

얼굴 모션 추정과 표정 복제에 의한 3차원 얼굴 애니메이션

권 오 룬[†] · 전 준 철^{††}

요 약

본 논문에서는 강건한 얼굴 포즈 추정과 실시간 표정제어가 가능한 비전 기반 3차원 얼굴 모델의 자동 표정 생성 방법 및 시스템을 제안한다. 기존의 비전 기반 3차원 얼굴 애니메이션에 관한 연구는 얼굴의 움직임을 나타내는 모션 추정을 반영하지 못하고 얼굴 표정 생성에 초점을 맞추고 있다. 그러나, 얼굴 포즈를 정확히 추정하여 반영하는 작업은 현실감 있는 얼굴 애니메이션을 위해서 중요한 이슈로 인식되고 있다. 본 연구에서는 얼굴 포즈추정과 얼굴 표정제어가 동시에 가능한 통합 애니메이션 시스템을 제안 하였다. 제안된 얼굴 모델의 표정 생성 시스템은 크게 얼굴 검출, 얼굴 모션 추정, 표정 제어로 구성되어 있다. 얼굴 검출은 비모수적 HT 컬러 모델과 템플릿 매칭을 통해 수행된다. 검출된 얼굴 영역으로부터 얼굴 모션 추정과 얼굴 표정 제어를 수행한다. 얼굴 모션 추정을 위하여 3차원 실린더 모델을 검출된 얼굴 영역에 투영하고 광류(optical flow) 알고리즘을 이용하여 얼굴의 모션을 추정하며 추정된 결과를 3차원 얼굴 모델에 적용한다. 얼굴 모델의 표정을 생성하기 위해 특징점 기반의 얼굴 모델 표정 생성 방법을 적용한다. 얼굴의 구조적 정보와 템플릿 매칭을 이용하여 주요 얼굴 특징점을 검출하며 광류 알고리즘에 의하여 특징점을 추적한다. 추적된 특징점의 위치는 얼굴의 모션 정보와 표정 정보의 조합으로 이루어져있기 때문에 기하학적 변환을 이용하여 얼굴의 방향이 정면이었을 경우의 특징점의 변위인 애니메이션 매개변수(parameters)를 계산한다. 결국 얼굴 표정 복제는 두 개의 정합과정을 통해 수행된다. 애니메이션 매개변수 3차원 얼굴 모델의 주요 특징점(제어점)의 이동은 획득된 애니메이션 매개변수를 적용하여 수행하며, 정점 주위의 부가적 정점의 위치는 RBF(Radial Basis Function) 보간법을 통해 변형한다. 실험결과 본 논문에서 제안된 비전기반 애니메이션 시스템은 비디오 영상으로부터 강건한 얼굴 포즈 추정과 얼굴의 표정변화를 잘 반영하여 현실감 있는 애니메이션을 생성함을 입증할 수 있었다.

키워드 : 3차원 얼굴 애니메이션, 복제, 모션 추정, 표정 제어

3D Facial Animation with Head Motion Estimation and Facial Expression Cloning

Kwon, Ohryun[†] · Chun, Junchul^{††}

ABSTRACT

This paper presents vision-based 3D facial expression animation technique and system which provide the robust 3D head pose estimation and real-time facial expression control. Many researches of 3D face animation have been done for the facial expression control itself rather than focusing on 3D head motion tracking. However, the head motion tracking is one of critical issues to be solved for developing realistic facial animation. In this research, we developed an integrated animation system that includes 3D head motion tracking and facial expression control at the same time. The proposed system consists of three major phases: face detection, 3D head motion tracking, and facial expression control. For face detection, with the non-parametric HT skin color model and template matching, we can detect the facial region efficiently from video frame. For 3D head motion tracking, we exploit the cylindrical head model that is projected to the initial head motion template. Given an initial reference template of the face image and the corresponding head motion, the cylindrical head model is created and the full head motion is traced based on the optical flow method. For the facial expression cloning we utilize the feature-based method. The major facial feature points are detected by the geometry of information of the face with template matching and traced by optical flow. Since the locations of varying feature points are composed of head motion and facial expression information, the animation parameters which describe the variation of the facial features are acquired from geometrically transformed frontal head pose image. Finally, the facial expression cloning is done by two fitting process. The control points of the 3D model are varied applying the animation parameters to the face model, and the non-feature points around the control points are changed by use of Radial Basis Function(RBF). From the experiment, we can prove that the developed vision-based animation system can create realistic facial animation with robust head pose estimation and facial variation from input video image.

Key Words : 3D Facial Animation, Cloning, Motion Estimation, Expression Control

※ 본 연구는 21세기 프론티어 연구개발사업의 일환으로 추진되고 있는 정보통신부의 슈퍼컴퓨터스 컴퓨팅 및 네트워크 원천기반기술개발사업의 지원에 의한 것임.

† 준 회 원 : (주) 한글과 컴퓨터 연구원

†† 정 회 원 : 경기대학교 정보과학부 컴퓨터과학전공 교수 (교신저자)
논문접수 : 2007년 4월 2일, 심사완료 : 2007년 6월 7일

1. 서 론

최근 비전기반의 얼굴 모델링, 모션 추정 및 애니메이션에 관한 관심이 증대되고 있다. 기존에 수행된 다수의 비전기반 3차원 얼굴 애니메이션에 관한 연구는 얼굴의 움직임을 나타내는 모션 추정을 배제하고 얼굴 표정 생성에 관한 연구에 그 초점을 맞추고 있다. 얼굴 모션 추정과 표정 제어가 통합된 연구에서는 대부분의 경우 마커나 센서와 같은 특수 장비를 사용된다. 그러나 실감 있는 얼굴의 애니메이션을 구축 하기위해서 얼굴 표정의 변화뿐만 아니라 얼굴의 움직임에 따른 방향성을 효과적으로 반영하는 작업이 중요하며 이에 대한 관심이 증가되고 있다.[1,2,3]

3차원 얼굴의 위치와 방향을 결정하는 얼굴의 모션 추정은 비전 기반의 컨트롤, HCI, 제스처 인식과 감시 시스템 등 컴퓨터 비전 분야의 많은 애플리케이션을 위한 중요한 작업 중에 하나이다. 또한 3차원 얼굴 모션 추정을 통해 수정된 이미지를 사용하여 얼굴 인식이나 얼굴 표정 분석에 관한 연구의 문제점을 해결할 수 있다. 3차원 얼굴의 모션을 추정하는 많은 응용분야의 문제점은 얼굴의 방향이나 크기 변화와 같은 얼굴 모션에 대한 정확한 추정이 어렵다는 것이며, 부가적으로 실시간 모션을 추정하기에는 많은 연산 시간을 필요로 한다는 점이다.

