

# 고해상도 캡처 기반 실시간 얼굴 모델링과 표정 애니메이션

변 혜 원<sup>\*</sup>

## 요 약

최근에 다양한 분야에서 캐릭터 표정을 생성하는데 연기자 기반 표정 캡처 방법을 널리 사용하고 있다. 특히, 방송이나 게임 등의 분야에서는 실시간 요소와 연기자와 외형상의 차이가 큰 캐릭터의 경우에도 표정 표현이 가능한 다양성이 중요한 문제가 된다. 본 논문에서는 연기자 얼굴에서 캡처한 표정을 다양한 얼굴 모델에 적용하는 새로운 표정 애니메이션 방법을 제안한다. 이 방법을 사용하면 최소의 실시간 비용으로 캐릭터 표정을 생성하고 제어하는 것이 가능하다. 본 논문에서는 세 가지 핵심 이슈로서 얼굴 표정 캡처, 표정 대입, 표정 애니메이션을 다룬다. 그리고 연기자와 얼굴 모델 간의 특성이 매우 상이한 다양한 경우에 대하여 실험결과를 보여줌으로써 접근방식의 유효성을 증명한다.

## Real-time Facial Modeling and Animation based on High Resolution Capture

Hae Won Byun<sup>\*</sup>

## ABSTRACT

Recently, performance-driven facial animation has been popular in various area. In television or game, it is important to guarantee real-time animation for various characters with different appearances between a performer and a character. In this paper, we present a new facial animation approach based on motion capture. For this purpose, we address three issues: facial expression capture, expression mapping and facial animation. Finally, we show the results of various examination for different types of face models.

**Key words:** Animation, Performance-driven Animation(연기자 기반 애니메이션), Facial Expression(얼굴 표정), Motion Capture(모션 캡처), Expression Mapping(표정 대입)

## 1. 서 론

최근에 캐릭터 애니메이션 작업에서 연기자 표정 캡처 방식이 널리 각광을 받고 있다. 영화, 광고 등의 다양한 분야에서 고가의 표정 캡처 장비를 동원하여 정교한 표정 연출을 시도하고 있다. 그러나 생방송이나 게임과 같은 응용분야에서는 신속하게 표정 애니메이션을 제작하여 기획 단계에서부터 프로토타입

을 확인할 수 있는지 실시간으로 작동하여 생방송에 사용할 수 있는지가 보다 중요한 요소가 된다. 특히, TV 화면에서는 실제 사람의 형상을 가진 캐릭터뿐만 아니라 의인화된 동물이나 괴물, 외계인 등과 같은 독특한 캐릭터를 부각시키는데 더욱 관심이 높다. 다양한 캐릭터의 표정을 실시간으로 표현하기 위해서는 표정 캡처, 표정 대입, 그리고 표정 애니메이션의 3가지 측면이 고려되어야 한다.

※ 교신저자(Corresponding Author): 변혜원, 주소: 서울시 성북구 동선동 3가(249-1), 전화: 02)920-7615, FAX: 02)920-2250, E-mail: hyewon@sungshin.ac.kr  
접수일: 2008년 6월 14일, 완료일: 2008년 8월 4일

<sup>\*</sup> 정회원, 성신여자대학교 미디어정보학부 조교수

※ 이 논문은 2005년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2005-204-D00033)

본 논문에서는 연기자 기반 표정 애니메이션의 새로운 방법론을 제안한다. 고가의 부가장비 없이 연기자의 맨 얼굴에서 표정 데이터를 실시간으로 추적하고 이를 연기자와 특성이 상이한 얼굴 모델에 대입하여 표정 애니메이션을 효과적으로 생성한다. 그림 1에서 볼 수 있듯이, 표정 캡처에서는 한 대의 카메라를 이용하여 연속된 이미지에서 표정 데이터를 실시간으로 추적하는 방법을 제안한다. 표정 대입 부분에서는 캡처한 표정 데이터로부터 표정 예제들을 혼합하는 가중치를 계산하는 새로운 표정 대입 방법을 제안한다. 마지막 단계인 표정 애니메이션 부분에서는 표정 예제들을 가중치 기반으로 혼합하여 새로운 표정을 실시간으로 생성한다.

## 2. 관련연구

표정 애니메이션 연구 분야에서는 키프레임 보간법[1,2], 수학적 근육 모델[3], 물리적 근육 시뮬레이션[4,5], 예제 기반 애니메이션[6,7] 등이 시도되었으며, 표정 데이터의 입력을 자동화하는 수단으로서 연기자 표정 캡처 기반 애니메이션[8,9], 스피치 기반 애니메이션[10,11] 등이 제안되었다. 본 논문은 연기자 표정 캡처 기반 애니메이션과 예제 기반 애니메이션을 결합하는 연구로서 두 분야를 중심으로 관련 연구를 서술한다.

