로 나타내어 변형된 모델과 비교한 것이다.

대부분의 샘플 포인트에서 실제 측면 사진과 유사한 깊이 값으로 변형된 것을 알 수 있다. 얼굴 영역 하단부분의 샘플 포인트들이 상대적으로 오차가 큰 이유는 실험 환경에서 조명이 책상에 반사되어 얼굴 영역 하단의 밝기가 3차원 공간상의 밝기와 유사하

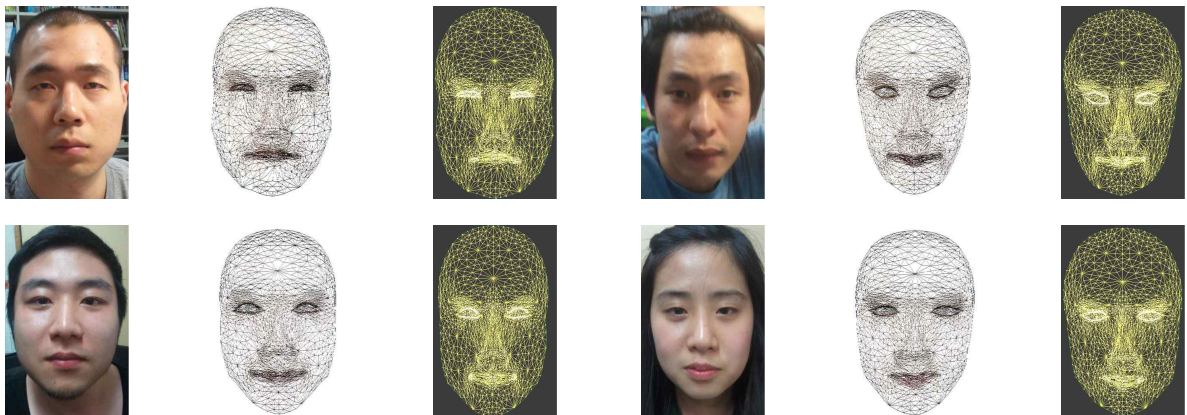
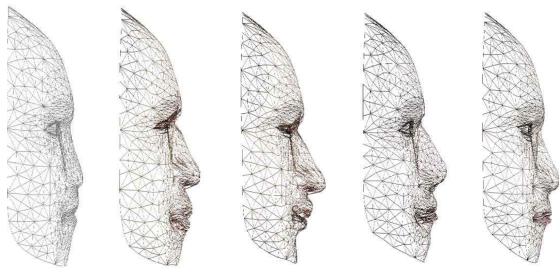


그림 8. 메쉬 워핑된 3D 모델과 UV 좌표



일반 모델 실험 1 실험 2 실험 3 실험 4

그림 9. 각 실험 영상에 대한 깊이 값 변형

게 맞춰지지 않았기 때문이다.

#### 4.4 텍스처 맵핑

3D 일반 모델의 변형을 한 후 메쉬 위핑으로 위핑

된 UV 좌표로 텍스처 맵핑을 수행한다. 별도의 텍스처 맵을 구성하지 않고 DirectX에서 제공하는 텍스처 맵핑 함수를 이용한다. 그림 11은 완성된 3D 얼굴 모델링 결과를 보여준다.

#### 4.5 비교 분석

두 장의 영상을 이용한 스마트폰에서의 3차원 얼굴 모델링[6]에서는 특징에 기반한 RBF 보간 함수로 3D 일반 모델을 변형하다보니 그림 12에서 보이는 것과 같이 턱과 같은 부분에서는 변형이 자연스럽게 못하다. 반면 본 논문에서는 메쉬 위핑을 사용하여 메쉬 영역을 기반으로 위핑이 이루어지기 때문에 RBF 보간과 같이 적용 범위를 별도로 지정해 줄 필