얼굴 모션 추정에 관한 대표적 방법은 외형기반 방법과 모델기반 방법이 있다. 외형 기반 방법은 3차원 얼굴 모션과 얼굴 영상 속성간의 관계를 이용하여 얼굴의 모션을 추정하는 방법으로 스테레오 영상으로부터 얼굴 특징점 위치의 기하학적인 속성을 이용하여 얼굴 모션을 추정하는 방법[4], 고유공간을 이용한 방법[5] 등이 있다. 모델기반 방법에는 3차원 얼굴 모델을 이용한 방법, 2차원 평면 모델을 이용한 방법, 3차원 실린더 모델과 얼굴 텍스처를 이용한 방법[6,7] 등이 있다. 매개변수 모델 기반 방법은 매개변수들에 의해 얼굴 표정 변화가 일어난다는 가정을 기반으로 표정을 생성하는 방법으로, MPEG-4 표준에 의해 정의된 얼굴 애니메이션 매개변수와 얼굴의 AU(Action Unit)을 이용하여 표정을 생성하는 방법이 소개되었다[8,9,10].

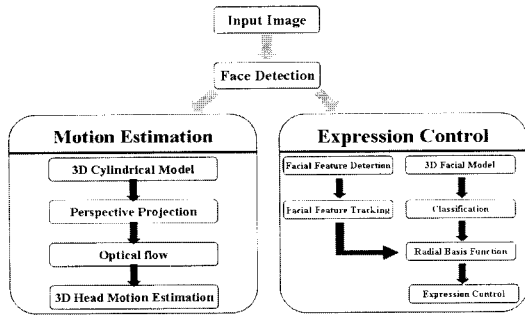
한편, 3차원 얼굴 모델을 이용한 얼굴 표정의 제어 방법에는 전통적인 키프레임 애니메이션, 근육 모델, 매개변수 모델을 이용한 방법 등 다양한 방법이 있다. 키프레임 애니메이션은 중요한 키프레임 사이에 보간법을 이용하여 생성하는 방법으로 Pighin의 연구에서 선형 보간법을 이용하여 애니메이션을 생성하였다[11]. 근육 모델 방법은 실제 사람의 근육을 얼굴 모델에 적용한 방법으로 벡터기반 근육 모델 방법이 가장 많이 사용되고 있다. 벡터기반 근육 모델 방법은 근육에 영향 받는 지역을 모션 필드로 구성하고 근육의 움직임에 따라 얼굴 모델을 변형시키는 방법으로써 얼굴 모델의 매쉬 구조에 독립적이고 각 근육에 따른 좋은 표현을 나타낼 수 있기 때문에 현재 얼굴 표정 생성에 많이 사용되고 있다[12]. 그러나 이 방법은 얼굴 모델에 근육을 제대로 위치시키는 것이 용의하지 않으며, 오히려 많은

근육의 위치는 부자연스럽고 예상하지 못한 얼굴 애니메이션을 생성할 수 있다는 문제점이 있다.

모션 추정과 표정 제어가 통합된 3차원 얼굴 애니메이션 생성에 관한 기존의 연구는 다음과 같다. Chai[1]는 얼굴 모션 추정과 얼굴 표정 제어가 통합된 얼굴 애니메이션 시스템 제안하였다. 제안된 시스템은 얼굴의 모션을 추정하는 단계와 얼굴 표정 생성에 대한 단계, 그리고 통합 단계로 구성된다. 특히, 얼굴 모션 추정 단계에서는 3차원 실린더 모델과 광류(optical flow)에 의해 생성되는 모션 필드를 이용하여 얼굴 모션 추정을 하였다. 얼굴 모델의 표정 생성은 데이터베이스에 정의된 기저들의 조합에 의해 얼굴 모델의 표정을 생성한다. 이와 같은 처리과정을 통해 생성된 얼굴 모션 매개변수와 얼굴 모델의 표정 생성을 아바타에 적용함으로써 얼굴 모델의 애니메이션을 생성하였다. Wei[2]은 얼굴의 기하학적 특징과 보간법을 이용한 얼굴 애니메이션 시스템을 제안하였다. 적외선 카메라의 일종인 IR-RED 기반의 눈 추적 시스템을 이용하여 연속된 프레임에서 눈을 검출한 후 눈의 위치를 기반으로 얼굴의 모션을 추정하는 방법을 제시하였다. 얼굴 표정을 생성하기 위해 제어점인 코, 입 주위의 얼굴 특징점을 가버 웨이블릿(Gabor Wavelet)을 이용하여 검출하며 칼만 필터를 통해 추적한다. 이때, 추적된 얼굴의 특징점과 얼굴 모델의 제어점과의 1:1 매핑관계를 이루게 되며, 추적된 특징점의 위치가 이동하게 되면 매핑되어 있는 제어점이 이동하며 벡터 보간을 통해 주위의 정점들을 이동시켜 얼굴의 표정을 생성하였다.

본 논문에서는 카메라로부터 실시간으로 입력되는 2차원 영상을 분석하여 얼굴의 모션과 표정 정보를 3차원 얼굴 모델 매핑을 통해 얼굴 모션 추정과 표정 제어가 통합된 3차원 얼굴 애니메이션 생성과 복제하는 방법을 제안하였다. 제안된 방법은 3차원 얼굴 애니메이션을 생성하기 위해 입력되는 비디오 영상으로부터 얼굴 영역을 검출하고, 검출된 얼굴에 3차원 실린더 모델을 투영하여 실린더 모델과 연속된 프레임으로부터 움직임을 나타내는 광류를 적용하여 얼굴의 포즈를 효과적으로 추적한다. 아울러 추출된 주요 표정을 나타내는 특징점 들의 움직임을 포즈 추적과 함께 추적한다. 특징점의 변위로부터 3차원 모델의 얼굴 애니메이션 매개변수(parameter)를 추출하고 얼굴 모델의 변형에 반영시킨다. 동시에 주요 특징점 주변의 정점들의 변형은 RBF(Radial Basis Function) 보간을 통해 생성한다. 본 논문에서 제안 및 개발된 애니메이션 시스템은 입력영상으로부터 얼굴 모션 추정과 표정의 추적 및 복제를 통해 3차원 얼굴 애니메이션을 실시간으로 생성하는 시스템으로, 사람 및 동물 등 다양한 3차원 얼굴 모델의 표정 복제 및 애니메이션에 적용, 그 성능을 평가하였다.