연기자의 표정 캡처를 제안한 연구는 1990년 Williams[8] 등에 의해서 처음으로 제안되었다. Terzopolous와 Waters 등[12]은 눈, 코, 입술 부분의 윤곽선을 추적하기 위해 'Snakes'[13]라는 능동적 윤곽선 모델을 도입하였다. Thalmann[14] 등은 인체 측정학적 정보를 이용하여 눈썹, 눈, 입술의 가로, 세로 길이 등과 같이 비교적 단순한 기하학적 정보를 추출하였다. Essa[15] 등은 광흐름(optical flow)을 이용하여 3차원 얼굴 모델을 구성하는 정점들의 위치를 계산하는 방법을 제안하였다. Chai[9] 등은 저비용의 비전 기반 표정 캡처 시스템이 내재적으로

가지고 있는 잡음, 저해상도, 인식 오류 등의 문제를 극복하기 위해서 모션 캡처 데이터를 결합하는 방법을 제안하였다. 기존 연구들과 달리 본 논문에서는 실시간 요소에 초점을 맞추고 있으며 눈썹, 눈, 입술 등 표정 변화에 시각적 단서가 되는 특징 부분들의 윤곽선을 정확하게 추출할 뿐만 아니라 조명이나 연기자가 변경되는 주변 상황의 변화에도 적응적인 방법을 제안한다.

실시간 표정 애니메이션 분야에서는 표정 예제를 도입하는 방법들이 널리 연구되고 있다. Blanz[16]와 Pighin[17] 등은 스캔한 3차원 얼굴 모델로 구성되는 데이터베이스에서 선택한 몇 개의 표정 예제들을 선형 혼합함으로써 새로운 얼굴을 자동으로 모델링하는 방법을 제안하였다. Joshi[18] 등은 표정 예제 기반의 방법들에서 표현하기 어려운 국소적 변형(local deformation)을 극복하기 위해서 표정 예제들을 분석하고 이를 토대로 얼굴 모델을 자동으로 분할하는 방법을 제안하였다.

최근에는 연기자 표정 캡처와 예제 기반 애니메이션을 혼합한 연구도 제안되었다. Buck[19] 등은 캡처한 표정 데이터를 사용하여 비사실적으로 렌더링된 얼굴 이미지를 제어하고 Chuang[20] 등은 3차원 얼굴 모델을 제어하는 시스템을 제안하였다.

## 3. 얼굴 표정 캡처

이 논문에서 사용하는 표정 캡처 방식은 머리에 부착하는 카메라나 얼굴에 붙이는 마커 등의 부가장비를 전혀 사용하지 않고 맨 얼굴에서 눈, 코, 입술 부분의 윤곽선을 실시간으로 추출하는 것이다. 연기자 전방에 하나의 카메라를 고정시킨 후, 연속적으로 획득한 영상을 분석하여 표정 데이터를 추출한다. 표정 캡처 과정은 색상 공간 변환, 블러프 구성, 윤곽선 추출의 세 부분으로 구성된다.

**색상공간 변환.** 눈, 코, 입술의 서로 다른 색상 성분을 동시에 추적하기 위하여 그림 2와 같이 색상 모델을 흑백 모델로 변환하는 새로운 색상 변환 함수를 제안한다. *CYMK* (Cyan, Yellow, Magenta, Black) 색상 모델에서 피부 영역은 적은 *M*, *K* 값을 보이며, *HSV* (Hue, Saturation, Value) 색상 모델에서 눈썹, 눈동자, 콧구멍과 같은 어두운 영역은 적은 *V*값을 가지고, 입술 영역은 피부색과 확연하게 구분되는 색채의 범

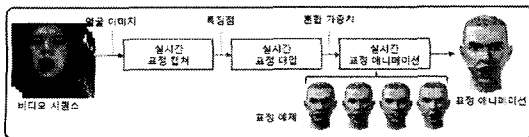


그림 1. 온라인 연기자 기반 표정 애니메이션의 전체 개요



그림 2. 색상 변환 (왼)원본 영상 (오)색상 변환된 영상

위  $H$ 를 가진다. 이러한 관찰에 근거하여 다음과 같은 색상 변환식에 의해서 픽셀  $(u, v)$ 의 명암(intensity)을 결정한다.

$$I(u, v) = w_1 M(u, v) + w_2 K(u, v) + w_3 V(u, v) + w_4 G(H(u, v)). \quad (1)$$

이 식에서  $G$ 는 입술 색상 강화 함수이다.

