

3차원 얼굴 표정 애니메이션을 위한 기대효과의 자동 생성

(Automatic Anticipation Generation for 3D Facial Animation)

최정주[†] 김동선^{**} 이인권^{***}
(Jung-Ju Choi) (Dong-Sun Kim) (In-Kwon Lee)

요약 전통적인 2차원 애니메이션 제작기법에 의하면, 애니메이션의 기대효과(Anticipation)는 동작에 대한 준비단계로서 사실적이고 풍부한 표현을 하는데 중요한 역할을 한다고 알려져 있다. 얼굴 표정 애니메이션 데이터가 충분히 길다고 가정하면, 특정한 얼굴 표정에 대한 기대효과가 주어진 애니메이션 데이터에 존재할 가능성이 높다. 본 논문에서는 주어진 얼굴 표정 애니메이션 데이터로부터 얼굴 표정의 기대효과를 자동으로 추출하여 추가하는 방법을 제시한다. 먼저 애니메이션 데이터에 대해 중요요소분석법(Principal Component Analysis)을 적용하여 얼굴 모델의 모든 정점을 비슷한 움직임의 방향을 갖는 요소(Component)의 집합으로 분류한다. 각각의 요소에 대하여, 반대 방향의 움직임을 갖는 표정 애니메이션 데이터를 각 요소에 대한 기대효과로서 추출하고, 이 중에서 얼굴 모델의 위상정보를 보존하는 최적의 기대효과를 주어진 표정의 기대효과로 선택한다. 선택된 기대효과는 애니메이션 데이터의 연속성과 재생 시간을 유지하도록 주어진 표정과 적절히 합성(Blend)한다. 모션캡처 및 키프레임 작업에 의한 3차원 얼굴 표정 애니메이션 데이터를 입력으로 하여 기대효과를 자동으로 생성하는 실험결과를 제시한다. 본 논문은 전통적인 2차원 애니메이션 제작방법 중 하나인 기대효과를 3차원 얼굴 표정 애니메이션에 적용하는 새로운 방법을 제시하여, 애니메이션 제작자가 얼굴 표정을 선택하는 간단한 상호작용만으로도 기대효과가 추가된 3차원 얼굴 표정 애니메이션을 손쉽게 얻을 수 있다.

키워드 : 얼굴 표정 애니메이션, 기대효과, 중요요소분석법

Abstract According to traditional 2D animation techniques, anticipation makes an animation much convincing and expressive. We present an automatic method for inserting anticipation effects to an existing facial animation. Our approach assumes that an anticipatory facial expression can be found within an existing facial animation if it is long enough. Vertices of the face model are classified into a set of components using principal components analysis directly from a given key-framed and/or motion-captured facial animation data. The vertices in a single component will have similar directions of motion in the animation. For each component, the animation is examined to find an anticipation effect for the given facial expression. One of those anticipation effects is selected as the best anticipation effect, which preserves the topology of the face model. The best anticipation effect is automatically blended with the original facial animation while preserving the continuity and the entire duration of the animation. We show experimental results for given motion-captured and key-framed facial animations. This paper deals with a part of broad subject : an application of the principles of traditional 2D animation techniques to 3D animation. We show how to incorporate anticipation into 3D facial animations. Animators can produce 3D facial animation with anticipation simply by selecting the facial expression in the animation.

Key words : facial animation, anticipation, principal component analysis

· 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 육성지원 사업의 연구결과로 수행되었음

† 통신회원 : 아주대학교 미디어학부 교수

jungju@ajou.ac.kr

** 비회원 : (주)엔젠 연구원

janus98@nxen.com

*** 정회원 : 연세대 컴퓨터과학과 교수

iklee@yonsei.ac.kr

논문접수 : 2004년 7월 31일

심사완료 : 2004년 10월 15일

1. 서론

컴퓨터를 이용한 애니메이션 분야에 얼굴 표정 애니메이션 기법이 소개된 이후로, 사실적인 얼굴 표정을 생성하기 위한 수 많은 연구들이 수행되었다[1]. 애니메이션을 보는 관객들은 애니메이션에 등장하는 캐릭터의 얼굴에 시선을 집중하는 경향이 있기 때문에 얼굴 표정 애니메이션은 애니메이션에서 매우 중요한 역할을 담당하고 있다. 그러나 많은 경우에 관객은 애니메이션에서 사실적인 표정을 기대하거나 원하지 않으며, 따라서 사실적 애니메이션에 대한 관객의 만족도는 높지 않은 편이다. 이러한 이유로 컴퓨터 그래픽스의 렌더링 분야에 도입된 비사실적 표현 기법이 최근 애니메이션 분야에도 도입되기에 이르렀다[2-4]. 실제로 3차원 애니메이션 제작 분야에서도 비사실적인 표현을 위주로 하는 전통적 2차원 애니메이션 기법이 적용될 수 있으며[5], 실제 3차원 애니메이션 제작 사례에서도 전통적 2차원 애니메이션 기법이 중요하게 사용되고 있다[6].

관객에게 비사실적 애니메이션 기법이 일반적으로 만족도가 높은데에는 다음과 같은 몇가지 이유가 있다. 첫째 관객들은 애니메이션을 시청할 때 애니메이션을 사실적 그림의 연속으로 보지 않고, 상황의 연속으로 받아들인다는 점이다. 보통 비사실적 애니메이션에서는 등장하는 캐릭터의 동작에 대한 과장이나, 준비동작, 반작용 등을 사용하여 상황을 보다 잘 전달하는 것으로 알려져 있다. 또 다른 중요한 이유는 인간의 인지적 특징에 기인한다. 관객들은 대개 등장 캐릭터의 특정한 동작 전후의 움직임들을 통해서 동작에 대한 기대와 그 동작의 결과를 무의식 중에 미리 예상하고 즐기는 경향이 있다[5,6]. 따라서 관객은 애니메이션에서 동작 자체가 급격히 발생하는데 따른 거부감을 느끼며, 준비동작에 따른 기대효과나 결과에 따른 반작용이 없으면 애니메이션이 자연스럽지 못하다고 느끼게 된다. 실제로 잘 훈련된 애니메이션 제작자들은 특정한 동작의 앞에 기대효과를 삽입하여 전체적인 동작의 자연스러움을 유지하고 있다. 그러나, 좋은 애니메이션 제작자를 훈련시키는데는 많은 시간이 필요하며, 좋은 애니메이션 제작자에게도 기대효과를 사용한 자연스러운 애니메이션을 제작하는 것은 복잡하고 어려운 과정이다.