(그림 1)은 본문에서 제안된 통합 애니메이션 시스템의 구조를 나타내고 있으며, 크게 얼굴 검출, 얼굴 모션 추정 및 얼굴 표정 생성 단계로 구성되어 있다. 입력 영상으로부터 피부색 모델을 이용하여 얼굴의 후보 영역을 검출하고 후보 영역으로부터 PCA, 템플릿 매칭을 통해 정확한 얼굴



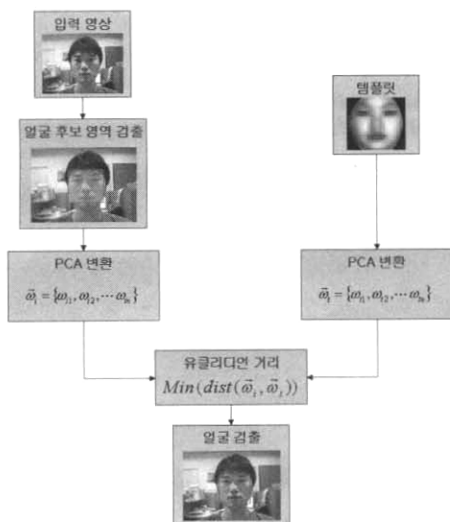
(그림 1) 얼굴 모션 추정과 표정 복제 시스템의 구성도

영역을 검출한다. 얼굴 모션 추정 단계에서는 검출된 얼굴 영역에 3차원 실린더 모델을 투영하고 광류 알고리즘을 이용하여 얼굴의 모션을 추정한다. 얼굴 표정 생성 단계에서는 검출된 얼굴 영역에서 얼굴의 주요 특징점을 검출하고 연속된 영상에서 그들을 추적한다. 얼굴 특징점의 위치와 추정된 모션 벡터로부터 기하학적 특성을 이용하여 애니메이션 매개변수를 생성한다. 애니메이션 매개변수는 3차원 얼굴 모델의 제어점에 적용되며 RBF 보간법을 이용하여 제어점 주위의 정점을 보간 함으로써 모델의 표정을 실감 있게 생성한다.

2. 얼굴 검출 및 얼굴 모션 측정

2.1 얼굴 검출

얼굴 영역의 검출은 입력영상으로부터 빠르게 얼굴 영역을 효과적으로 분리할 수 있는 비모수적 스킨 컬러 모델인 HT 색상 모델[13]을 적용하여 얼굴 후보 영역 검출을 검출한 후, 템플릿 매칭에 의하여 정확한 얼굴 영역을 검출한다. 이는 일반적으로 색상 모델을 얼굴 검출에 적용할 때 복잡한 배경이나 주변의 조명에 영향을 많이 받는 단점을 효과적으로 보완해주기 위해서이다. 템플릿 매칭 시 비교 연산



(그림 2) 얼굴 검출 과정

을 감소시키기 위하여 후보 영역을 PCA 변환을 통해 고차원의 데이터를 저차원의 특징 벡터로 변환 후 템플릿 매칭을 수행한다. 다음 (그림 2)는 이러한 얼굴검출 과정을 도식화한 것이다.

2.2 얼굴 포즈 추정

본 연구에서는 얼굴 움직임 정보를 얻기 위한 3차원 얼굴 모션 추정 과정은 3차원 실린더 모델의 투영과 광류 알고리즘을 적용하였다. 검출된 얼굴 영역에 3차원 실린더 모델을 원근 투영하게 되며 투영된 영역에 대해 광류 알고리즘을 적용하여 얼굴의 모션 정보를 추정한다. 강인한 얼굴 모션 추정을 위해 조명 보정 방법을 적용하며, 동적 템플릿을 이용하여 얼굴 영역이 외부 장애물에 의해 가려졌을 경우에도 강인한 얼굴 모션 추정을 유지 할 수 있도록 하였다.

본 논문에서 적용하는 광류 알고리즘은 연속된 두 영상에서의 화소의 이동 변위를 계산하고 나타내는 방법으로서 일반적으로 물체의 이동 또는 운동에 관한 사항을 측정하거나 인식하는데 적용될 수 있다. 본 논문에서는 Lucas-Kanade[14]가 제안한 LK 광류 알고리즘을 적용하였다. LK 알고리즘은 영상의 특징이 되는 국부영역($n \times n$ 특징 윈도우) 안의 움직임 속도는 동일하다는 가정 하에, 영상의 밝기 값의 차이를 이용하여 움직임 정보인 모션 필드를 측정한다. 즉, LK 알고리즘은 특징 윈도우 영역인 국부영역에 대한 모션 필드를 계산하는 방법으로 다른 광류 알고리즘에 비하여 정확하고 잡음에 강인한 장점을 가지고 있으며, 기본적인 특성은 다음과 같다.

영상에서의 화소를 $\mathbf{u}=(u,v)$ 라고 하고 시간 t 에서의 영상은 $I(\mathbf{u}, t)$ 로 표현할 수 있으며, 시간 $t+1$ 에서 픽셀 \mathbf{u} 는 모션 벡터에 의해 $\mathbf{u}'=F(\mathbf{u}, \mu)$ 으로 이동하게 된다. 여기서 μ 는 모션 벡터, $F(\mathbf{u}, \mu)$ 는 모션 모델이라 한다. 만일 주변 조명이 일정하다고 가정한다면 영상에서의 화소 좌표는 이동하지만($\mathbf{u} \rightarrow \mathbf{u}'$), 화소의 밝기 값은 변하지 않기 때문에 (식 1)과 같이 나타낼 수 있다.

$$I(F(\mathbf{u}, \mu), t+1) = I(\mathbf{u}, t) \quad (\text{식 1})$$

이 때, LK 알고리즘은 모션 벡터 μ 를 계산하기 위해 (식 2)와 같이 밝기 값 변화량의 제곱 값을 최소화함으로써 얻을 수 있다.

$$\min E(\mu) = \sum_{\mathbf{u} \in \Omega} (I(F(\mathbf{u}, \mu), t+1) - I(\mathbf{u}, t))^2 \quad (\text{식 2})$$

위 식에서 Ω 는 측정하고자 하는 영상의 영역이며, 모션 벡터를 얻기 위해서는 영역 Ω 안에 존재하는 화소를 이용하여 계산한다. (식 3)은 일차 테일러 확장을 통해 (식 2)를 전개한 결과이다.

$$\mu = -\left(\sum_{\Omega} (I_{\mathbf{u}} F_{\mu})^T (I_{\mathbf{u}} F_{\mu})\right)^{-1} \sum_{\Omega} I_t (I_{\mathbf{u}} F_{\mu})^T \quad (\text{식 3})$$

이 때, I_u 와 I_t 는 각각 공간적 영상 변화량과 시간적 영상 변화량을 의미하며, F_μ 는 함수 F 에서 μ 에 대해 편미분한 것이다.

3차원 실린더 모델을 원근 투영하고 투영된 모델을 이용하여 얼굴 모션을 추정하는 방법은 다음과 같다. 먼저 시간 t 에서 실린더 모델의 좌표 $X = [x, y, z]^T$ 는 (식 4)와 같이 $t + 1$ 에서의 모델 좌표는 회전과 이동 변화에 의해 결정된다.

$$X(t+1) = M \cdot X(t) = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \cdot X(t) \quad (\text{식 4})$$

행렬 M 은 (식 5)와 같이 3차원 회전 변환 행렬인 $R_{3 \times 3}$ 과 3차원 이동 변환 행렬인 $T_{3 \times 1}$ 을 동차좌표계로 표현한 것이다.

$$M = \begin{bmatrix} 1 & -\omega_z & \omega_y & t_x \\ \omega_z & 1 & -\omega_x & t_y \\ -\omega_y & \omega_x & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{식 5})$$

이 때, $[\omega_x, \omega_y, \omega_z]$ 는 3차원 회전에 관한 변수이고, $[t_x, t_y, t_z]$ 는 3차원 이동에 관한 변수를 나타낸다.

