

실감 있는 얼굴 표정 애니메이션 및 3차원 얼굴 합성*

Realistic Facial Expression Animation and 3D Face Synthesis

한 태 우**, 이 주 호**, 양 현 승**
Taewoo Han, Ju ho Lee, Hyun S. Yang

요 약 컴퓨터 하드웨어 기술과 멀티미디어 기술의 발달로 멀티미디어 입출력 장치를 이용한 고급 인터페이스의 필요성이 대두되었다. 친근감 있는 사용자 인터페이스를 제공하기 위해 실감 있는 얼굴 애니메이션에 대한 요구가 증대되고 있다. 본 논문에서는 사람의 내적 상태를 잘 표현하는 얼굴의 표정을 3차원 모델을 이용하여 애니메이션을 수행한다. 애니메이션에 실재감을 더하기 위해 실제 얼굴 영상을 사용하여 3차원의 얼굴 모델을 변형하고, 여러 방향에서 얻은 얼굴 영상을 이용하여 텍스처 매핑을 한다. 변형된 3차원 모델을 이용하여 얼굴 표정을 애니메이션 하기 위해서 해부학에 기반한 Waters의 근육 모델을 수정하여 사용한다. 그리고, Ekman이 제안한 대표적인 6가지 표정들을 합성한다.

1. 서 론

사람과 컴퓨터 사이의 인터페이스가 처음에는 단순한 텍스트 위주의 입출력이었지만, 현재 대부분의 시스템들은 그래픽 사용자 인터페이스(GUI)를 이용하여 편리한 환경을 제공한다. 그러나, 그래픽 사용자 인터페이스에 사용되는 윈도우, 아이콘, 입력 장치(마우스, 펜 등)는 기계적 성질이 강하기 때문에 사용자에게 편안함이나 친근감을 제공하지 못한다. 뿐만 아니라, 이러한 방법으로는 효율적으로 생각이나 정보를 전달하기에 부족함이 있다.

사람에게 가장 친숙한 의사전달 방법은 얼굴을 마주보고 서로의 생각을 교환하는 것이라 할 수 있다. 이때 의사 전달의 수단으로서 음성, 표정, 몸짓 등이 사용된다. 음성을 통한 정보 전달이 가장 확실한 방법이겠지만, 컴퓨터가 사람에게 친숙한 인터페이스를 제공하기 위해서는 위의 3가지를 모두 갖추어야 한다. 특히 표정은 사람의 감정을 잘 표현하고 정보 전

달의 의미를 보다 명확하게 해 줄 뿐만 아니라 사람의 심적 상태를 자연스럽게 표출한다.

본 논문에서는 해부적 구조에 기반한 애니메이션 방법을 사용하여 얼굴 표정을 애니메이션하며, 실감 있는 얼굴 표정 애니메이션을 하기 위하여 사람의 모습에 가까운 3차원 얼굴 모델을 합성한다. 각각의 합성된 얼굴에 대하여 동일한 파라미터에 의해 동일한 얼굴 표정 애니메이션을 수행하게 하는 시스템을 구현하는 것이 본 연구의 목적이다.

얼굴 애니메이션 기술은 컴퓨터 그래픽스 분야에서 많은 관심을 끌어오기 시작했다. 얼굴의 풍부한 표현력이 컴퓨터 그래픽스 모델 제작자(modeler)들과 만화 영화 제작자들에게 매력적으로 보이기 때문이다. 이 얼굴 애니메이션 기술은 지능형 에이전트(intelligent agent), 가상 배우(virtual actor), 멀티미디어 통신 등 여러 분야에서 활용되고 있다.

본 논문의 구성은 다음과 같다. 2장에서는 기존에 이루어진 얼굴 애니메이션 방법들에 대해 살펴본다. 3장에서는 해부학에 기반한 동적 근육 모델을 제안하고, 이 모델에 의해 작동되는 얼굴 표정 애니메이션을 설명한다. 그리고, 4장에서는 얼굴 합성에 대해 설

* 본 논문은 G7 감성공학 과제에 의하여 수행되었음

** 한국과학기술원 전산학과

Department of Computer Science, Korea Advanced Institute of Science and Technology
E-mail : bluebird@paradise.kaist.ac.kr

명한다. 특정 사람의 얼굴 영상을 이용하여 일반 모델을 변형하는 방법을 제안하고, 실제감 있는 얼굴 모델을 생성하고자 텍스처를 매핑 시킨다. 그리고, 5장에서 본 논문의 구현 결과를 보이고, 6장에서 결론을 맺는다.