본 논문은 비사실적 2차원 애니메이션 제작기법을 3차원 애니메이션 제작에 응용하는 큰 주제 중에서 일부의 문제를 다룬다. 특히, 여러 가지 비사실적 2차원 애니메이션 제작 기법 중에서 기대효과(Anticipation)를 3차원 얼굴 표정 애니메이션에 적용하는 새로운 방법을 제시한다. 전반적인 적용 방법은 다음과 같다. 먼저 주어진 얼굴 표정 애니메이션 데이터로부터 기대효과를 추가할 얼굴 표정을 추출한다. 얼굴 표정 애니메이션 데

이터는 얼굴 모델을 구성하는 모든 정점에 대한 움직임을 가지고 있고, 얼굴 모델은 많은 수의 정점으로 이루어져 있기 때문에 계산의 효율성을 위하여 모든 정점을 얼굴 구성 요소(Facial Component)의 집합으로 분류한다. 이때, 정점의 움직임에 대한 중요요소분석법(Principal Component Analysis)을 통해서 얻어지는 고유벡터(Eigenvector)의 방향 유사성을 기준으로 모든 정점을 얼굴 구성 요소의 집합으로 분류한다. 얼굴 표정을 이루는 각각의 얼굴 구성 요소들에 대하여 반대 방향의 움직임을 갖는 요소의 움직임을 검색하여 해당 요소의 기대효과로 추출하고, 추출된 기대효과 중에서 얼굴 모델의 위상 정보를 유지하는 최적의 표정을 선택한다. 선택된 최적의 표정을 주어진 얼굴 표정의 앞에 위치시키고, 애니메이션 데이터의 연속성과 재생 시간을 변화시키지 않도록 합성(Blend)한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 연구의 배경이 되는 비사실적 애니메이션 기법을 3차원 애니메이션에 적용한 기존의 연구와, 전통적인 2차원 애니메이션 제작 기법으로서 기대효과의 일반적 특징을 살펴본다. 3장에서는 기대효과를 주어진 애니메이션 데이터로부터 자동으로 생성하는 자세한 기법을 설명한다. 4장 및 5장에서는 각각 본 연구의 실험 결과와 결론을 기술한다.

2. 연구의 배경

2.1 관련 연구

1970년대에 Parke에 의해서 얼굴 표정 애니메이션 분야의 연구가 도입된 이래로[7,8], 수 많은 방법들이 연구되어 왔다[1,9]. 최근까지 얼굴 표정 애니메이션의 주된 관심사는 사실적인 표정을 생성하는 방법에 맞추어져 왔다. 최근 2차원 애니메이션 제작에 사용되는 비사실적 기법들이 3차원 애니메이션의 제작에 빠르게 도입되고 있으나, 비사실적 애니메이션 기법을 얼굴 표정 애니메이션에 적용한 연구는 현재까지는 알려진 사례가 없는 상황이다.

1930년대 이래 디즈니와 같은 전통적 애니메이션 제작자들은 나름대로의 만화적인 비사실성에 입각한 전통적 2차원 애니메이션 기법을 개발하여 기술 축적을 해왔으며, 이러한 기법 중에서 상당수는 현재 3차원 애니메이션 제작 과정에도 도입되었다[5]. 예를 들어 Opaloch와 Maddoc[10]은 음향수 곡면을 이용한 공룡의 모델링 및 애니메이션에 2차원 애니메이션의 다양한 효과를 도입하였다[10]. Rademacher는 2차원 애니메이션에서 자주 관찰되는 물체의 불규칙한 변형을 3차원 모델에 적용하여, 변형된 모델을 키프레임으로 지정하고 키프레임 사이를 보간(Interpolation)하여 3차원 애니메이

션을 완성하는 기법을 사용하였다[11]. Bregier는 2차원 애니메이션 비디오로부터 직접 동작을 획득(Capture)하여 3차원 모델에 적용하는 방법을 사용하였으며[12], Agarwala는 실사 영상으로부터 획득한 동작을 만화적 애니메이션에 적용하는 반자동화된 기법을 제시하였다[2]. Chenny는 물체의 질감 표현에 효과적인 것으로 알려진 2차원 애니메이션의 전통적인 기법 중 하나인 squash and stretch 기법을 물리 모의 실험을 통하여 적용하는 방법을 제시하였다[3].

2.2 기대효과의 특징

본 논문에서 제시하는 방법은 전통적인 2차원 애니메이션의 기법 중 하나인 기대효과의 시각적 특징에 근거하고 있다[6]. 많은 애니메이션 제작자들은 어떤 특정한 동작을 디자인할 때, 기대효과와 동작 자체, 그리고 동작에 따른 반작용의 연속된 순서를 하나의 독립된 동작 단위로 본다. 이때 기대효과는 동작에 대한 준비동작의 역할을 하게 된다. 인간의 많은 동작은 실제로 기대효과를 포함하고 있다. 그 주된 이유로서 인간은 어떤 특정한 동작을 취하기 전에 무의식적으로 동작에 대한 생각을 통해서 준비동작을 하게 되며, 준비동작을 통해서 원하는 동작에 대한 에너지를 얻게 된다. 따라서 일반적으로 큰 동작일 수록 큰 준비동작을 가지게 된다. 방향을 가리키기 위하여 팔을 뻗는 동작 보다는 의자에 앉아 있다가 일어나는 동작의 움직임이 더 크기 때문에 더 큰 준비동작, 즉 더 큰 기대효과를 가지게 된다. 화살을 더 멀리 쏘기 위해서 활시위를 더 많이 당겨야 하는 것은 자명하며, 이때 활시위를 당기는 것이 화살을 쏘는 동작에 대한 준비동작의 역할을 하게 된다.

기대효과는 얼굴 표정의 애니메이션에서도 관찰된다. 예를 들면, 놀라운 광경을 목격했을 때 사람은 보통 중립 표정에 비해서 입과 눈이 커지는 현상이 발생한다. 이때 잘 훈련된 애니메이션 제작자는 반대되는 표정으로서 감은 눈과 작게 오무린 입을 갖는 표정을 중립 표정과 놀란 표정의 사이에 삽입하여 놀란 표정의 사실성을 증대시킨다. 이 경우 반대되는 표정이 놀란 표정의 기대효과로 작용하게 된다.