시간 $t + 1$ 에서 원근 투영에 의해 실린더 모델의 정점 X 가 영상에 투영되는 화소 \mathbf{u} 는 모션 모델 $F(\mathbf{u}, \mu)$ 에 의해 계산될 수 있으며 모션 모델은 (식 6)과 같이 모션 벡터 $\mu = [\omega_x \ \omega_y \ \omega_z \ t_x \ t_y \ t_z]$ 에 의해 표현될 수 있다.

$$\mathbf{u}(t+1) = \begin{bmatrix} x - y\omega_z + z\omega_y + t_x \\ x\omega_z + y - z\omega_x + t_y \end{bmatrix} \cdot \frac{f_L}{-x\omega_y + y\omega_x + z + t_z} \quad (\text{식 6})$$

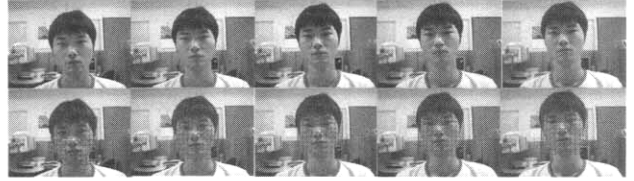
위 식에서 f_L 은 초점 거리를 나타낸다. 시간 t 에서 $\mu = 0$ 일 때, F_μ 는 (식 7)에 의해 계산될 수 있다.

$$F_\mu = \begin{bmatrix} -xy & x^2 + z^2 - yz & z & 0 & -x \\ -(y^2 + z^2) & xy & xz & 0 & z - y \end{bmatrix} \cdot \frac{f_L}{z^2}(t) \quad (\text{식 7})$$

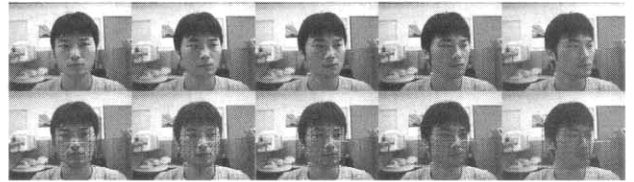
I_u, I_t, F_μ 를 이용하여 3차원 실린더 영역 안의 밝기 패턴의 움직임을 계산할 수 있으며 식 3에 의해 오차가 최소화되는 모션 벡터 μ 를 계산함으로써 3차원 얼굴 모션을 추정한다.

위의 LK 알고리즘을 이용하여 얼굴 모션을 추정할 수 있지만 좀 더 정확한 모션을 추정하기 위해 조명에 대한 보정과 반복적 가중 최소 제곱법, 에지를 포함한 픽셀에 대한 가중치, 템플릿의 밀도가 동일하지 않은 것에 대한 가중치를 적용하며 (식 3)을 다음의 (식 8)과 같이 수정하여 정의할 수 있다.

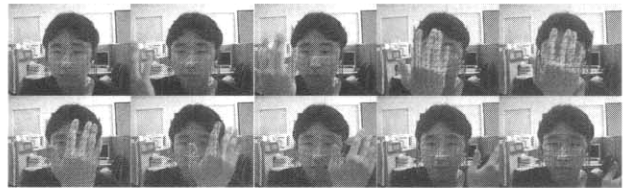
$$\mu = - \left(\sum_{\Omega} (w(I_u F_\mu)^T (I_u F_\mu)) \right)^{-1} \sum_{\Omega} (w(I_t (I_u F_\mu)^T)) \quad (\text{식 8})$$



(a) 이동 변환에 의한 모션 추정



(b) 회전 변환에 의한 모션 추정
(그림 3) 얼굴 모션 추정 결과



(그림 4) 얼굴이 가려진 경우의 모션 추정 결과

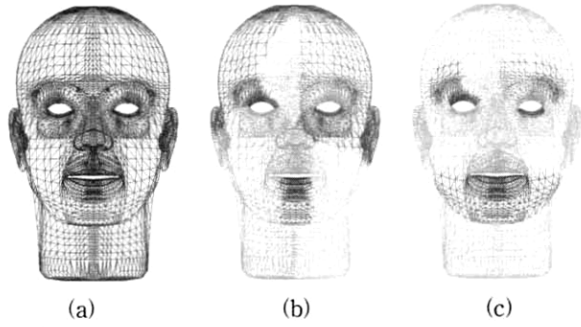
다음 (그림 3)은 연속적인 입력 영상에 대해 얼굴 모션 추정 결과이다.

또한, 동적 템플릿을 적용함으로써 얼굴에 가려짐 현상(occlusion)이 발생하였을 경우에도 (그림 4)와 같이 강한 얼굴 모션 추정을 할 수 있다.

2.3 얼굴 모델의 표정 제어

일반적으로 모델의 변형은 전역적 변형과 지역적 변형으로 나눌 수 있다. 정점에서 어느 정도 멀리 있는지의 거리에 상관없이 모든 제어점에 의해 영향을 받는다면 전역적 변형이며 정점에서 일정거리 이상 떨어진 제어점 들로부터는 영향을 받지 않는다면 지역적 변형이라 한다. 지역적 변형은 제어점 들이 충분히 가까운 경우에만 영향력을 미치게 되고 이를 표현하기 위해 근접한 정도에 따라 가중치 값을 달리하여 영향력을 표현한다. 얼굴 모델의 표정 생성을 위한 변형으로 지역적 변형을 적용한다. 지역적 변형을 위한 공간의 결정은 K 평균 알고리즘을 이용하여 눈, 눈썹, 입 주위에 존재하는 12개의 제어점에 따라 군집을 형성한다. 군집을 형성한 후 군집 내의 정점들 중 영향을 받는 영역을 정하기 위해 RBF 함수의 너비 매개변수에 의해 변형의 영향을 주는 영역으로 설정한다.

다음 (그림 5)는 (a) 12개의 제어점, (b) 제어점을 중심으로 한 K평균 알고리즘에 의한 분류 영역 및 (c) 분류 영역 내 RBF 너비 매개변수에 따라 제어점 변경시 실제 영향을 받는 영역을 각각 나타내고 있다.



(그림 5) 제어점 및 RBF 보간법에 의한 얼굴 모델 변형 대상 영역

<표 1> RBF 함수의 기저함수 형태

기저함수	$h(r)$
선형 거리	r_i
다중 이차	$\sqrt{r^2 + \sigma^2}$
TPS(Thin Plate Splines)	$r^2 \log r$
가우시언	$e^{(-r^2/\sigma)}$

본 논문에서는 3차원 얼굴 모델의 표정을 제어하기 위해 특징점 기반 모델 변형 방법을 이용하며 RBF 보간법을 적용한다. RBF 보간법은 각 제어점이 대응되는 제어점으로 이동하고 영향을 받는 영역 안에 있는 나머지 정점들을 중간적 위치로 사상하도록 하는 보간을 제공하며 3차원 기하학적 모델을 자연스럽게 부드럽게 변형시키는 특징을 가지고 있다. RBF 함수의 일반적인 형태의 변형은 다음과 같다.