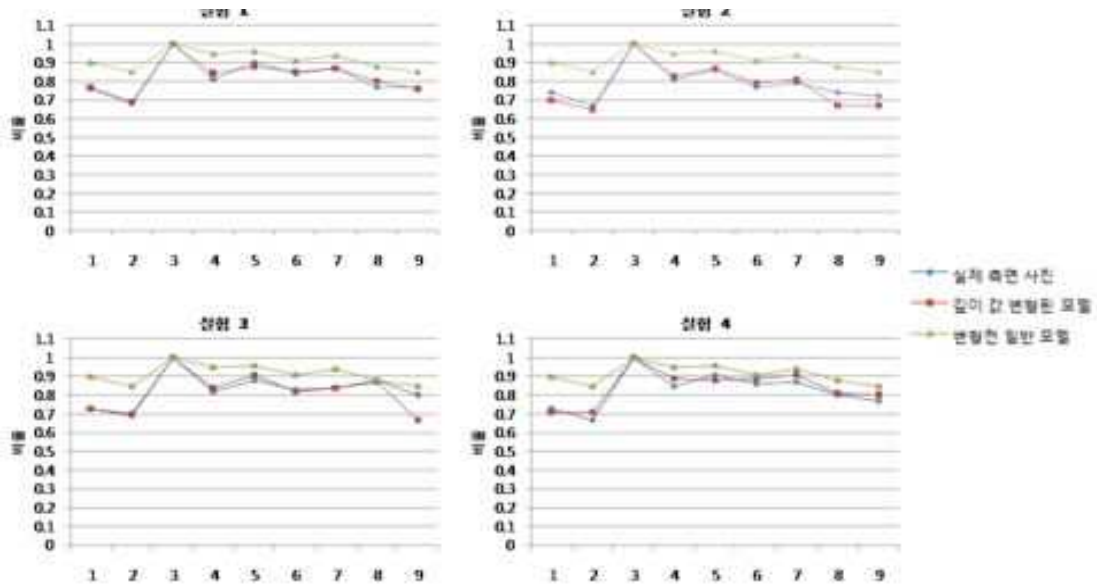


그림 10. 실제 측면사진과 변형된 모델의 깊이 값 비교

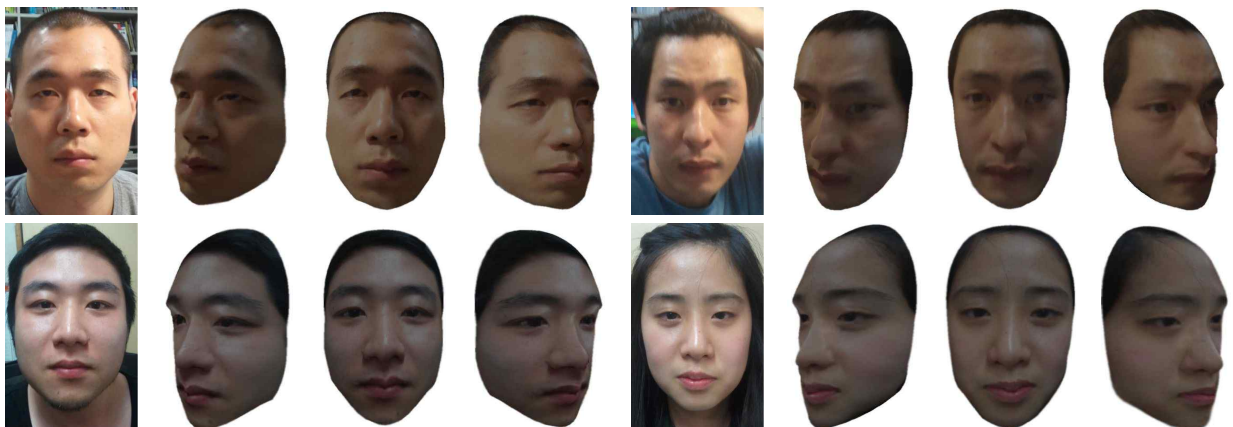
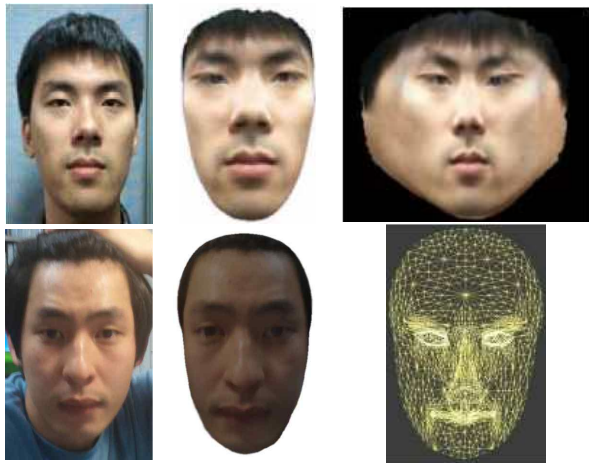