이상의 관찰을 정리하여 보면, 특정한 동작에 대한 기대효과는 일반적으로 주어진 동작의 반대 방향으로 일어나며, 주어진 동작의 크기와 속도에 비해서는 상대적으로 작고 느리게 일어난다. 따라서 기대효과를 적용한 동작은 결과적으로 좀 더 빠르고 상대적으로 크게 움직이는 것처럼 보여지게 된다. 이러한 기대효과의 특징은 얼굴 표정 애니메이션에도 그대로 적용된다. 다만 얼굴은 눈, 코, 입, 눈썹 등의 여러 얼굴 구성 요소들로 이루어져 있고, 표정은 이러한 얼굴 구성 요소의 조합으로 나타나기 때문에 얼굴 구성 요소의 기대효과로부터 표

정의 기대효과를 찾아야 하는 문제를 추가적으로 해결해야 한다. 본 논문에서는 얼굴 표정 애니메이션 데이터로부터 얼굴 구성 요소들을 분할하는 방법, 기대효과의 일반적인 특징에 따라 얼굴 구성 요소별 기대효과를 추출해 내는 방법, 얼굴 구성 요소별 기대효과로부터 얼굴 표정에 대한 최적의 기대효과를 선택하는 방법을 제시한다.

3. 기대효과의 자동 생성

이번 장에서는 주어진 얼굴 표정 애니메이션 데이터로부터 기대효과를 자동으로 생성하는 자세한 방법을 설명한다. 본 논문에서 제시하는 방법은 먼저 애니메이션 데이터로부터 기대효과를 추가할 표정을 찾고, 애니메이션 데이터를 검색하여 해당 표정에 대한 적절한 기대효과를 추출하고, 추출된 기대효과를 주어진 표정과 합성한다. 이에 대한 개략적인 절차는 다음과 같다.

입력 : 3차원 얼굴 표정 애니메이션 데이터(키프레임 혹은 모션 캡처 애니메이션 데이터)

출력 : 기대효과가 추가된 3차원 얼굴 표정 애니메이션 데이터

단계 1 : 입력에 대한 필터링 작업을 수행하여 부드러운 애니메이션 데이터를 획득한다.

단계 2 : 얼굴 모델의 모든 정점에 대한 움직임을 분석하여 공통된 움직임의 방향을 갖는 얼굴 구성 요소의 집합으로 분할한다.

단계 3 : 기대효과를 추가할 얼굴 표정을 선택한다.

단계 4 : 선택된 얼굴 표정을 이루는 모든 얼굴 구성 요소에 대하여 반대 방향의 움직임을 갖는 표정을 추출한다.

단계 5 : 추출된 표정 중에서 최적의 표정을 선택한다.

단계 6 : 최적의 표정을 단계 3에서 선택된 표정의 앞부분에 추가하고 합성한다.

3.1 애니메이션 데이터의 필터링

실제 얼굴 표정을 실시간으로 캡처하여 만든 얼굴 표정 애니메이션 데이터는 많은 잡음(Noise)을 포함하고 있을 가능성이 높다. 이러한 잡음은 다음 단계에서 데이터를 분석하여 얼굴 구성 요소를 생성하는데 문제를 일으킨다. 본 논문에서는 잡음을 제거하여 데이터를 부드럽게 하는데 영상처리 분야에서 사용하는 가우시안 필터(Gaussian Filter)를 애니메이션 데이터에 적용하는 방법을 사용하였다[13].

프레임의 수가 m 인 애니메이션 데이터가 주어지고, $2^{-n} \leq m \leq 2^{n+1}$ 일 때, $f_b = n$ 을 만족하는 주파수 밴드 수 f_b 를 정할 수 있다. 필터링을 한 후에는 데이터의 크기가 줄어들고 다음 단계에서 다시 필터를 적용하기 전에

데이터를 스케일 한 후에 다시 필터링 작업을 하는 것이 일반적이지만, 필터 커널 w 에 적절히 0을 삽입하여 커널을 확장한 후 사용할 수 있다.

$$\begin{aligned} w_1 &= [c, b, a, b, c] \\ w_2 &= [c, 0, b, 0, a, 0, b, 0, c] \\ w_3 &= [c, 0, 0, b, 0, 0, a, 0, 0, b, 0, 0, c] \\ &\dots \end{aligned}$$

이때 $a=3/8, b=1/4, c=1/16$ 와 같다. 기본적으로 G_0 를 원본 애니메이션 데이터라고 할때, 필터를 적용한 후 다음 밴드는 점화식 $G_{k+1} = w_{k+1} \otimes G_k, 0 \leq k < fb$ 을 통해서 얻어진다. 위의 필터 커널을 사용하면,

$$G_{k+1}(i) = \sum_{m=-2}^2 w_{k+1}(m) G_k(i + m \cdot 2^k)$$

와 같다. 본 논문에서는 원본 얼굴 표정 애니메이션 데이터에 대해 위의 방법을 적용하여 잡음을 제거한 G_3 을 사용하여 얼굴 구성 요소를 생성한다(그림 1 참조).

3.2 얼굴 구성 요소의 생성

본 논문에서 얼굴 구성 요소는 얼굴 모델을 구성하는 모든 정점 중에서 유사한 움직임을 갖는 정점의 집합으로 정의되며, 보통 얼굴의 눈, 코, 입 등 얼굴을 구성하는 특징적인 요소(Facial Features)에 대응된다. 기존의 연구를 통하여 얼굴의 구성 요소를 찾는 다양한 방법이 제시되었다[14-16]. 기존의 연구 결과들은 매쉬로 표현된 얼굴 모델에서 정점의 기하학적 위치에 기반하여 얼굴 구성 요소를 찾는 방법을 주로 사용하여, 필연적으로 사용자의 상호작용(Interaction)을 요구하게 된다. 따라서 주어진 얼굴 표정 애니메이션 데이터에서 정점의 움직임이 유사함에도 불구하고 같은 얼굴 구성 요소로 분류되지 않거나, 혹은 전혀 다른 움직임을 갖는 정점이 같은 요소로 분류되는 가능성이 존재하게 된다. 예를 들면, 입의 윗 입술과 아랫 입술은 주어진 애니메이션 데이터에서 같은 방향으로의 움직임이 있을 수도 있고, 때로는 다른 방향의 움직임이 있을 수도 있다. 이때 단순히 얼굴 모델 데이터의 정점 위치 정보만으로 얼굴 구성 요소를 분류하면 애니메이션 데이터를 고려했을 때

올바른 얼굴 구성 요소를 분류할 수 없는 결과를 얻게 된다. 본 논문에서는 정점의 위치 정보를 기반으로 사용자의 상호작용이 필요했던 기존의 방법과는 달리, 정점의 애니메이션 정보를 기반으로 사용자의 상호작용이 필요 없이 얼굴 구성 요소를 분류하는 중요요소분석법에 기반한 새로운 방법을 제시한다.