변형하고자 하는 정점들을 $f(\vec{x})$ 라고 한다면 보간 함수는 $F(\vec{x}; \vec{c})$ 으로 정의할 수 있다. \vec{x} 는 정점을 나타내는 벡터이며 \vec{c} 는 각 제어점 으로부터 영향을 받는 가중치를 나타낸다. RBF 함수를 수학적으로 표현하면 (식 9)와 같다.

$$F(\vec{x}) = \sum_{i=1}^N c_i h(r) \tag{식 9}$$

N 은 제어점의 개수를 나타낸다. 또한, $h(r)$ 은 RBF 함수의 기저함수를 나타내며 매개변수인 r 은 제어점과 정점간의 유클리디언 거리를 나타낸다. RBF 함수는 기저함수에 의해 여러 가지 형태로 변형할 수 있으며 다양한 형태로 존재하는데 <표 1>과 같이 정의된다.

본 논문에서는 RBF 기저함수 중 가우시언 함수를 적용한다. 초기 제어점의 위치와 이동한 제어점의 위치 사이의 매핑관계는 $\vec{x} = x_1, x_2, \dots, x_{12}$ 으로 정의되며 RBF 함수는 (식 10)과 같이 표현된다.

$$F(\vec{x}) = \sum_{i=1}^N c_i e^{(-r_i^2/\sigma)} \tag{식 10}$$

이와 같이 RBF 함수의 해를 구하기 위해 제어점의 설정

과 가중치 c_i , 그리고 너비 매개변수 σ_i 의 값을 구해야 한다. 먼저 제어점의 위치는 눈썹, 눈, 입 주위에 존재하는 12개의 정점이다.

c_i 는 제어점의 변위를 이용하여 계산할 수 있다. 얼굴 모델 변형에서 $F(\vec{x})$ 는 초기의 얼굴 모델에 정의되어 있는 정점들이 RBF 함수에 의해 새로운 정점들의 위치로 매핑이 될 수 있도록 변형이 이루어질 때의 변위를 의미한다. 12개의 제어점 위치는 초기의 위치와 변형된 새로운 위치를 모두 알고 있으므로 (식 11)과 같이 c_i 를 계산할 수 있다.

$$\vec{c} = h^{-1} \vec{x}^{target} \tag{식 11}$$

하지만 제어점의 데이터가 노이즈를 가질 수 있기 때문에 (식 12)와 같이 행렬 h 에서 대각선 성분을 수정하여 노이즈를 줄일 수 있다.

$$\vec{c} = (c_1, c_2, \dots, c_n) = (h + \lambda I)^{-1} \vec{x}^{target}$$

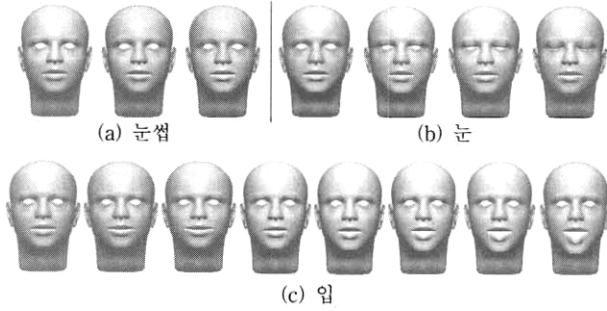
$$h = \begin{pmatrix} h(0) + \lambda I & h(\|x_2 - x_1\|) & \dots & h(\|x_{12} - x_1\|) \\ h(\|x_1 - x_2\|) & h(0) + \lambda I & \dots & h(\|x_{12} - x_2\|) \\ \vdots & \vdots & \ddots & \vdots \\ h(\|x_1 - x_{11}\|) & h(\|x_2 - x_{11}\|) & \dots & h(\|x_{12} - x_{11}\|) \\ h(\|x_1 - x_{12}\|) & h(\|x_2 - x_{12}\|) & \dots & h(0) + \lambda I \end{pmatrix} \tag{식 12}$$

λ 는 노이즈에 대한 보정을 하기 위한 계수로서 작은 영향을 주기 위한 실험에 의한 값으로 0.01을 적용한다. 만약 λ 의 값이 0이면 (식 11)과 동일하다.

RBF 함수에 의해 얼굴 모델이 지역적으로 국한된 영역에 한해서만 변형이 일어나야 하는 경우 각 기저 함수의 영향이 미치는 범위를 정할 수 있어야 한다. 기저 함수의 영향이 미치는 범위는 너비 매개변수 σ 를 조정함으로써 정할 수 있다. 너비 매개변수는 제어점 주위에 존재하는 정점들의 움직임에 영향을 미치는지에 대한 범위와 관련된 변수이다. 너비 매개변수에 따라 가우시언 함수의 특징인 종 모양의 기저 함수에 영향이 미치는 정점의 범위를 정의할 수 있다.

가우시언 함수의 특징에 따라 제어점 으로부터 멀리 떨어진 정점에는 작은 가중치를 부여하며, 가까이 있는 정점에는 큰 가중치의 값을 부여하여 제어점 주위의 정점들을 변형시킨다. 이러한 너비 매개변수에 따라 변형의 부드러운 정도가 결정되며, 너비 매개변수의 값이 커질수록 변화는 완만해지며, 값이 작아질수록 급격한 변화를 가져온다. 또한, 너비 매개변수가 증가하면 멀리 있는 제어점에 영향을 받기 때문에 전역적 변형이 일어날 수 있다.

본 논문에서 적용하는 너비 매개변수의 값은 얼굴 모델에서 정점의 분포가 균일하지 않고 산재되어 있기 때문에 각 제어점에 따른 서로 다른 너비 매개변수의 값을 적용한다. 따라서 너비 매개변수는 각 제어점에 대해 균집화된 영역에 있는 정점들 중 제어점과의 거리가 가장 긴 거리를 이용하고 각 군집마다 가중치를 적용하여 너비 매개변수의 값을 계산한다(식 13).



(그림 6) RBF 보간법에 의한 얼굴 모델 변형

$$\sigma_i = \max \| \vec{x}_i^{source} - \vec{x}_j \| \times weight \quad (i \neq j) \quad (식 13)$$

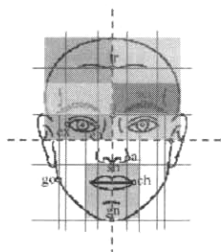
(그림 6)은 제안된 방법에 의한 얼굴 모델의 변형을 통해 생성할 수 있는 표정에 대한 결과이다.