2. 관련 연구

지금까지 얼굴 애니메이션에 대한 연구는 많은 사람들에 의해 진행되어 왔다[1][2][3][4]. 초기 연구에서는 2차원에서의 만화와 같은 방식으로 시작되었지만, 지금은 해부적 구조에 기반하여 물리적인 피부 조직을 모델링하고 텍스처를 매핑 하는 등 실제 사람의 모습에 가까워지고 있다. 현재까지 연구된 3차원의 얼굴 애니메이션 방법으로는 키-프레임 방법, 파라미터를 이용한 인터플레이션 방법, 해부적 구조에 기반한 방법, 물리적 성질에 기반한 방법 등이 있다.

2.1. 키-프레임 방법

키-프레임 방법은 얼굴 애니메이션에 사용된 초기 방법 중의 하나이며 캐릭터 애니메이션과 같은 간단한 형태의 애니메이션에 사용되어 왔다. 표정들에 대한 각각의 키-프레임들이 각각의 프레임마다 다르고, 키-프레임 사이의 프레임들은 인터플레이션에 의해 생성된다. 이 방법에서는 제작자가 전체 애니메이션의 키-프레임에 모델의 위치와 방향, 그리고 모든 점의 위치를 세부적으로 지정해 주어야 한다는 단점이 있다.

2.2. 파라미터에 의한 방법

파라미터에 기반한 모델에서는 각 점의 평상시의 위치와 최대로 움직였을 때의 위치를 파라미터로 지정한다. Parke가 제안한 모델에서의 파라미터는 얼굴의 형태를 나타내는 파라미터와 표정을 나타내는 파라미터로 나누어진다. 이런 파라미터에 기반한 모델은 점들을 한 묶음으로 묶어 지정된 애니메이션을 수행한다.

그러나, 이런 파라미터들은 모두 지정해 주어야 하는데, 일일이 지정하기란 쉽지 않을 뿐만 아니라 힘든 일이다. 더구나, 모델의 위상(topology)이 바뀌면 모든 파라미터들이 수정되어야 하는 불편함이 있다. 그리고, 이 모델에서는 각 점이 움직이게 되는 힘의 원천에 대해서도 고려하지 않고 있다.

2.3. 해부적 구조에 기반한 방법

Keith Waters는 이런 힘의 원천을 얼굴 피부 속에 묻혀 있는 근육으로 보고, 근육에 의해 얼굴 애니메이션을 수행하였다. 근육의 구조를 살펴보면 한쪽은 얼굴의 골격에 붙어 있고, 다른 한 부분은 피부 조직에 묻혀 있다. 각 근육에 대해 근육의 영향이 미치는 범위, 최대 움직임, 근육의 움직임이 감소(fall-off)하기 시작(start)하는 부분과 끝(finish)나는 부분을 정의하여 각 근육이 움직임에 따라 그와 연관된 피부의 점들이 움직이게 된다.

해부적 구조에 기반한 방법은 실제로 가까운 애니메이션을 할 수 있을 뿐만 아니라 애니메이션에 사용된 움직임의 동기를 자연스럽게 이해할 수 있다. 그러나, 실제 해부적 구조에 가깝게 얼굴을 모델링하는 게 쉽지 않을 뿐만 아니라, 이 방법에서는 주름살이나 피부가 부풀어오르는 것 같은 세밀한 차이를 표현하지 못한다[5].

2.4. 물리적 성질에 기반한 모델

Yuencheng Lee와 Demetri Terzopoulos, 그리고 Keith Waters는 얼굴을 물리적으로 모델링 하였다. 이들은 3D 모델을 변형 가능한 격자 조직으로 구성하였다. 이 합성 조직은 탄성을 지닌 스프링에 의해 연결된 질량을 가진 작은 크기의 덩어리들로 구성되어 있으며 3개의 층을 이룬다. 맨 아래층은 뼈가 붙어 있게 될 뼈 표면이고, 맨 위층은 표피층이며, 그 사이는 근막층으로 근육이 위치하게 된다. 활성화된 근육 섬유에 의한 작은 점 덩어리들에 가해지는 압력이 격자를 통해 계속적으로 전달됨으로써 큰 얼굴 조직의 변형을 계산해 낸다. 여기에 탄성, 질량, 부피 보존의 힘 등의 여러 가지 물리적인 성질을 고려하여 기본 파라미터를 설정하여 수치적으로 시뮬레이션 하였다.