그러면, 이제 식(1)에서 가중치  $w_i$ 를 추정하기 위해서 3-레이어 뉴럴 네트워크를 도입한다. 다양한 조명 환경과 사람들의 얼굴을 캡처한 영상들을 이용하여 학습시킨 후, 새로 캡처한 얼굴 영상에 적합한 가중치  $w_i$ 를 실시간으로 추정한다. 입력 레이어는 11개로 구성되며, 피부, 눈, 눈썹, 콧구멍, 입술, 뺨 등 11개의 얼굴 영역에 대한 명암 대표값을 각각 입력받는다. 출력 레이어는 식(1)의 가중치들  $w_1, w_2, w_3, w_4$ 의 4개이다.

**블립 구성.** 블립은 색상이나 명암 등의 특성이 유사한 픽셀들의 집합으로서 얼굴 영상에서 눈, 코, 입술 영역 등이 블립으로 구성된다. 얼굴 영상에서 블립을 보다 빠르게 찾기 위해서 눈, 코, 입술 등의 상대적 위치에 관한 인체측정학적(anthropometric) 지식을 사용하여 각 요소들을 제한하는 경계상자(bounding box)의 크기를 초기화한다. 이때, 얼굴 영상에서 눈, 코, 입술 블립이 사라지지 않는다는 가정을 전제한다. 경계상자 내부에서 샘플링한 몇 개의 픽셀, 즉 종자점(seed point)을 중심으로 블립 성장(blob growing) 알고리즘과 모폴로지 연산을 적용하여 각 경계상자마다 가장 커다란 블립 한 개를 특징 블립으로 구성한다.

**윤곽선 추출.** 각 눈, 코, 입술의 외곽선을 추출하기 위하여 Kass 등이 제안한 Snake를 도입한다[12]. Snake는 자유도가 매우 높은 움직임에 적합하게 설계된 스플라인 곡선이므로 얼굴 영상에서 Snake가 원하지 않는 경계선으로 이동할 수 있다. 본 논문에서는 Snake 수식에서 내부적 힘을 제거하고 비교적

단순한 움직임의 특성을 보이는 눈, 코, 입술 등의 윤곽선을 표현하기 위해서 4개의 제어점을 갖는 3차 베지어 곡선을 도입한다. 이와 같은 단순화는 얼굴의 눈, 코, 입술 등을 추적하는데 과도한 자유도를 어느 정도 포기하는 대신에 효율성과 안정성을 증가시키는 효과가 있다.

제안하는 윤곽선 모델의 에너지 함수는 다음과 같이 2가지 요소로 구성된다.

$$E(v) = \int_0^1 E_{img}(v(s)) + E_{con}(v(s)) ds \quad (2)$$

이 식에서  $E_{img}$ 와  $E_{con}$ 은 각각 영상의 힘과 제한조건에 관한 힘이고,  $v(s)$ 는 베지어 곡선으로서 눈, 코, 입술 등의 윤곽선을 표현한다. 이 식에서  $E(v)$ 를 최소화함으로써 곡선  $v(s)$ 의 4개의 미지수인 제어점(control point)를 계산한다.

첫 번째 요소에서 에너지  $E_{img}$ 는 경계선 추적 함수(edge detection function)로서 다음과 같이 정의된다.

$$E_{img}(v(s)) = -w_1 |\nabla I(u, v)|^2 \quad (3)$$

여기에서  $(u, v)$ 는 3차 베지어 곡선  $v(s)$  상의 한 점이며,  $I(u, v)$ 는 식 (1)에서 구한 한 점  $(u, v)$ 의 밝기이며,  $\nabla I(u, v)$ 는 밝기의 변화(gradient)를 나타낸다. 식 (3)은 픽셀 간의 밝기 차이만을 사용하기 때문에 찾고자 하는 입술의 윤곽선을 정확하게 찾지 못하는 결과를 초래할 수 있다. 예를 들어, 그림 3에서 입술 윗부분의 윤곽선 (A)와 아래 부분의 윤곽선 (B)를 구분할 방법이 없다.