중요요소분석법은 다차원 정보로부터 특징적인 패턴을 찾는 수학적 도구이다[17]. 얼굴 모델을 구성하는 모든 정점의 애니메이션 데이터에 대해서 중요요소분석법을 수행하면, 정점의 움직임에 대한 고유벡터(Eigenvector)를 얻을 수 있다. 이때 얻어지는 고유벡터의 수는 애니메이션 데이터의 공간 차원과 같다. 이 중에서 최대의 고유치(Eigenvalue)를 갖는 고유벡터의 방향을 해당 정점의 주방향(Principal Direction)으로 정한다. 얼굴 모델에 존재하는 임의의 두 정점에 대해서, 두 정점의 주방향의 차이가 주어진 한계값(Error Tolerance) 내에 있으면 해당 두 정점을 하나의 얼굴 구성 요소로 분류한다. 결과적으로 얼굴 모델에 존재하는 모든 정점에 대해서 주방향을 너비우선탐색방식(Breadth-First-Search Paradigm)을 사용하여 비교하면 얼굴의 모든 정점을 움직임의 방향이 비슷한 얼굴 구성 요소의 집합으로 분류할 수 있게 된다. 그림 2는 얼굴 표정 애니메

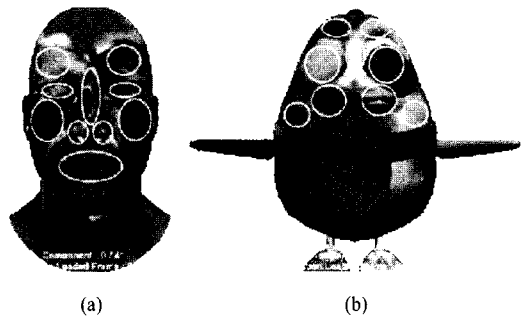


그림 2 얼굴 모델 정점에 대한 중요요소분석법을 통해 분류한 얼굴 구성 요소

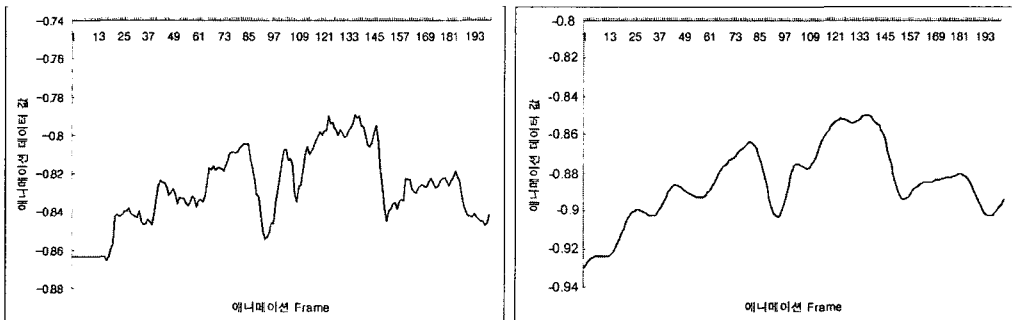


그림 1 원본 애니메이션 데이터와 가우시안 필터링을 사용하여 잡음을 제거한 결과

이션 데이터로부터 중요요소분석법을 통하여 얼굴 모델의 정점을 얼굴 구성 요소의 집합으로 분류한 결과이다. 그림 10(a)에 주어진 3차원 얼굴 애니메이션 데이터에 대한 실험 결과 그림 2(a)의 얼굴 모델에서는 41개의 얼굴 구성 요소를 얻었으며, 그림 11(a)의 데이터에 대한 실험결과 그림 2(b)의 얼굴 모델에서는 24개의 얼굴 구성 요소를 얻을 수 있었다. 중요요소분석법에 기반한 얼굴 구성 요소의 분류에서는 동일한 얼굴 모델이라도 입력되는 애니메이션 데이터가 다르면 분류되는 얼굴 구성 요소의 개수가 다를 수 있다.

3.3 얼굴 표정의 추출

일단 얼굴 애니메이션 데이터의 분석을 통해서 얼굴 구성 요소들이 생성되면, 얼굴 표정 애니메이션 데이터로부터 얼굴 표정의 시작과 끝 프레임 정보를 바탕으로 얼굴 표정들을 분할한다. 모든 얼굴 구성 요소에 대하여 3.2절에서 구한 주방향을 기준으로 얼굴 구성 요소의 애니메이션 그래프 G 를 재구성한다(그림 3 참조). 그림 3에서 세로축은 해당 얼굴 구성 요소가 주방향을 따라서 움직인 양을 나타내며, 가로축은 시간 축을 나타낸다. 이때 세로축의 값 0은 해당 얼굴 구성 요소가 중립 표정에 있을 때를 나타낸다.

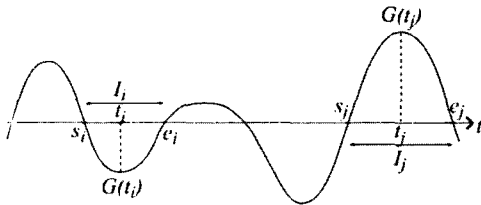


그림 3 주방향을 따라서 재구성한 애니메이션 그래프의 예

먼저 프레임 번호에 대한 함수로 입력되는 얼굴 애니메이션 그래프 G 를 0보다 작지 않은, 혹은 0보다 크지 않은 세그먼트 g 의 집합으로 다음과 같이 분할한다.

$$G = \cup g_i, \quad 1 \leq i \leq r$$

이때 g_i 는 시간의 흐름을 따라서 애니메이션 그래프의 i 번째 세그먼트를 나타내며, 해당 얼굴 표정 애니메이션에 r 개의 세그먼트가 존재한다고 가정한다. 세그먼트 g_i 의 시작 프레임을 s_i , 종료 프레임을 e_i 라 하고, 프레임 간격 I_i 을 $[s_i, e_i]$ 와 같이 정의한다. 세그먼트 g_i 에 대해서 $G(t) \geq 0$ (혹은 $G(t) \leq 0$), $\forall t \in I_i$ 이면 g_i 를 양수(혹은 음수)라 부른다. 세그먼트 g_i 가 양수(혹은 음수)이면 프레임 간격 I_i 내에서 최대값(혹은 최소값)을 가지며, 최대(혹은 최소)가 되는 프레임 번호를 t_i 라 한다. 결과적으로 얼굴 표정의 움직임에 상응하는 세그먼트

트 g_i 에 대하여 $G(t_i)$ 는 i 번째 얼굴 표정(Expression)이라고 할 수 있다.