물리적 성질에 기반한 애니메이션 방법은 가장 자연스럽게 사실에 가까운 모델을 합성해 낸다. 그러나, 수치 문제를 다루게 되어 많은 계산량을 필요로 한다.

본 연구에서는 물리적 성질에 기반한 방법보다 실행속도가 빠르고, 파라미터에 의한 방법보다 성능이 좋은 Waters가 제안한 근육에 기초한 방법에 기반하여 애니메이션을 수행하도록 한다.

3. 얼굴 표정 애니메이션

얼굴 애니메이션을 하기 위해서는 피부 메쉬들이 움직일 수 있는 동력이 주어져야 한다. 이런 힘의 원천들이 Parke의 파라미터에 기반한 모델에서는 고려되지 않고 있으나, 사실적인 표정 연출을 해 내기 위해서는 이런 힘에 의한 애니메이션이 필요하다. 이러한 고려로 사람 얼굴을 근육과 피부 조직으로 나눈 해부학에 기초한 근육 모델이 나오게 되었다.

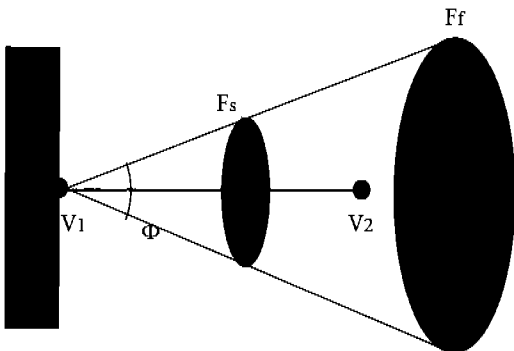
3.1. 얼굴의 근육 구조

근(muscle)은 근 양단에 결합조직성인 건(tendon)이 있어, 이것을 통하여 근은 뼈 또는 다른 근의 근막(fascia), 건 및 피부 등에 부착한다. 근이 부착되어 있는 곳은 기시(origin)라는 수축 운동의 범위가 작고 근을 고정시키고 있는 부분이다. 그리고, 근 운동을 시작하고 수축 범위가 큰 부분을 정지(insertion)라고 한다[6].

얼굴에 있는 근육들은 대부분 뼈와 피부 사이 또는 피부와 피부 사이에 개재되어 있다. 이런 얼굴 근육들을 안면근이라 하는데, 20여종의 작은 피근(cutaneous muscle)으로 이루어져 있다. 안면근은 얼굴 피부와 피부 밑의 섬유 조직을 이동시켜서 얼굴의 특유한 표정을 생성하는 것을 담당한다. 그래서 안면근을 표정근(expression muscle)이라고도 부른다.

3.2. 동적인 근육 모델

이런 안면근들을 애니메이션에 사용하기 위해 근육



V1 : 근육 벡터의 시작점(고정점)

V2 : 근육 벡터의 방향

그림 1 : 동적 근육 모델

들을 그림 1과 같은 콘(cone) 형태로 모델링 한다. 기본적으로 근육은 하나의 벡터량이다. 뼈에 붙어 있어서 움직일 수 없는 점 V1 이 벡터의 시작점이 되고, 피부 조직 속에 파묻혀 있는 점 V2 는 V1 과 함께 벡터의 방향을 나타내게 된다. 뼈에 붙어 있는 기시 V1 은 움직일 수 없고, 정지인 V2 부근에서 근육이 가장 많이 움직일 수 있다.

이러한 움직임을 명시하기 위해 근육 조직의 움직임이 감소(fall-off)하기 시작하는 부분과 감소가 끝나는 부분을 정해야 한다. 이 부분들은 근육 벡터의 길이에 대한 상대적인 위치로 정해진다. 그리고 이 근육의 영향력이 미치는 범위를 각 θ 로 지정한다.

Waters의 근육 모델은 근육 자체가 얼굴 피부에 특징적으로 설계되어 있어서 모델을 변형하는 경우에 정확한 동작을 하지 못할 수 있다. 그래서 이를 논리적 구조로 바꾼다. 논리적 근육 구조는 얼굴 모델이 변형될 때 근육의 위치나 크기가 자동적으로 결정되어 변형된 얼굴에 대해서는 같은 애니메이션을 수행할 수 있다.