이 문제를 해결하기 위해서 식 (2)에서 두 번째 항의 제한조건에 관한 에너지  $E_{con}$ 을 도입한다. 제한조건으로서 차감 곡선(offset curve)를 도입하는데, 이는 머리 벡터의 반대방향으로 입술 윤곽선의 제어점들을 약간씩 이동시킴으로써 만들어진다. 차감 곡선은 입술 내부에 생성되는 내부곡선  $v_{in}(s)$ 과 외

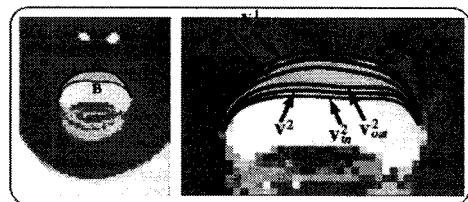


그림 3. (왼)윗입술을 나타내는 후보곡선 2개 (오)차감 곡선

부에 생성되는 외부곡선  $v_{out}(s)$ 으로 구성된다.  $v(s)$  곡선의 제한조건에 관한 에너지는 다음과 같이 정의할 수 있다.

$$E_{con}(v(s)) = w_{out}I(v_{out}(s)) - w_{in}I(v_{in}(s)) \quad (4)$$

그림 3에서 설명한 것과 같이  $w_{out}$ 보다  $w_{in}$ 이 훨씬 큰 값을 갖는 경우, 제한조건에 관한 에너지  $E_{con}(v(s))$ 의 값은 곡선  $v_2$ (그림 3에서  $v^2$ 의 경우)에 대해서 양수인 반면, 찾고자 하는 곡선  $v_1$ (그림 3에서  $v^1$ 의 경우)에 대해서는 음수가 된다.

**특징점 추출.** 얼굴 표정 캡처의 마지막 단계에서는 윤곽선을 표현하는 3차 베지어 곡선으로부터 등간격의 특징점들을 추출한다. 이때, 등간격 조건을 만족하기 위하여 원호길이를 파라미터로 이용한다. 먼저, 베지어 곡선  $v(u)$ 의 전체 길이를 구하기 위하여 파라미터  $u_s$ 와  $u_e$  사이의 곡선 길이를 나타내는 식을 다음과 정의한다.

$$L_{[u_s, u_e]}^* = \int_{u_s}^{u_e} \| \dot{v} \| du, \quad u_s \leq u_e \quad (5)$$

식 (5)의 근사값을 계산하기 위하여 적분 형태를 다음 식(6)과 같이 불연속적인 합으로 변형하면,  $\Delta u \rightarrow 0$ 일때  $L_{[u_s, u_e]}$ 는  $L_{[u_s, u_e]}^*$ 에 수렴하므로 이등분법을 적용하여 곡선길이의 근사치를 구할 수 있다. 따라서  $v(u)$ 로부터 추출하는  $N$ 개의 특징점들은  $v(0)$ 로부터의 곡선 길이가  $0, \frac{1}{N-1}L_{[0,1]}, \frac{2}{N-1}L_{[0,1]}, \dots, 1$ 이 되는 점들이다.

$$L_{[u_s, u_e]} = \sum_i \| \Delta v_i \| = \sum_i \left\| \frac{\Delta v_i}{\Delta u} \right\| \Delta u \quad (6)$$

### 4. 표정대입

표정대입 문제는 연기자 얼굴에서 추출한 특징점들의 움직임이 주어질 때, 각 프레임에서 연기자의 얼굴 표정을 모방하도록 표정 예제들을 어떻게 혼합할지를 찾는 것이다. 이 문제를 해결하기 위하여 연기자 표정에 해당하는 소스 예제와 이에 대응되는 타겟 표정 예제를 사용한다. 즉, 소스 예제들을 사용하여 타겟 예제들을 파라미터 공간에 위치시키고 타겟 예제들을 근사하는 가중치 함수를 구함으로써, 실시간으로 캡처되는 소스 입력에 대해 타겟 예제들의

혼합 가중치를 계산한다.

#### 4.1. 파라미터화

파라미터화 과정에서는 소스 예제들을 사용하여 타겟 예제들을 파라미터 공간에 위치시킨다. 소스 예제는 연기자 얼굴로부터 추출한 특징점들의 변위를 연결하여 구성된 변위 벡터이다. 이를 이용하여 타겟 예제들을 파라미터화한다. 변위 벡터  $v_i$ 는 소스 예제  $S_i$ 로부터 무표정의 기본 예제  $S_b$ 를 차감한 벡터로서 다음과 같이 정의된다.

$$v_i = s_i - s_b, \quad 1 \leq i \leq M \quad (7)$$

여기에서  $s_b$ 는 각각 무표정의 소스 예제  $S_b$ 의 특징점들의 3차원 좌표를 연결하여 구성된 벡터이고, 이와 유사하게  $s_i$ 는  $i$ 번째 소스 예제  $S_i$ 의 특징점들의 3차원 좌표를 연결하여 구성된 벡터이며,  $M$ 은 소스 예제의 개수이다. 이때, 변위벡터  $v_i$ 는 소스 예제  $S_i$ 에 대응되는 타겟 예제  $T_i$ 를  $N$ 차원 파라미터 공간에 위치시킨다. 이때,  $N$ 은  $S_i$  벡터를 구성하는 요소의 개수, 즉, 특징점들의 개수의 세배가 된다.