얼굴 애니메이션 그래프로부터 얼굴 표정을 추출하기 위하여 그래프를 3차 허미트 곡선으로 보간한다. 3차 허미트 곡선으로 보간된 애니메이션 그래프 G 에 대한 해를 구하면 모든 세그먼트에 대한 시작 및 종료 프레임을 구할 수 있고, 이는 얼굴 표정의 움직임에 해당한다. 또한 임의의 $t \in (s_i, t_i)$ 에 대하여 $G(t)$ 의 부호를 검사하여 해당 세그먼트가 양수인지 혹은 음수인지를 결정할 수 있다. 미분 그래프 $\frac{d}{dt} G(t)$ 에 대한 해를 구하면 그래프 G 의 최대값 혹은 최소값을 갖는 프레임을 구할 수 있다. 그림 4에서는 하나의 얼굴 구성 요소에 대하여 이상과 같은 방법으로 얼굴 표정을 추출한 예를 제시하였다.

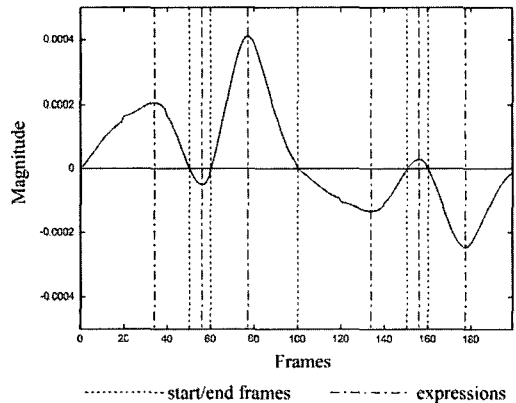


그림 4 애니메이션 그래프로부터 표정을 추출한 결과

3.4 기대효과 탐색 및 최적 기대효과의 결정

기대효과를 추가할 표정이 사용자의 선택에 의해서 결정되면, 적절한 기대효과를 찾기 위하여 주어진 얼굴 표정 애니메이션 데이터를 탐색한다. 주어진 얼굴 애니메이션 데이터가 충분히 길고 다양한 표정이 존재한다면, 적절한 기대효과를 찾을 가능성이 높아진다.

기대효과를 추가할 얼굴 표정으로 g_j 가 선택되었다고 가정하고, 해당 표정에 대한 기대효과 표정을 g_j 라 하자. 이때 기대효과의 일반적 특성에 의하여 g_j 의 부호(양수 혹은 음수)가 먼저 고려되어야 한다. 만일 g_j 가 양수라면 g_j 는 반드시 음수여야 한다. 반대로 g_j 가 음수라면 g_j 는 반드시 양수여야 한다. 주어진 두 상수 w_1, w_2 에 대하여 기대효과 표정의 인덱스 j 는 다음과 같은 식으로 계산된다.

$$j = \arg \min_{1 \leq k \leq s, k \neq i} \{w_1 \Delta G_{i,k} + w_2 \Delta I_{i,k}\} \quad (1)$$

여기서,

$$\begin{aligned}\Delta G_{i,k} &= ||G(t_i) - |G(t_k)||, \\ \Delta I_{i,k} &= |I_i - I_k|,\end{aligned}$$

와 같이 정의되며, \arg 함수는 결과에 대한 인덱스를 반환하는 함수로 정의한다. 이때, 기대효과 표정 움직임의 방향이 주어진 표정 움직임의 반대가 되게 하기 위하여 $G(t_i)G(t_k) < 0$ 의 조건을 만족해야 한다. 식 (1)은 주어진 얼굴 표정과 기대효과 표정 사이의 움직임의 크기와 더불어 프레임 간격의 유사성을 동시에 고려하고 있으며, 두 가지 고려사항에 대한 가중치는 w_1, w_2 의 상대적인 크기에 의해서 결정된다.

식 (1)은 모든 얼굴 구성 요소별로 계산되며, 따라서 주어진 하나의 표정에 대하여 모든 얼굴 구성 요소를 고려하면 서로 다른 기대효과를 서로 다른 프레임의 위치에서 얻게될 가능성이 있다. 즉, 얼굴 모델에서 3.2절에서 주어진 방법으로 분류한 얼굴 구성 요소의 개수를 h 라 하면, 임의로 주어진 얼굴 표정 $G(t_i)$ 에 대해서 $G(t_{i,1}), G(t_{i,2}), \dots, G(t_{i,h})$ 등 최대 h 개의 서로 다른 기대효과 표정을 얻을 수 있다. 만일 이러한 얼굴 구성 요소별로 얻어진 기대효과 표정의 움직임을 그대로 병합하여 애니메이션을 만들게 되면, 사람의 얼굴 등에서는 나올 수 없는 표정이 나타나게 되거나 혹은 얼굴 모델의 위상정보가 깨지는 현상이 보이게 된다(그림 5 참조).

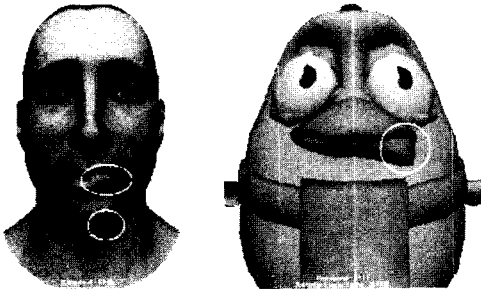


그림 5 얼굴 구성 요소별 단순 병합에 의하여 모델의 위상이 깨지는 경우

따라서 h 개의 얼굴 구성 요소를 갖는 얼굴 애니메이션 데이터에 대해서 얻어진 기대효과 표정 중에서 최적의 기대효과 $G(t_{i,l}), 1 \leq l \leq h$ 를 선택하는 방법이 필요하다. 본 논문에서는 애니메이션을 시청하는 관객의 인지적 특징에 기반하여 최적의 기대효과를 선택하는 방법을 제시한다. 먼저 관객은 일반적으로 작은 움직임 보다 큰 움직임에 대해서 관심을 집중하는 경향이 있다. 따라서 기대효과를 적용했을 때 얼굴 표정의 차이가 최대가 되는 방향으로 최적의 기대효과를 다음과 같은 계산식에 의하여 선택할 수 있다.

$$l = \arg \max_{1 \leq k \leq h} |G(t_i) - G(t_{i,k})| \quad (2)$$

식 (2)에 의하여 계산된 인덱스를 가지고 얻어진 표정 $G(t_{i,l})$ 을 주어진 표정 $G(t_i)$ 의 기대효과 표정으로 적용하면 얼굴 모델의 위상정보가 깨어지는 현상은 막을 수 있다. 인덱스 l 은 얼굴 구성 요소와 관련되어 있기 때문에, 식 (2)에 의하여 얻어진 기대효과 표정은 l 의 인덱스를 갖는 얼굴 구성 요소에 대해서는 최대의 기대효과 표정을 가지게 되지만, 그 외의 얼굴 구성 요소에 대해서는 기대효과를 갖는다고 보장할 수 없다.