그림 11. 텍스처 맵핑으로 완성된 3D 얼굴 모델





그림 12. 제안한 방법과 일반 모델 변형 방법과의 비교



(a) 입력 영상 (b) 정면 영상 (c) 텍스처 맵

그림 13. 변형가능 형상 모델과의 비교

요가 없고 좀 더 자연스러운 변형이 가능하다. 또한 기존 연구는 정면과 측면, 두 장의 사진을 사용하지만 본 논문에서는 한 장의 사진으로 충분하다. 메쉬 워핑을 사용하여 얼굴을 변형시키는 시간은 30초 정도이지만 이는 워핑 과정의 반복

처리를 최적화 할 경우 스마트폰에서 활용하기에도 유용할 것이다.

3D 변형가능 형상 모델 기반 3D 얼굴 모델링[11]에서는 얼굴 형상을 구성하기 위한 데이터의 수가 많이 필요하고 텍스처 맵을 구성하는데 시간이 오래 걸려 총 소요시간이 길어진다는 단점이 있다. 또한 일련의 영상을 입력받아 3D 정점 값을 계산하는 SfM을 사용하여 처리과정이 복잡하고 입력환경에 제한이 있다. 반면에 본 논문에서는 UV좌표에도 메쉬 워핑을 적용함으로써 입력 영상을 두고 UV를 맞추는 처리과정이 아닌 일반 모델의 UV자체를 변형 시킴으로써 좀 더 정확한 맵핑이 가능하다. 조명조건이 고정되어야하는 단점이 있지만 한 장의 사진만을 입력받음으로 환경의 제한이 좀 더 자유롭다는 장점이 있다. 본 논문에서는 그림 13에서 보듯 많은 데이터와 처리과정을 요구하는 [11]에 비해 떨어지지 않는 결과물을 볼 수 있다.

그림 14는 변형된 얼굴 모델이 입력 사진과 얼마나 유사한지 직관적으로 나타낸다. 그림 14의 (a)는

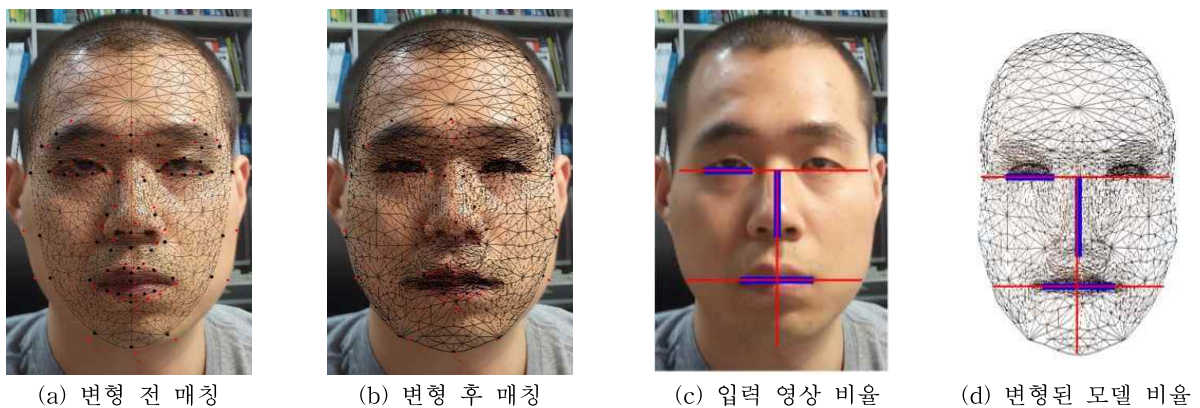


그림 14. 입력 사진과 변형된 얼굴의 비교

입력영상에 대한 특징점(빨간색)과 일반모델에 대한 특징점(검정색)을 표시하여 변형되기 전의 상태를 나타내었고, (b)는 두 단계 메쉬 워핑을 통해 변형된 얼굴 모델과 입력 사진의 매칭 상태를 보여준다.