3.3. 얼굴 피부 조직의 변형

근육의 힘이 미치는 영역 안에 있는 피부 조직의 한 정점(vertex)의 이동량은 근육의 시작점으로부터의 거리에 의한 이동 요인 R, 근육 벡터와 이루는 각에 의한 이동 요인 A, 탄성 상수 K와 근육에 주어질 힘 F에 의해 결정된다.

$$\Delta P = g(F, K, A, R)$$

거리에 의한 이동 요인은 근육의 시작점과 피부의 한 점 사이의 거리에 의해 결정되는 것으로 근육 수축의 시작 부분과 끝 부분에서는 이동 요인이 거의 없으며 중간 부분에서 가장 많이 이동할 수 있는 잠재력을 가지고 있다. 각에 의한 이동 요인은 근육 벡터와 이루는 각이 작을수록 많이 움직일 수 있고, 근육의 영향력이 미치는 영역 안에 있는 경우에만 움직인다. 한 정점이 여러 근육에 의해서 움직인 위치는 각 근육에 의한 정점의 이동량의 합에 의해 결정된다.

$$P = P + \sum_i g_i(F, K, A, R)$$

근육에 최소한의 자극이 주어지면 근육은 수축하게 되는데, 근육 수축과 완화 기간이 있기 전에 잠재기

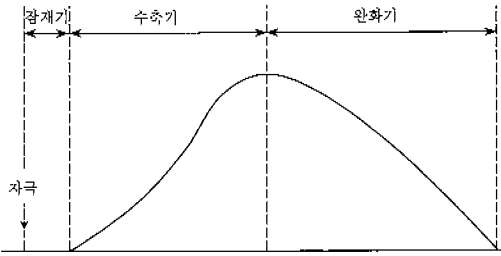


그림 2 : 근육의 수축 과정

라는 시간이 존재한다. 이 잠재기(latent period)는 근육에 자극이 주어져서 수축 현상이 일어나기 전까지의 시간이다. 그리고, 근육이 완화하는데 드는 시간은 수축하는데 필요한 시간보다 더 길다(7). 근육에 자극이 주어져 수축하고 완화된되기까지의 과정은 그림 2와 같다.

3.4. 기본적인 얼굴 표정 합성

각 근육 모델에 주어질 힘의 크기를 조절함으로써 근육들에 의한 얼굴 메쉬의 변형된 형태를 얻을 수 있다. 여기에 눈을 뜨거나 감는 정도, 입의 벌리거나 닫는 정도, 눈동자의 회전 정도 등을 함께 고려하여 Ekman이 분석한 기본적인 6가지의 얼굴 표정을 생성해 낼 수 있다(8).

4. 얼굴 합성

더 사실적이고 친근감 있는 애니메이션을 하기 위해서는 셰이딩(shading)에 의한 렌더링(rendering) 효과만으로는 부족하다. 그래서 실제 사람의 얼굴 영상을 3D 모델 상에 텍스처 매핑 시켜 특정 사람의 얼굴을 합성한다. 텍스처 매핑을 하기 전에 얼굴 모델을 얼굴 영상에 맞게 변형시킬 필요가 있다. 얼굴 모델을 변형하지 않고 텍스처 매핑을 하는 경우에는 텍스처가 늘어나고 줄어드는 정도가 얼굴 부분마다 다르기 때문에 다른 사람의 얼굴처럼 보일 수 있다.

한 얼굴 영상에 대해 일반 모델을 변형시키려면 먼저 일반 모델을 영상 평면에 프로젝션 시킨다. 그리고, 얼굴 영상과 프로젝션 된 일반 모델 사이의 매핑 관계를 정의하고, 이에 의해 일반 모델을 2차원 상에서 변형한다.