따라서 h 개의 얼굴 구성 요소를 모두 고려하여 가장 많은 수의 얼굴 구성 요소가 기대효과로 선택한 표정을 최적의 기대효과로 선택하여 이러한 문제를 해결할 수 있다. 그러나 단지 어떤 일정한 구간에서 가장 많은 수의 얼굴 구성 요소에 대한 기대효과가 발견된다고 해서, 단순히 이 구간의 표정을 기대효과로 선택하면 지나치게 작은 움직임에 대해서만 기대효과를 얻게되고, 큰 움직임에 대한 기대효과를 얻지 못하는 잘못된 결과를 얻을 수도 있다. 따라서 기대효과와 주어진 표정 사이의 차이를 최대화하면서, 동시에 최대 개수의 얼굴 구성 요소들이 기대효과를 갖는 프레임 구간의 표정을 고려하여 최적의 기대효과를 선택하여야 한다. 함수 $F_s(t)$ 를 주어진 작은 상수 ϵ 에 대하여 다음과 같이 정의하면,

$$F_s(t) = \begin{cases} 1, & \text{if } s - \epsilon < t < s + \epsilon \\ 0, & \text{otherwise} \end{cases}$$

최적의 기대효과는 다음과 같은 식에 의하여 계산된다.

$$l = \arg \max_{1 \leq k \leq h} \left(\sum_{q=1}^4 F_{t_{i,q}}(t_{i,q}) \Delta G_{i,j_q} \right) \quad (3)$$

여기서 $\Delta G_{i,j_q} = ||G(t_i) - |G(t_{i,q})||$ 와 같다. 식 (3)은 얼굴 모델의 위상 정보를 유지하면서, 가장 많은 수의 얼굴 구성 요소가 적절한 기대효과 표정을 가지고 동시에 표정의 차이가 최대화되도록 최적의 기대효과 표정을 선택한다.

3.5 기대효과의 합성

일단 주어진 얼굴 표정에 대한 최적의 기대효과 표정을 얻고나면, 기대효과 표정을 주어진 얼굴 표정의 앞에 삽입한다. 그러나 단순히 기대효과를 삽입하면, 애니메이션 그래프의 연속성이 깨어지고 아울러 애니메이션의 재생 시간이 늘어나게 된다(그림 6의 (a), (b) 참조). 따라서 애니메이션 그래프의 연속성과 재생 시간을 유지하면서 기대효과를 삽입하여 합성(Blending)하는 방법을 사용해야 한다(그림 6의 (c) 참조). 이때 기대효과의 일반적인 특성에 따라서 기대효과 표정의 움직임은 주어진 얼굴 표정에 비해서 작고, 느리게 조절하여 합성한다. 결과적으로 초기에 주어진 얼굴의 표정이 상대적으로 원본에 비해서 빠르고 크게 느껴지게 되며 이는 기대효과에 의한 움직임의 특징에 부합하는 결과이다. 본

논문에서는 실험적인 결과에 따라서 기대효과와 움직임의 크기와 속도를 원본으로 입력되는 얼굴 표정의 1/3에 해당하는 양을 사용하여 합성하였다.

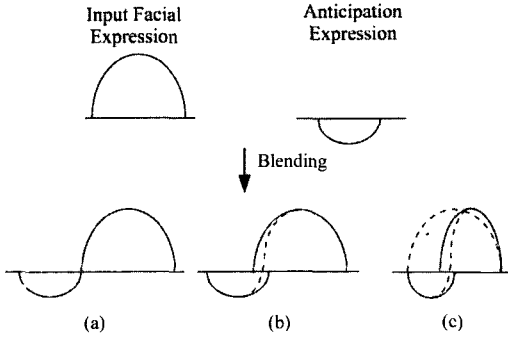


그림 6 입력 얼굴 표정과 기대효과 표정을 합성하는 예

4. 실험 결과 및 고찰

본 논문에서 제시된 방법에 따라서 얼굴 모션 캡처를 통해서 얻어진 200 프레임 길이의 얼굴 표정 애니메이션 데이터를 입력으로 사용하였다(그림 7참조).

입력 애니메이션 데이터에 대해서 3.2절에서 설명한 방법을 적용한 결과 얼굴 모델의 모든 정점이 41개의 얼굴 구성 요소로 분류되었다. 30번째 프레임에 존재하는 얼굴 표정(그림 7(a))을 선택하여, 3.4절의 식 (1)에서 제시된 방법에 따라서 얼굴 구성 요소별로 선택한 기대효과 표정은 그림 8과 같다.

이 중에서 3.4절의 식 (3)에 의하여 선택되는 최적의 기대효과 표정은 그림 8의 (a)로서 입력 애니메이션 데이터에서 80번째 프레임의 표정(그림 7(b))에 해당한다. 여기서 그림 8(b)와 그림 8(c)에 비해서 그림 8(a)가 더 적은 수의 얼굴 구성 요소가 기대효과 표정으로 선택되었음에도 불구하고 최적의 기대효과 표정으로 선택되었

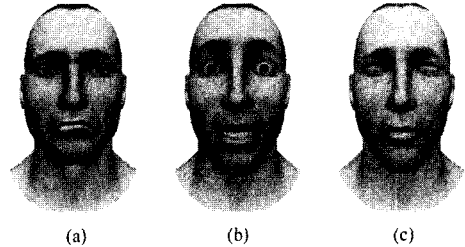
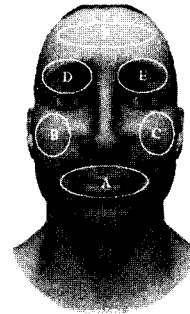


그림 8 주어진 표정에 대한 얼굴 구성 요소별 기대효과 : (a) 얼굴 요소 F에 대한 기대효과, (b) 얼굴 구성 요소 B와 C에 대한 기대효과, (c) 얼굴 구성 요소 A, D, E에 대한 기대효과

다. 이는 식 (3)이 두 표정 사이의 차이를 최대화하는 것을 동시에 고려하고 있기 때문에 발생하는 현상이다.