표 1은 그림 14의 (c), (d)로 입력사진의 얼굴 구성요소(눈, 코, 입) 비율과 변형된 모델의 얼굴 구성요소 비율을 비교하여 나타낸 것이다. 특징점을 기준으로 전체적인 길이를 빨간색, 얼굴 구성요소의 길이는 파란색 선으로 나타냈다. 입 부분의 약간의 오차를 제외하면 입력 사진과 변형된 모델의 비율은 대부분 비슷한 것을 볼 수 있다.

표 1. 입력 사진과 변형된 얼굴의 눈,코,입 비율 비교 (단위: pixel)

	눈	코	입
입력 사진	0.2459	0.4058	0.4539
변형된 모델	0.2486	0.3823	0.5368

### 5. 결 론

본 논문에서는 3D 일반 모델 기반의 메쉬 워핑에 의한 3D 얼굴 모델링 방법을 제안하였다. 기존의 3D 일반 모델을 사용하는 방법보다 정점수를 크게 증가시켜 정교한 얼굴 모델을 구축하고 실험하였다. AAM을 이용하여 입력 얼굴 영상의 특징점을 추출하였고, 두 단계 메쉬 워핑을 적용하여 3D 얼굴 모델의 정점들과 UV좌표를 변형하였다. 그리고 명도 기반의 깊이 값 추출을 통해 한 장의 이미지에서 깊이 값을 추출하고 3D 모델의 깊이 값을 변형하였다. 마지막으로 변형된 3D 얼굴 모델에 두 단계 메쉬 워핑으로 변형된 UV좌표를 사용하여 텍스처를 맵핑하였다.

본 논문의 3D 얼굴 모델링은 기존의 학습 모델 기반의 3D 얼굴 모델링 방법과 다른 방법을 사용함으로써 데이터 수집의 어려움을 줄이고, 3D 일반 모델 기반 변형 방법들의 과다한 계산량과 처리 과정을 줄였다. 3D 스캐너를 활용한 모델링과 같은 방법들에 비해 정교함 면에서는 부족하지만 대상의 참여도, 활용성 등의 문제에 자유롭기 때문에 3D 게임이나 가상 성형 등 인터랙티브한 활용에 더욱 편리하고 빠른 모델링 방법이 될 수 있을 것이다.

향후, 얼굴의 옆면과 뒷면까지 포함하도록 얼굴 모델을 확장하고 개선된 AAM 알고리즘을 적용하여

얼굴 특징점 추출 성능을 높일 것이다. 또한 두 단계 메쉬 워핑의 알고리즘을 개선하여 메쉬의 역전이 가능하도록 하고 메쉬 보간시 부드러운 스플라인 함수를 사용하여 결과물의 질을 높일 것이다. 깊이 값 추출과 변형에 있어 더욱 정교한 결과를 얻을 수 있도록 구조적인 조명을 사용하고 알고리즘을 개선할 것이다. 이러한 작업들을 통해 더욱 개선된 3D 얼굴 모델링 시스템을 구축할 예정이다.

### 참 고 문 헌

[1] 이원우, 우운택, “실감 3D 콘텐츠 제작을 위한 영상 기반 객체 모델링 연구 동향”, 한국 멀티미디어학회지, 제10권, 제3호, pp. 24-33, 2006.

[2] A.D. Crocombe, A.D. Linney, J. Campos, and R. Richards, “Non-Contact Anthropometry using Projected Laser Line Distortion: Three Dimensional Graphic Visualisation and Applications,” *Optics and Lasers in Engineering*, Vol. 28, No. 2, pp. 137-155, 1997.