$v_i$ 를 구성하는 대부분의 성분들은 서로 상관관계가 있고 파라미터 공간의 차원  $N$ 은 표정 예제들의 개수  $M$ 보다 훨씬 크다. 이와 같은 중복성과 파라미터 공간의 차원을 줄이기 위하여 PCA(Principal Component Analysis)를 도입한다. 본 실험에서는 파라미터 공간의 차원을 60에서 18로 줄였으며 이에 따른 시각적 오차는 크지 않았다.

#### 4.2. 가중치 함수

이제, 파라미터 공간에 배치된 타겟 예제들을 근사하는 가중치 함수를 구할 차례이다. 파라미터 공간에 매우 성긴 형태로 배치되어 있는 타겟 예제들을 근사하는 문제는 산재 데이터 보간법으로 해결할 수 있다. 파라미터화 단계에서 구한 파라미터 벡터  $p$ 에서의 각 타겟 예제  $T_i (1 \leq i \leq M)$ 의 가중치 함수  $w_i(p)$ 를 다음과 같이 정의한다. 다음 식의 첫째 항은 선형 근사, 둘째 항은 방사형 근사를 의미하다.

$$w_i(p) = \sum_{l=0}^{\bar{N}} a_{il} A_l(p) + \sum_{j=1}^M r_{ji} R_j(p) \quad (8)$$

$A_l(p)$ 와  $a_{il}$ 은 각각 선형 기저함수와 선형 계수이

며,  $R_j(p)$ 와  $r_{ij}$ 는 각각 방사형 기저 함수와 방사형 계수이다. 그리고  $p_i (1 \leq i \leq M)$ 는 타겟 예제  $T_i$ 를 나타내는 파라미터 벡터이다. 이때, 타겟 예제를 정확하게 보간하기 위해서는 표정 예제  $T_i$ 의 가중치가  $p_i$ 에서 1이어야 하고  $p_j (i \neq j)$ 에서 0이어야만 한다.

식 (8)의 첫째 항인  $\sum_{i=0}^{\bar{N}} a_{il} A_i(p)$ 을 먼저 결정하기 위해서 선형 계수  $a_{il}$ 을 계산한다.

$$w_i(p) = \sum_{l=0}^{\bar{N}} a_{il} A_i(p) \tag{9}$$

여기에서  $A_l(p) = p^l, 1 \leq l \leq \bar{N}$ 은 선형 기저 함수로서  $p^l$ 은  $p$ 의  $l$ 번째 요소이고  $A_0(p) = 1$ 이다. 이 식에서 각 타겟 예제의 파라미터 벡터  $p_i$ 와 그 가중치  $w_i(p_i)$ 가 주어질 때, 최적화 방법(least square method)을 도입하여 미지수인 선형계수  $a_{il}$ 을 구한다.

선형 근사를 행한 이후에도 선형 근사한 초과평면(hyper-plane)과 타겟 예제 간에 여전히 남아 있는 잔여 오차(residual error)는 다음과 같이 계산된다.

$$w'_i(p) = w_i(p) - \sum_{l=0}^{\bar{N}} a_{il} A_l(p) = \sum_{j=1}^M r_{ji} R_j(p) \tag{10}$$

여기에서, 방사형 기저 함수  $R_j(p)$ 는 파라미터 공간상에 존재하는 한 점  $p$ 와  $p_j$ 간의 유클리드 거리의 함수로서 다음과 같이 정의된다. 미지수인 방사형 계수  $r_{ij}$ 는 다음의 행렬 방정식을 계산해서 구한다.

$$rR = w' \tag{11}$$

### 5. 표정 애니메이션

표정 대입 단계에서 소스 예제들을 이용하여 타겟 예제들을 근사하는 가중치 함수를 구하였다. 표정 애니메이션 단계에서는 실시간으로 입력되는 연기자의 표정 데이터와 가중치 함수를 이용하여 타겟 예제들을 혼합할 가중치를 계산함으로써 새로운 표정을 생성한다.