4.1. 특징선에 기초한 모델 변형

얼굴 모델을 변형하기 위해서 영상과 모델 사이의 매핑 관계를 정의한다. 이 매핑을 모든 점들에 대해

정의해 주는 것은 시간이 많이 소요되는 번거로운 작업이다. 그래서 이런 매핑에 대해 특징점들 몇 개만 지정해 주고, 그 이외의 비특징점들은 특징점들에 가해진 수정 벡터들을 인터폴레이션 하는 방법을 사용한다

그런데, 특징점들 사이를 선분에 의해 연결하면 특징점들을 하나씩 지정하는 방법보다 더 많은 특징점들을 손쉽게 지정할 수 있다. 이 방법은 특징 기반의 영상 변형 방법(Feature-based Image Metamorphosis)을 3D 모델의 변형에 적용한 것이다(9). 프로젝션 된 3차원 모델과 영상 사이에 대응하는 선분들을 지정하여 선분들에 의해 결정된 매핑 관계에 의해 모델을 변형한다.

4.2. 파노라믹 영상에 의한 텍스처 매핑

한 장의 영상을 이용하여 얼굴 합성을 하는 경우에는 옆모습이나 뒷모습을 모델링 할 수 없다. 사람 얼굴의 모든 방향을 모델링하기 위해서는 여러 장의 영상을 이용하거나 3D 레인지 스캐너(range scanner)와 같은 전문 장비를 사용한다(10). 그러나 아직까지는 3차원 스캐너와 같은 장비를 사용하는 데는 많은 비용이 든다. 본 논문에서는 여러 방향에서 얻은 영



그림 3 : 3D 일반 모델

상들을 합성하여 3차원 모델의 표면에 텍스처 매핑 시킨다.

5. 구현 결과

본 논문에서 사용한 3D 얼굴의 일반 모델은 그림 3과 같다. 이 모델은 약 2000개의 정점(vertex)과 3800여 개의 삼각형 메쉬(polygon mesh)로 구성되어 있다.

정면 영상 한 장을 이용하여 일반 모델을 변형하고 텍스처 매핑을 한 후 Ekman이 제안한 6가지의 대표적인 표정을 생성한 결과를 그림 4에 보였다.

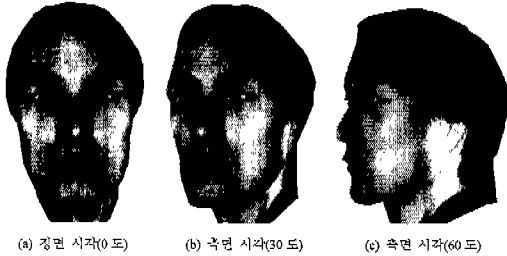


그림 4 : 텍스처 매핑 된 3차원 얼굴 모델

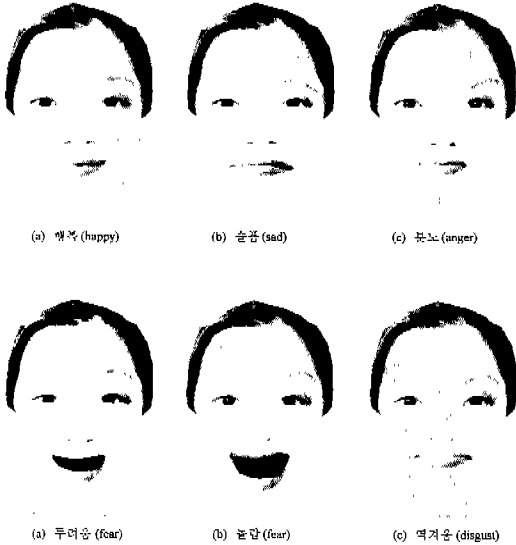


그림 5 : 6가지 대표적인 얼굴 표정 (단일 영상)

Ekman은 대표적인 사람의 얼굴 표정들을 정성적으로 기술하였기 때문에 이를 정량화 하여 표정을 생성하는데 많은 어려움이 있었다.

단일 영상으로는 얼굴 합성을 3차원적으로 완벽하게 할 수 없기 때문에 여러 장의 영상을 합성하여 다음과 같이 텍스처를 매핑 하였다.

이렇게 합성된 얼굴에 대해 본 논문에서의 얼굴 표정 애니메이션을 수행한 결과를 그림 6에 나타내었다.