[3] J. Ahlberg, *Candide-3-An Updated Parameterised Face*, Report no. LiTH-ISY -R-2326, Dept. of Electrical Engineering, 2001.

[4] Y. Lee, D. Terzopoulos, and K. Waters, “Realistic Modeling for Facial Animation,” *Proc. SIGGRAPH*, pp. 55-61, 1995.

[5] In Kyu Park, Hui Zhang, and Vladimir Vezhnevets, “Image-Based 3D Face Modeling System,” *EURASIP Journal on Applied Signal Processing*, pp. 2072-2090, 2005.

[6] 이원범, 이만희, 박인규, “두 장의 영상을 이용한 스마트폰에서의 3차원 얼굴 모델링,” 한국 정보과학회논문지: 시스템 및 이론, 제38권, 제5호, 2011.

[7] V. Blanz and T Vetter, “A Morphable Model for the Synthesis of 3D faces,” *Proc of SIGGRAPH*, pp. 187-194, 1999.

[8] 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.

[9] Y. Hu, D. Jiang, S. Yan, L. Zhang, and H.

zhang, "Automatic 3D Reconstruction for Face Recognition," *Proc. of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 843-848, 2004.

[10] Z. Zhang, Z. Liu, D. Adler, M.F. Cohen, E. Hanson, and U. 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.

[11] 장용석, 정선태, 김부균, 조성원, "3D 변형가능 형상 모델 기반 3D 얼굴 모델링," 한국콘텐츠학회논문지, 제8권, 제1호, pp. 212-227, 2008.

[12] 정도준, 장재식, 박세현, 김항준, "Active Contours를 사용한 얼굴 검출," 한국멀티미디어학회 2002년도 춘계학술발표논문집(상), pp. 195-199, 2002.

[13] T.F. Cootes, C.J. Taylor, D. Cooper, and J. Graham, "Active Shape Models-Their Training and Application," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, 1995.

[14] T.F. Cootes, D.J. Edwards, S.J. Taylor, "Active Appearance Models," *IEEE Trans. Pattern Anal.* Vol. 23, No. 6, pp. 681-658, 2001.

[15] 엄기문, 안충현, 이수인, 3차원 모델링 기술 동향, 정보통신산업진흥원, [IITA] 정보통신 연구진흥원 학술정보 주간 기술 동향 1185호, 2005.

[16] L. Torresani, A. Hertzmann, and C. Bregler, "Learning Non-Rigid 3D Shape from 2D Motion," *Proc. Of Neural Information Processing Systems*, pp. 1555-1562, 2003.

[17] G. Wolberg, "Digital Image Warping," *IEEE Computer Society Press Monograph*, pp. 222- 240, 1990.

[18] 황규현, 박상훈, "동영상 합성 및 편집을 위한 특징점 기반 조명 및 그림자 추정," 한국컴퓨터그래픽스학회, Vol. 18, No. 1, pp. 1-9, 2012.



김 중 식

2010년 2월 동아방송대학교 방송 기술학과 졸업(전문학사)  
 2010년 8월 평생교육진흥원 매체 공학 졸업(공학사)  
 2012년 8월 동국대학교 멀티미디어학과 졸업(공학석사)

관심분야: 컴퓨터 그래픽스, 영상처리, 컴퓨터 비전, 방송 기술



김 진 모

2006년 2월 동국대학교 멀티미디어공학과 졸업(공학사)  
 2008년 2월 동국대학교 멀티미디어학과 졸업(예술공학석사)  
 2012년 8월 동국대학교 멀티미디어학과 졸업(공학박사)

2012년~현재 동국대학교 영상문화콘텐츠연구원 전임 연구원  
 관심분야: 컴퓨터 그래픽스, 게임공학, 절차적 모델링, 실시간 렌더링, 영상처리



조 형 제

1973년 부산대학교 전자공학과(학사)  
 1975년 한국과학기술원 전기·전자공학과(공학석사)  
 1986년 한국과학기술원 전기·전자공학과(공학박사)

1986년~현재 동국대학교 멀티미디어학과 교수  
 관심분야: 컴퓨터 그래픽스, 게임공학, 컴퓨터 비전, 디지털사운드처리