만일 주어진 얼굴 표정 애니메이션에서 130번째 프레임의 표정(그림 7(c))을 선택하여, 각 얼굴 구성 요소별로 기대효과를 식 (1)에 의하여 구하면 그림 9와 같다. 이때 최적의 기대효과 표정으로는 그림 9(c)가 선택되며 이는 입력 애니메이션 데이터에서 180번째 프레임의 표정(그림 7(d))에 해당된다.

그림 10과 그림 11에서 각각 얼굴 모션 캡처에 의하여 얻어진 얼굴 표정 애니메이션과, 키프레임 애니메이션 작업을 통하여 얻어진 얼굴 표정 애니메이션 데이터

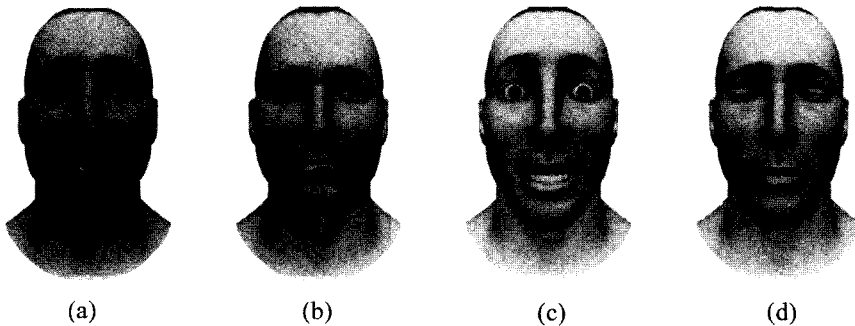


그림 7 기대효과가 없는 얼굴 표정 애니메이션 데이터 : (a) 30번째 프레임의 표정, (b) 80번째 프레임의 표정, (c) 130번째 프레임의 표정, (d) 180번째 프레임의 표정.

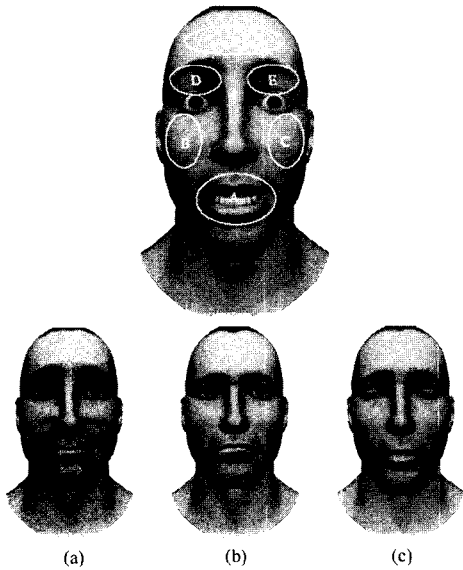
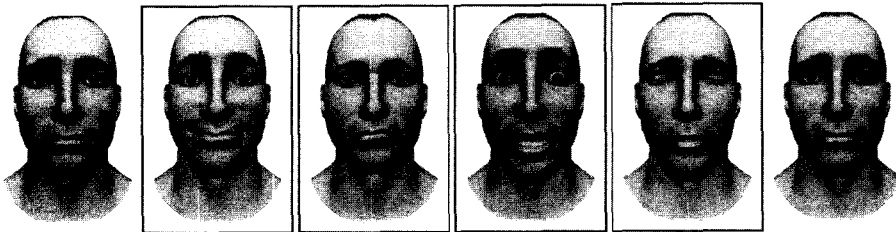


그림 9 주어진 표정에 대한 얼굴 구성 요소별 기대효과
 : (a) 얼굴 요소 B와 C에 대한 기대효과, (b) 얼굴 구성 요소 A에 대한 기대효과, (c) 얼굴 구성 요소 D, E, F에 대한 기대효과

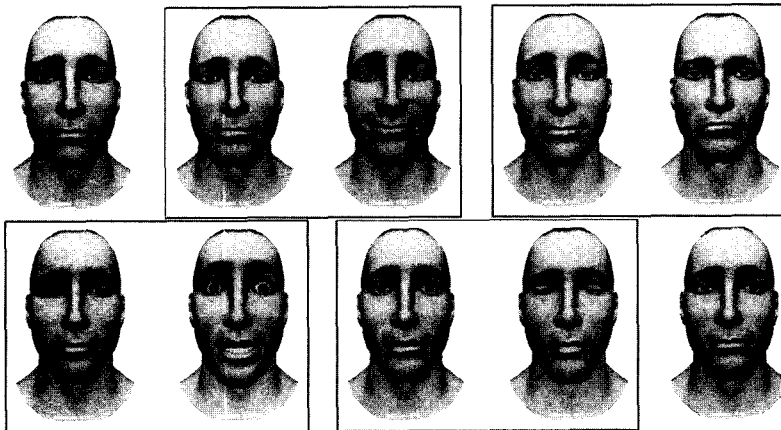
에 대해서 매 표정에 대한 기대효과를 적용한 결과를 제시하였다. 기대효과 표정의 크기가 합성과정에서 상대적으로 줄어든 결과를 확인할 수 있다.

5. 결론

본 논문에서는 주어진 얼굴 표정 애니메이션 데이터로부터, 기대효과를 자동으로 추가하는 방법을 제시하였다. 얼굴 모델에 존재하는 모든 정점에 대하여 입력 애니메이션 데이터로부터 중요요소분석법(Principal Component Analysis)을 통하여 모든 정점을 얼굴 구성 요소의 집합으로 분류하였다. 각 얼굴 구성 요소내에 존재하는 모든 정점들의 움직임은 비슷한 주방향(Principal Direction)을 갖는다. 기대효과를 추가할 표정이 주어지면, 기대효과 특징을 사용하여 모든 얼굴 구성 요소별로 반대방향의 움직임을 갖는 표정들을 추출한다. 이들 얼굴 구성 요소별 기대효과 표정들 중에서, 얼굴 모델의 위상정보를 유지하면서 주어진 표정과 기대효과 사이의 움직임의 차이를 최대화하고, 동시에 최대 개수의 얼굴 구성 요소들이 적절한 기대효과 표정을 가질 수 있는 최적의 기대효과 표정을 선택하였다. 최적의 기대효과를



(a)



(b)

그림 10 (a) 얼굴 모션 캡처에 의하여 입력된 얼굴 표정 애니메이션 데이터, (b) 기대효과 표정을 추가하여 얼굴 표정 애니메이션을 완성한 예

원본 입력 표정의 앞에 위치시키고, 애니메이션 데이터의 연속성과 재생 시간을 보존하면서 합성하였다.