타겟 예제는 희로애락 등의 다양한 표정을 반영하는 표정 예제(facial examples)로서 디자이너의 의도를 반영할 수 있는 장점을 지닌다. 타겟 예제는 말하는 입모양을 반영하는 얼굴 모델과 감정을 반영하는 얼굴 모델로 구성된다. 입모양 모델은 '아' '에' '이' '오' '우' '으' '이'의 모음과 'ㄱ' 'ㄴ' 'ㅇ'의 닫힌

자음으로 구성된다. 감정 모델은 Ekman[21]이 제시한 인간의 기본적인 6가지 감정, '기쁨' '슬픔' '화남' '놀람' '두려움' '분노'로 구성된다.

연기자로부터 캡처한 표정 데이터를 입력으로 하는 실시간 시스템은 크게 세 단계로 구성된다. 첫째, 연기자 얼굴에서 추출한 특징점들의 변화를 연결하여  $N$ 차원의 변위벡터  $d_{in}$ 을 구성한다. 두 번째 단계는 방정식 (12)를 사용하여 변위 벡터  $d_{in}$ 에 표정 대입 단계에서 구한 특징 행렬  $F$ 를 곱하여  $N'$ 차원의 파라미터 벡터  $p_{in}$ 을 다음과 같이 구한다.

$$p_{in} = Fd_{in} \tag{12}$$

마지막 단계에서는 방정식 (8)에서 주어진 타겟 표정 예제  $T_i$ 에 대한 가중치 함수  $w_i$ 를 이용하여 파라미터 벡터  $p_{in}$ 에서의 가중치  $w_i(p_{in})$ 를 계산한 후, 각 표정 예제에 가중치를 곱하여 혼합함으로써 새로운 표정  $T_n$ 을 생성한다.

$$T_n(p_{in}) = T_B + \sum_{i=1}^M w_i(p_{in})(T_i - T_B) \tag{13}$$

여기에서  $T_B$ 는 무표정의 타겟 기본 모델이다.

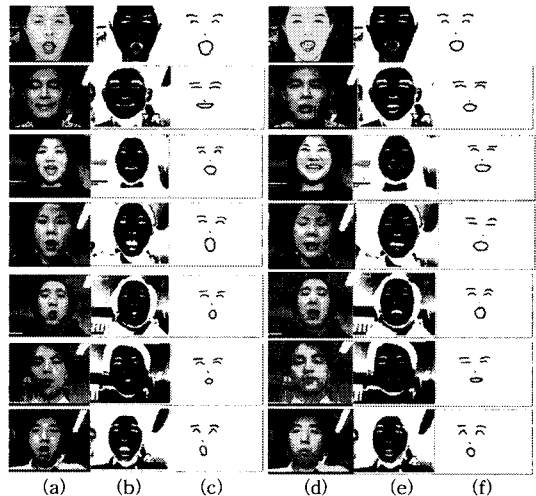


그림 4. 원본영상, 색상변환 영상, 추출 곡선

### 6. 실험결과

제안된 방법의 효용성을 실험하기 위해서 Pentium IV 1.2GHz CPU와 512MB 메모리를 장착한

PC 환경에서 실험을 진행하였다. 한 대의 디지털 카메라와 PC에 장착된 비디오 캡처 보드를 사용하여 초당 30프레임의 속도로 얼굴 표정 이미지를 캡처하였다. 연기자 표정을 보다 효과적으로 캡처하기 위하여 13W 벌브의 스탠드 조명 2개를 사용하여 연기자 얼굴을 비추게 하였다. 그림 4에서 확인할 수 있듯이 연기자 얼굴에 마커를 부착하지 않았으며 머리에 장착하는 카메라와 같은 부가장비를 사용하지 않았다. 연기자는 표정 연기 도중에 제한된 범위 내에서 머리를 약간씩 이동하거나 회전할 수 있도록 허용된다.

그림 4의 (a)와 (d)는 7명의 연기자 표정을 캡처한 영상이고, (b)와 (e)는 색상 공간 변환 과정을 수행한 후의 결과 영상이다. 색상 공간이 변환된 영상을 보면, 피부 영역의 픽셀 명암과 눈썹, 눈, 입술 등 얼굴 특징 부분의 명암이 확연하게 차이가 나는 것을 관찰할 수 있다. 그림 4의 (c)와 (f)는 얼굴 영상에서 추출

한 윤곽선을 보여 준다. 본 논문에서 제안하는 표정 캡처 방식은 초당 100 프레임 이상을 처리할 수 있어서 실시간 연기자 기반 표정 애니메이션 응용에 효과적으로 활용할 수 있다.