3. 통합 얼굴 애니메이션 시스템

본 논문에서는 앞서 언급한 얼굴 모션 추정과 얼굴 모델의 표정 제어가 통합된 얼굴 애니메이션을 생성한다. 특징점을 기반으로 얼굴 모델의 표정 제어가 이루어지기 때문에 연속된 프레임에서 얼굴 특징점을 추적하게 된다. 추적된 얼굴의 특징점 위치는 얼굴의 모션과 얼굴 표정 변화의 조합으로 이루어져 있다. 따라서, 얼굴의 표정 변화를 나타내는 애니메이션 매개변수는 추적된 얼굴 특징점의 위치와 모션 추정 결과인 모션 매개변수로부터 기하학적 변환을 이용하여 얼굴의 방향이 정면이었을 때 특징점의 위치를 계산하게 되며 초기 특징점의 위치로부터 변위인 얼굴 애니메이션 매개변수를 계산할 수 있다. 이와 같이 애니메이션 매개변수와 3차원 얼굴 모델의 제어점 이동벡터간의 매핑관계를 통해 얼굴 모델을 변형시키며 모션 매개변수를 적용하여 통합된 3차원 얼굴 애니메이션을 생성한다.

3.1 얼굴 특징점 검출 및 추적

검출된 얼굴 영역으로부터 얼굴의 구조적 정보, 템플릿매칭, 그리고 히스토그램을 이용하여 특징점을 검출한다. (그림 7)은 얼굴의 구조적 정보를 표현한 것이다. 얼굴 구조적 정보로부터 얼굴의 특징을 검출하기 위한 영역을 정하게 되며 해당 영역 안에서 템플릿 매칭을 이용하여 특징의 위치를 찾게 된다.



(그림 7) 얼굴 구조적 정보



(그림 8) 눈 특징점 검출 과정



(a) 얼굴 검출 영상 (b) 얼굴 특징점 검출

(그림 9) 얼굴 특징점 검출 결과

눈을 검출하기 위해 얼굴의 윗부분 영역에서 템플릿 매칭을 수행하며 템플릿 영상과의 가장 유사한 영역을 눈 영역으로 검출하게 된다. 검출된 영역으로부터 필요한 특징점을 검출하기 위해 재배치 과정을 수행한다. 재배치 과정은 (그림 8)과 같이 검출된 영역을 Canny 에지 필터를 적용하여 검출된 에지의 히스토그램을 이용하여 특징점의 위치를 재배치한다.

눈썹과 입에 대한 특징점을 검출하는 과정은 위의 과정과 동일하다. 이와 같은 방법으로 12개의 특징점을 검출한다. (그림 9)는 얼굴의 특징점을 검출한 결과이다.

초기 프레임에서 검출된 얼굴의 특징점은 3장 3절에서 언급한 광류를 이용하여 추적한다. 광류는 LK 알고리즘을 이용하여 추적된 얼굴의 특징점의 위치가 정확하지 않았을 경우 얼굴의 구조적 정보를 이용하여 재조정한다.

얼굴의 눈썹과 입 주위에 존재하는 특징점과 달리 눈 주위에 존재하는 특징점은 눈의 특성상 눈의 움직임을 나타내는 특징점은 좌우에 위치한 특징점이 아닌 상하에 위치한 특징점이다. 따라서, 눈에 존재하는 좌우의 특징점은 광류를 이용하여 추적하며 추적된 특징점 주위의 눈 영역에 대한 모션 추정결과에 따른 2차원 영상에 대한 변환을 적용한다. 좌우 특징점에 따른 눈 영역에 대해 명암을 이용하여 임계값으로부터 이진 데이터로 분할한다(식 14).

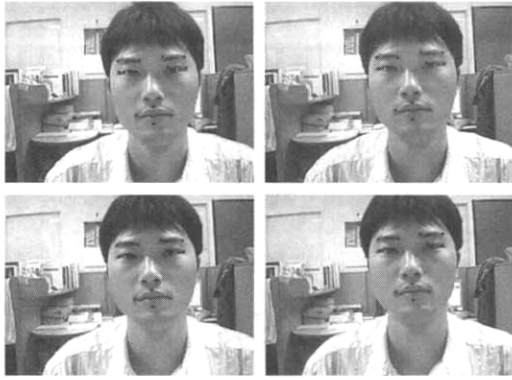
$$I(x,y) = \begin{cases} 255 & (I(x,y) > \sigma) \\ 0 & (I(x,y) \leq \sigma) \end{cases} \quad (식 14)$$

σ 는 임계값으로 60을 이용한다.

(그림 10)과 같이 눈 영역 처리 과정을 통해 눈의 상하에 위치한 특징점의 위치를 획득하며 분할된 이진 영상으로부터 히스토그램을 이용하여 눈의 상하 움직임을 측정한다. (그림 11)은 동영상으로부터 추출된 특징점을 추적한 결과이다.



(그림 10) 눈 영역 처리 과정



(a) 이동할 경우



(b) 입 부분이 움직이는 경우
(그림 11) 특징점 추적 결과

3.2 애니메이션 매개변수 추출

얼굴 특징점의 위치는 이미지에서 얼굴 모션의 변화나 표정 변화에 의해 이동한다. 얼굴 모션 변화에 의해 이동한 특징점의 위치는 기하학적 변환에 의해 얼굴의 방향이 정면이었을 때의 특징점 위치를 계산할 수 있다. 얼굴의 방향이 정면이고 초기 특징점의 위치는 \vec{v}_o 이며, 모션 변화와 표정 변화에 의해 이동한 특징점의 위치를 \vec{v}_p 라고 정의한다. (식 15)는 기하학적 변환에 의한 \vec{v}_o 와 \vec{v}_p 의 관계를 정의하고 있다.

$$\vec{v}_p = T \cdot R \cdot \vec{v}_f, \quad \vec{v}_f = \vec{\delta v} + \vec{v}_o \quad (식 15)$$

\vec{v}_f 는 현재 얼굴 방향이 정면일 때 특징점의 위치를 나타낸다. $\vec{\delta v}$ 는 초기 특징점의 위치로부터 표정 변화에 의해 변형된 특징점의 변위를 나타내며 애니메이션 매개변수이다. T 와 R 은 각각 모션 추정 결과에 의한 이동 변환행렬과 회전 변환행렬을 나타낸다.