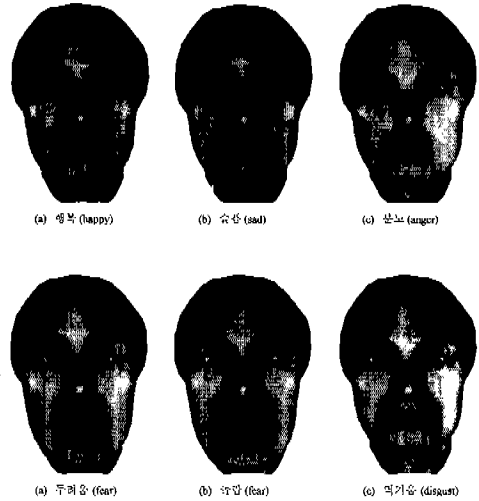


그림 6 : 6가지 대표적인 얼굴 표정 (파노라믹 영상)

6. 결 론

본 논문에서는 사용자에게 보다 친숙하고 편리한 인터페이스를 제공하고자 사실감 있는 얼굴 표정 애니메이션 시스템을 구현하였다. 구현한 시스템에서는 2차원 영상들을 이용하여 일반 모델을 변형하여 특정한 사람에 대한 3차원 모델을 얻었다. 그리고, 사람의 내적 상태를 잘 드러내고 있는 얼굴 표정들을 애니메이션 하였다. 특히, Waters의 근육 모델을 논리적 구조로 바꿈으로써 변형된 얼굴에 대해서도 같은 애니메이션을 할 수 있었다.

또한, 3차원의 얼굴 합성을 하기 위하여 여러 방향에서 얻은 얼굴 영상들을 이용하여 하나의 파노라믹 영상을 합성하고 이를 텍스처로서 사용하였다. 그리고, 모델을 2차원 상에서 변형하는 데는 특징선에 기초한 영상 변형 방법을 응용하여 사용하였다. 이 방법은 선분을 사용함으로써 기존의 특징점들을 매칭시켜 변형하는 방법에 비해 많은 특징(feature)들을 손쉽게 지정해 줄 수 있었다.

본 연구에서는 6가지 대표적인 표정에 대해서만 애니메이션을 수행하였지만, 인간의 심적 상태를 표현하는 얼굴의 표정들은 아주 많기 때문에 이에 관한 연구가 필요하다. 그리고, 모델을 자동적 또는 반자동으로 변형하는 연구가 향후 필요할 것이다.

참고 문헌

- Stephen M. Platt, Norman I. Badler, Animating Facial Expressions, In Computer Graphics, ACM, Vol. 15, No. 3, pp. 245-252, 1981.
- F.I. Parke, Parameterized models for facial animation, In IEEE Computer Graphics and Applications, Vol. 2, No. 9, pp. 61-68, Nov. 1982.
- Keith Waters, A Muscle Model for Animating Three-Dimensional Facial Expression, Proc. SIGGRAPH 87, In Computer Graphics, Vol. 21, No. 4, pp. 17-24, July 1987.
- Demetri Terzopoulos and Keith Waters, Physically-based Facial Modeling, Analysis, and Animation, In The Journal of Visualization and Computer Animation, Vol. 1, pp. 73-80, 1990.
- Y. Wu, N. Magnenat Thalmann and D. Thalmann, A Plastic Visco-Elastic Model for Wrinkles in Facial Animation and Skin Aging, In Journal of Visualization and Computer Animation, Vol. 6, No. 4, pp. 195-205, 1995.[5]
- Barbara R. Landau, Essential Human Anatomy and Physiology, 2nd Edition, Scott, Foresman and Company, 1980.
- Jacob and Francone, Structure and Function in Man, 3rd Edition, W.B. Saunders Company, 1974.
- P. Ekman, W.E. Friesen, Unmasking the Face, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1975.
- Thaddeus Beier & Shawn Neely, FeatureBased Image Metamorphosis, Proc. SIGGRAPH 92, In Computer Graphics, pp. 35-42, 1992.
- Yuencheng Lee, Demetri Terzopoulos, and Keith Waters, Realistic Modeling for Facial Animation, Proc. SIGGRAPH 95, In Computer Graphics, pp. 55-62, 1995.

Realistic Facial Expression Animation and 3D Face Synthesis

Taewoo Han, Ju ho Lee, Hyun S. Yang

(Department of Computer Science, Korea Advanced Institute of Science)

Abstract The demand of realistic facial animation has been increased to offer friendly user interface. In this thesis, we animate facial expressions representing the internal state of human using three -dimensional model. We deform 3D facial model and do texture-mapping using a synthetic image from 2D images of a specific person's face that were taken from several view angels. To animate facial expressions with the deformed 3D model, we use the modified Waters' muscle model based on anatomy. And, we synthesize the six representative facial expressions proposed by Ekman.