온라인 표정 대입의 유효성을 증명하기 위하여 연기자 얼굴과 특성이나 형태가 상이한 다양한 얼굴 예제에 대하여 실험을 진행하였다. 그림 5는 연기자 얼굴에서 캡처한 특징 곡선을 입력으로 하여 3차원 얼굴 모델의 표정을 유도한 결과를 보여 준다. 그림 5의 첫째 행은 캡처 영상을 보여주고, 둘째 행부터 다섯째 행까지는 서로 다른 모습과 형태를 지닌 4개의 얼굴 모델을 보여 준다. 표정 대입으로 생성된 애니메이션 결과는 각 얼굴 모델 본래의 특성을 유지하며 디자이너 원래의 의도를 반영하는 것을 관찰할 수 있다. 연기자 얼굴과 모델의 모습이 크게 상이한 경우에도 본 논문에서 제시한 온라인 표정 대입 방식이 효과적으로 적용된 결과를 확인할 수 있다.

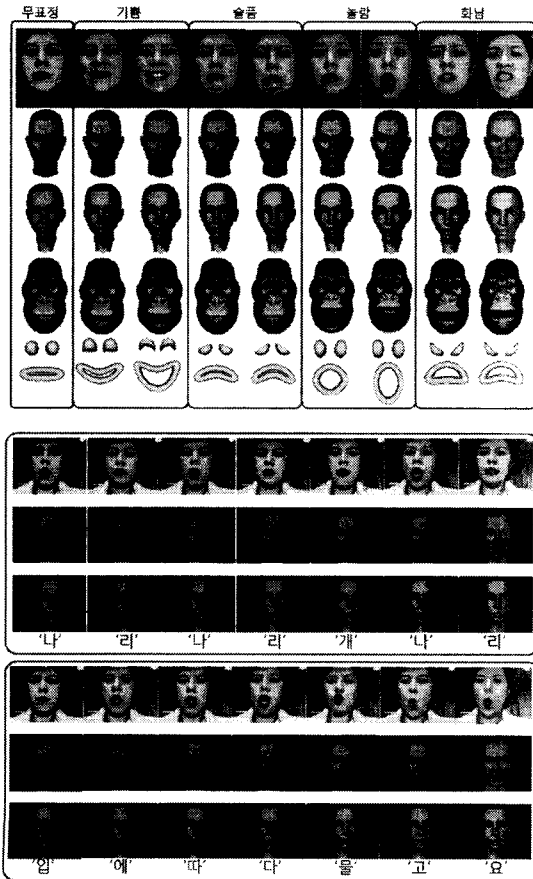


그림 5. 감정별, 입모양별 얼굴 표정 애니메이션

## 7. 결 론

본 논문에서는 고가의 부가장비 없이 편리하게 연기자의 맨 얼굴에서 표정 데이터를 실시간으로 추적하고 이를 연기자와 특성이 상이한 얼굴 모델에 대입하여 표정 애니메이션을 효과적으로 생성하는 연기자 기반 표정 애니메이션의 새로운 방법론을 제안한다. 이 방법은 연기자 표정 캡처와 표정 예제 기반 애니메이션을 결합함으로써 디자이너의 의도를 그대로 반영하며 표정 예제의 원래의 특성을 최대한 유지한다. 특히, 연기자 얼굴과 얼굴 모델 간의 특성이 크게 상이한 경우에도 올바르게 작동하는 실시간 표정 대입방법을 제안한다. 이 논문에서 제안하는 방법은 표정 캡처, 표정 대입 그리고 표정 애니메이션의 세 가지 문제를 해결한다. 표정 캡처 시, 머리에 부착하는 카메라나 얼굴에 붙이는 마커를 사용하지 않고 사용자 맨얼굴에서 실시간으로 표정을 추출한다. 표정 대입에서는 캡처한 연기자 표정을 형태나 모습이 다른 얼굴 모델에 적용하기 위한 표정 파라미터를 계산한다. 최종적으로 표정 대입 단계에서 구한 표정 파라미터를 이용하여 실시간으로 표정 애니메이션을 생성한다. 실험 결과에서도 볼 수 있듯이 본 논문에서 제안하는 표정 애니메이션은 연기자 표정을 실감나게 모방하면서도 디자이너의 의도가 반영

된 캐릭터의 표정 애니메이션을 확인할 수 있다.

이 방식에서 부족한 점은 새로운 얼굴 모델을 도입할 때마다 디자이너가 새로운 얼굴 모델에 대해 일일이 다양한 표정 예제들을 수작업으로 만들어야 한다는 점이다. 숙련된 디자이너일지라도 이와 같은 반복 작업은 오랜 시간과 지루한 노력이 필요하다. 따라서 향후에는 표정 복제를 용이하게 해결할 수 있어야 하며 눈동자 움직임이나 눈썹, 눈꺼풀, 혀 등의 작은 얼굴 움직임도 함께 제어함으로써 보다 생동감있고 풍부한 표정 표현에 관한 연구가 병행되어야 할 것이다.