애니메이션 매개변수 $\vec{\delta v}$ 를 계산하기 위해 \vec{v}_o 와 \vec{v}_p 의 기하학적 관계를 이용한다. \vec{v}_f 는 (식 16)과 같이 역변환을 통해 계산된다.

$$\vec{v}_f = R^{-1} \cdot T^{-1} \cdot \vec{v}_p \quad (식 16)$$

계산된 \vec{v}_f 로부터 (식 17)과 같이 애니메이션 매개변수를 계산한다.

$$\vec{\delta v} = \vec{v}_f - \vec{v}_o \quad (식 17)$$

4. 실험 결과

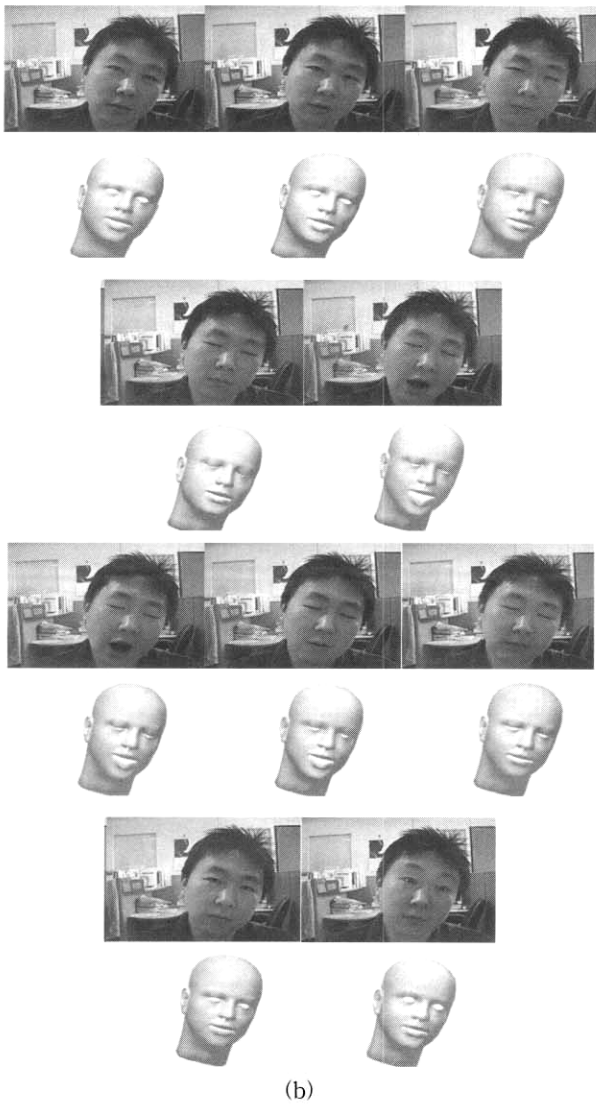
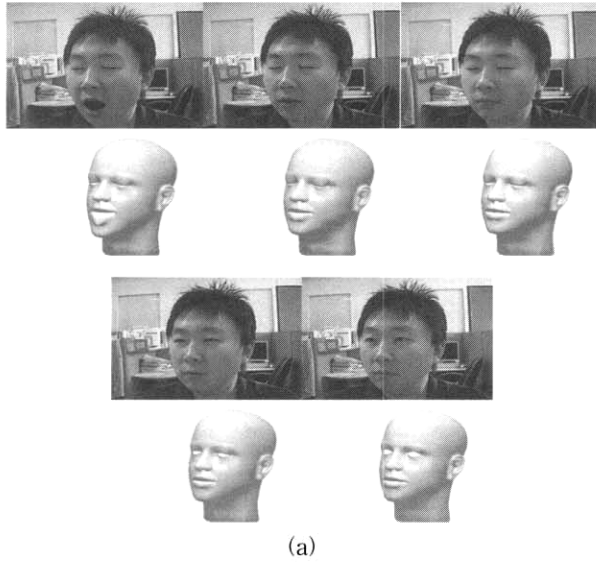
본 논문에서 구현한 얼굴 애니메이션 시스템의 실험 환경은 <표 2>와 같다. 윈도우 환경에서 Visual C++ .Net 버전을 이용하여 실험하였으며, 윈도우 프로그램을 위한 MFC 라이브러리, PC 캠 영상 획득을 위한 OpenCV 라이브러리, 3차원 얼굴 모델을 렌더링하기 위해 OpenGL 라이브러리를 사용하였다.

<표 2> 실험 환경

	구분	사양
Software	운영체제	Windows XP
	구현 언어	Visual C++
	라이브러리	MFC, OpenCV, OpenGL
	개발 도구	Visual C++ .Net 2003
Hardware	CPU	Intel Pentium 4 2.8Ghz
	RAM	DDR 512MB
	PC CAM	Lebeca PC Cam

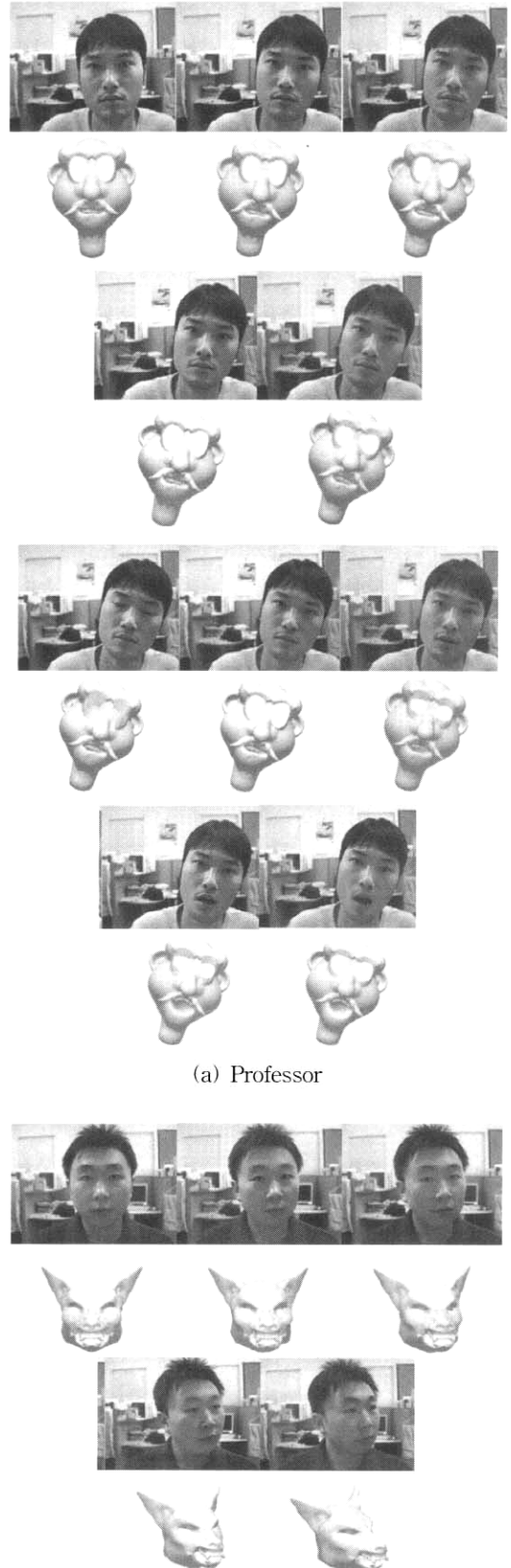
본 논문에서 제안한 방법을 통해 얼굴의 움직임이 반영된 3차원 얼굴 애니메이션 생성 결과는 (그림 12)와 같다. 애니메이션 생성 결과와 같이 얼굴의 방향 변화를 반영하면서 얼굴표정을 생성하는 것을 확인 할 수 있다.