기대효과는 전통적인 2차원 애니메이션의 제작 원리 중의 하나로, 동작을 보다 풍부하게 하고 사실적으로 표현하는데 적합한 원리로 알려져 있다. 본 논문에서는 이러한 기대효과를 3차원 얼굴 표정 애니메이션의 제작에 적용하여 보다 풍부하고, 사실적인 3차원 얼굴 애니메이션을 얻는 새로운 방법을 제시하였다. 본 논문에서 제시된 방법을 사용하면, 숙련되지 않은 애니메이션 제작자가 얼굴 표정을 선택하는 간단한 상호작용만으로도 기대효과가 적용된 3차원 얼굴 표정 애니메이션을 손쉽게 제작할 수 있다.

참 고 문 헌

- [1] Parke, F. I. and Waters, K., *Computer Facial Animation*, A. K. Peters, 1996.
- [2] Agarwala, A., "Snaketoonz : a semiautomatic approach to creating cel animation from video," *Proceedings of Non-Photorealistic Animation and Rendering*, pp. 139-148, 2002.
- [3] Chenny, S., Pingel, M., Iverson, R., and Szymanski, M., "Simulating cartoon style animation," *Proceedings of Non-Photorealistic Animation and Rendering*, pp. 133-138, 2002.
- [4] Korf, A., "Computer aided inbetween," *Proceedings of Non-Photorealistic Animation and Rendering*, pp. 125-132, 2002.
- [5] Lasseter, J., "Principles of traditional animation applied to 3D computer graphics," *Proceedings of SIGGRAPH '87*, pp. 35-44, 1987.
- [6] Williams, R., *The Animator's Survival Kit*, Faber and Faber, 2001.
- [7] Parke, F. I., "Computer generated animation of faces," *Proceedings of ACM annual conference*, pp. 451-457, 1972.
- [8] Parke, F. I., *A Parametric Model for Human Faces*, Ph.D. Thesis, University of Utah, 1974.
- [9] Noh, J. Y., and Neumann, U., "A survey of facial modeling and animation techniques," *Technical Report 99-705*, University of Southern California, 1999.
- [10] Opalach, A., and Maddoc, S., "Disney effects using implicit surfaces," *Proceedings of 5th Eurographics Workshop on Animation and Simulation*, 1994.
- [11] Rademacher, P., "View-dependent geometry," *Proceedings of SIGGRAPH '99*, pp. 439-446, 1999.
- [12] Bregler, C., Loeb, L., Chuang, E., and Deshpande, H., "Turning to the masters: Motion capturing cartoons," *ACM Transactions on Graphics(Proceedings of SIGGRAPH 2002)*, Vol. 22, No. 3, pp. 399-407, 2002.
- [13] Bruderlin, A., and Williams, L., "Motion signal processing," *Proceedings of SIGGRAPH '97*, pp. 97-104, 1997.
- [14] Lee, Y. C., Terzopoulos, D., and Waters, K., "Constructing physics-based facial models of individuals," *Proceedings of Graphics Interface*, pp. 1-8, 1993.
- [15] Lee, Y. C., Terzopoulos, D., and Waters, K., "Realistic modeling for facial animation," *Proceedings of SIGGRAPH '95*, pp. 55-62, 1995.
- [16] Kuo, C. J., Huang, R. S., and Lin, T. G., "Synthesizing lateral face from frontal facial image using anthropometric estimation," *Proceedings of International Conference on Image Processing*, Vol. 1, pp. 133-136, 1997.

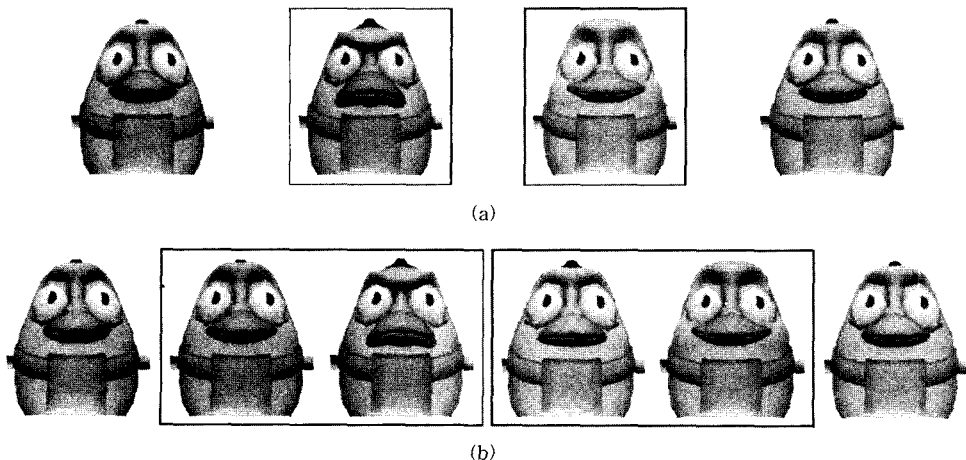


그림 11 (a) 키프레임 애니메이션 작업에 의하여 입력된 얼굴 표정 애니메이션 데이터, (b) 기대효과 표정을 추가하여 얼굴 표정 애니메이션을 완성한 예

[17] Jolliffe, I. T., Principal Component Analysis, Springer, 1986.



최 정 주

1990년 한국과학기술원 전산학과 공학사
1992년 포항공과대학교 컴퓨터공학과 공학석사. 1997년 포항공과대학교 컴퓨터공학과 공학박사. 1997년~2000년 한국전자통신연구원 선임연구원. 2000년~2002년 (주)디지털아리아 선임연구원. 2002년~현재 아주대학교 미디어학부 조교수. 관심분야는 컴퓨터 그래픽스, 기하 모델링, 애니메이션



김 동 선

2002년 아주대학교 미디어학부 공학사
2004년 아주대학교 미디어학과 공학석사
2004년~현재 (주)엔젤 연구원. 관심분야는 컴퓨터 그래픽스, 애니메이션



이 인 권

1989년 연세대 전산과학과 학사. 1992년 포항공대 전자계산학과 석사. 1997년 포항공대 전자계산학과 박사. 1997년~1999년 비엔나 공대 연구원. 1999년~2001년 포항공대 정보통신연구소 연구원
2001년~2003년 아주대 미디어학부 조교수. 2003년~현재 연세대 컴퓨터과학과 조교수. 관심분야는 컴퓨터 그래픽스, 기하 모델링, 컴퓨터 음악