### 참 고 문 헌

- [1] Parke, F. I., "Parameterized models for facial animation," *IEEE Computer Graphics*, Vol.2, No.9, pp. 61-68, 1982.
- [2] Parke, F. I. and Waters, K., *Computer Facial Animation*, AK Peters, 1994.
- [3] Waters, K., "A muscle model for animating three-dimensional facial expressions," *In Proc. of ACM SIGGRAPH*, Vol.21, No.3, pp. 17-24, 1987.
- [4] D. Terzopoulos and K. Waters, "Analysis and synthesis of facial images using physical and anatomical model," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.15, No.6, pp. 569-579, 1993.
- [5] Y. Lee, D. Terzopoulos, and K. Waters, "Realistic modeling for facial animation," *In Proc. of ACM SIGGRAPH*, pp. 55-62, 1995.
- [6] Lewis J. P., Cordner M., and Fong N., "Pose Space Deformation: A unified approach to shape interpolation and skeleton-driven deformation," *In Proc. of ACM SIGGRAPH*, pp. 165-172, 2000.
- [7] Sloan P. P., Rose c., and Cohen M. F., "Shape by Example," *In ACM Symposium on Interactive 3D Graphics*, pp. 135-143, 2001.
- [8] L. Williams, "Performance-driven facial animation," *In Proc. of ACM SIGGRAPH Conference*, pp. 235-242, 1990.
- [9] J. Chai, J. Xiao and J.K. Hodgins, "Vision-based control of 3D facial animation," *In Proc. of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2003.
- [10] Bregler, C. Covell, M., and Slaney, M., "Video rewrite: driving visual speech with audio," *In Proc. of ACM SIGGRAPH*, pp. 353-360, 1997.
- [11] Yong Cao, Wen C.Tien, Petros Faloutsos, and Frédéric Pighin, "Expressive speech-driven facial animation," *ACM Trans. on Graphics*, Vol.24, No.4, pp. 1283-1302, 2005.
- [12] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International J. of Computer Vision*, Vol.1, No.4, pp. 321-331, 1987..
- [13] D. Terzopoulos and K. Waters, "Analysis and synthesis of facial image sequences using physical and anatomical models," *IEEE Trans. of Pattern Analysis and Machine Intelligence*, Vol.15, No.6, pp. 569-579, 1993.
- [14] T. Goto, S. Kshirsagar, and N. Magnenat Thalmann, "Real time facial feature tracking and speech acquisition for cloned head," *IEEE Signal Processing Magazine*, Vol.18, No.3, pp. 17-25, 2001.
- [15] S. Basu, N. Oliver, and A. Pentland, "3D modeling and tracking of human lip motions," *In Proc. of ICCV*, 1998.
- [16] Volker Blanz and Thomas Vetter. "A morphable model for the synthesis of 3d faces", *In Proc. of ACM SIGGRAPH*, pp.187-194, 1999.
- [17] Frederic Pighin, Richard Szeliski, and David Salesin, "Resynthesizing facial animation through 3d model-based tracking," *International Conference on Computer Vision*, pp. 143-150, 1999.
- [18] Pushkar Joshi, Wen C. Tien, Mathieu Desbrun, and Frederic Pighin, "Learning controls for blend shape based realistic facial animation," *Eurographics/SIGGRAPH Symposium on Computer Animation*, 2003.

- [19] I. Buck, A. Finkelstein, C. Jacobs, A. Klein, D. H. Salesin, R. Szeliski, and K. Toyama, "Performance-driven hand-drawn animation," *In Symposium on Non Photorealistic Animation and Rendering*, pp. 101-108, 2000.
- [20] E. Chuang and C. Bregler, "Performance driven facial animation using blendshape interpolation," Computer Science Technical Report, Stanford University, 2002.
- [21] P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.



변혜원

1986년 3월 ~ 1990년 2월 연세대학교 전산학과 이학사  
1990년 3월 ~ 1992년 3월 KAIST 전산학과 공학석사  
1996년 3월 ~ 2004년 8월 KAIST 전산학과 공학박사  
1992년 4월 ~ 2006년 2월 KBS 방송기술연구소 연구원

2006년 3월 ~ 현재 성신여대 미디어정보학부 조교수  
관심분야 : 컴퓨터 그래픽스, 영상처리, HCI, 미디어아트