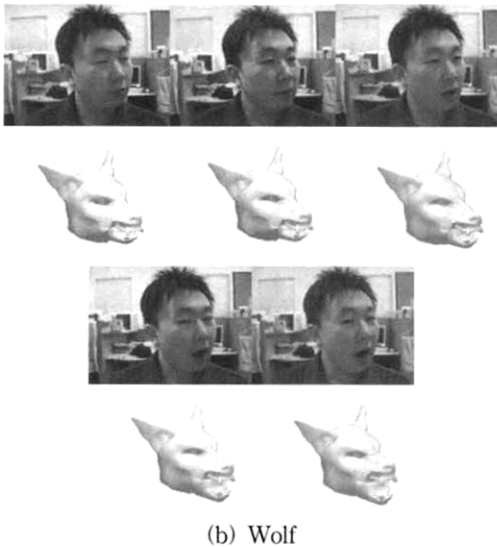




(그림 12) 얼굴 애니메이션 결과 ((a)얼굴의 회전 및 (b) 기울임)

또한, (그림 13)은 얼굴 모델이 아닌 캐릭터 모델(professor, wolf)에 표정변화를 복제한 결과이다.





(그림 13) 캐릭터 모델로의 표정 복제 결과

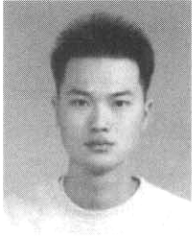
5. 결론 및 향후 연구

기존의 3차원 얼굴 애니메이션에 관한 연구는 주로 정면 얼굴의 표정 생성을 중심으로 연구되어 왔으며, 다양한 얼굴의 모션 추정 결과를 얼굴표정 생성과 결합하는 연구가 미흡하였다. 본 논문에서는 실시간으로 입력되는 2차원 영상을 분석하여 얼굴의 모션과 표정 정보를 3차원 얼굴 모델에 매핑하여 얼굴 모션 추정과 표정복제가 통합된 3차원 얼굴 애니메이션 생성 방법을 제안하였다. 3차원 얼굴 애니메이션을 실시간으로 생성하기 위하여 비디오 영상으로부터 얼굴 영역을 검출하고, 검출된 얼굴에 3차원 실린더 모델을 투영하여 실린더 모델과 연속된 프레임으로부터 움직임을 나타내는 광류를 적용하여 얼굴의 포즈를 효과적으로 추적하였다. 아울러 추출된 주요 표정을 나타내는 특징점들의 움직임을 포즈 추정과 함께 추적한다. 특징점의 변위로부터 3차원 모델의 얼굴 애니메이션 매개변수를 추출하고 얼굴 모델의 변형에 반영시켰다. 이때, 주요 특징점 주변의 정점들의 변형은 RBF(Radial Basis Function) 보간을 통해 생성된다. 본 논문에서 제안 및 개발된 애니메이션 시스템은 입력영상으로부터 얼굴 모션 추정과 표정의 추적 및 복제를 통해 3차원 얼굴 애니메이션을 실시간으로 생성하는 시스템으로, 사람 및 동물 등 다양한 3차원 얼굴 모델의 표정 복제 및 애니메이션에 적용, 그 성능을 평가하였다.

실험 결과 눈, 눈썹, 입 주위에 대한 표정을 생성과 얼굴 모션의 통합된 3차원 얼굴 애니메이션을 생성할 수 있었고 부드러운 애니메이션 결과를 얻을 수 있었다. 또한, 모션 정보와 애니메이션 매개변수를 다양한 캐릭터 모델에 적용하여 애니메이션 복제를 수행하였다. 향후 연구 과제로는 보다 다양하고 세밀한 얼굴 표정을 생성을 위한 연구가 필요하며, 이를 위하여 더 많은 얼굴 특징점을 자동 검출하기 위한 방법에 대한 연구가 있어야겠다.

참고 문헌

- [1] J. Chai, J. Xiao, and J. Hodgins, "Vision-based Control of 3D Facial Animation", In Proc. ACM/Eurographics Symposium on Computer Animation, pp. 193-206, 2003.
- [2] X. Wei, Z. Zhu, L. Yin, and Q. Ji, "A Real Time Face Tracking and Animation System", IEEE Conference on CVPR 2004, Vol. 5, pp. 71-78, 2004.
- [3] J.Y. Noh and U. Neumann, "Expression Cloning", Computer Graphics, Proceedings of ACM SIGGRAPH, pp. 277-288, 2001.
- [4] Z. Liu and Z. Ahang, "Robust Head Motion Computation by Taking Advantage of Physical Properties", Proc. HUMO2000, 2000.
- [5] D. DeCarlo and D. Metaxas, "The Integration of Optical Flow and Deformable Models with Applications to Human Face Shape and Motion Estimation", CVPR96, pp. 231-238, 1996.
- [6] G.D. Hager and P.N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination", PAMI, vol. 20, no. 10, pp. 1025-1039, 1998.
- [7] M.L. Cascia and S. Sclaroff, "Fast, Reliable Head Tracking under Varying Illumination", CVPR99, pp. 604-610, 1999.
- [8] J. Ostermann, "Animation of Synthetic Faces in MPEG-4", Computer Animation, pp. 49-51, 1998.
- [9] C. Chien, Y. Chang, and Y. Chen, "Facial Expression Analysis under Various Head Poses", In Proc. the Third IEEE Pacific Rim Conference on Multimedia, Vol. 2532, pp. 1193-1200, 2002.
- [10] A. Wojdel, L. J. M. Rothkrantz, "Parametric Generation of Facial Expressions Based on FACS", Computer Graphic Forum, Vol. 24, pp. 743-757, 2005.
- [11] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin, "Synthesizing Realistic Facial Expression from Photographs", Proc. Computer Graphics SIGGRAPH, pp. 231-242, 1998.
- [12] D. Terzopoulos and K. Waters, "Physically-based Facial Modeling, Analysis, and Animation", J. Visualization and Computer Animation, pp. 73-80, 1990.
- [13] J. Chun, K. Min, "A Realistic Human Face Modeling from Photographs by Use of Skin Color and Model Deformation," LNCS 3480, pp 1135-1143, 2005.
- [14] J.L. Barron and N.A. Thacker, "Tutorial: Computing 2D and 3D Optical Flow", Tina Memo No. 2004-012, 2005.



권 오 룬

e-mail : kor5663@hotmail.com
2005년 경기대학교 전자계산학과(학사)
2007년 경기대학교 대학원전자계산학과
(이학석사)
2007년~현재 (주)한글과 컴퓨터 연구원
관심분야: 컴퓨터 비전, 컴퓨터 애니메이션,
HCI 등



전 준 철

e-mail : jcchun@kyonggi.ac.kr
1984년 중앙대학교 전자계산학과(학사)
1991년 The Univ. of Connecticut, 컴퓨터공학과
(공학석사)
1995년 The Univ. of Connecticut 컴퓨터공학과
(공학박사)
1987년~1988년 삼성반도체 통신연구소 연구원
2001년~2002년 미시건주립대학 PRIP 연구실 객원교수
1995년~현재 경기대학교 컴퓨터과학전공 교수
관심분야: 비전기반 애니메이션, 3차원 얼굴모델링, 컴퓨터그래픽스,
HCI