

얼굴 구성 정보 기반의 실시간 애니메이션을 위한 매핑 알고리즘

이정훈 이찬 이필규
인하대학교 전자계산공학과
g2001689 @inhavision.inha.ac.kr,

A Mapping Algorithm for Real Time Animation Based on Facial Features

Yi Jung-Hoon, Lee Chan, Rhee Phill-Kyu
Dept of Computer Science & Engineering, Inha University

요약

본 논문에서는 가상 인터페이스로서 범용적으로 사용할 수 있는 실시간 비전 기반 얼굴 애니메이션을 제안한다. 이를 위해서 실시간으로 얼굴 구성요소를 추출하고, 이를 정량화 하였다. 정량화된 값과 자연스럽게 매핑하기 위해서 정함함수를 통해 3차원 모델에 맵핑하기 위한 방안을 제안한다. 일반적으로 3차원 애니메이션을 수행할 경우, 기본 모델을 중심으로 특정한 사용자만을 위주로 수행되나, 본 논문에서는 임의의 일반 사용자를 위한 3차원 애니메이션을 수행하였다. 여러 사용자에 대하여 얼굴 구성요소 추출을 이용한 3차원 얼굴 애니메이션 동작에 대하여 실험하였으며, 실험결과 여러 사용자 얼굴 애니메이션 동작에 대하여 만족할 만한 성능을 보였다.

1. 서론

컴퓨터 하드웨어 기술의 발전과 더불어 소프트웨어 기술의 발달도 매우 눈부시다. 현재의 하드웨어 시스템은 환상적인 그래픽과 사운드를 제공해 주고 있으며, 통신망과의 결합을 통하여 인간의 즐거움을 충족시킬 수 있는 완벽한 매개체가 될 수 있다. 이러한 기술의 발전은 가상현실의 전반적인 기술을 실제의 현실에 가깝게 끌어 올렸다. 흔히 볼 수 있는 게임에서부터 방송의 가상 스튜디오까지, 컴퓨터의 하드웨어와 소프트웨어가 만들어 내는 또 하나의 세계가 우리에게 더욱 친밀감 있게 다가오고 있다.

이와 같은 가상현실 기술의 발전은 가상의 공간을 형성하는 것 뿐만 아니라 가상의 캐릭터를 창조해 낼 수도 있다. 아바타(AVATA)로 일컬어지는 가상공간의 캐릭터는 가상현실의 실재감을 더욱 높일 수 있을 뿐 아니라, 컴퓨터와 인간과의 더욱 진보된 인터페이스 역할을 수행할 수 있다.

본 논문에서는 실재감을 갖는 얼굴 애니메이션이 실시간 트래킹과 결합하여 자연스럽게 이루어질 수 있도록 얼굴 영상을 3차원 모델에 매핑하기 위한 방안을 제시한다. 이를 위해서 실시간으로 얼굴 구성

요소를 추출하고, 이를 정량화 하였다. 또, 추출된 결과와 자연스럽게 매핑하기 위해서 정함함수를 통한 3차원 얼굴 모델 애니메이션으로의 매핑방안을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 얼굴 구성정보 추출 및 애니메이션에 대한 기존 연구에 대하여 기술한다. 3장에서는 얼굴 구성요소 검출을 위한 전처리 과정과 템플릿을 이용한 구성요소 검출 방법을 기술한다. 4장에서는 특징벡터 기반의 정함함수에 대해 기술하며, 5장에서는 실험 결과를 설명한다. 마지막으로 6장에서는 결론 및 향후 연구 방향에 대하여 기술한다.

2. 기존연구

3차원 모델에 의한 얼굴 애니메이션을 수행할 경우, 애니메이션의 움직임에 위한 파라미터 조절을 위해서는 얼굴 정보의 검출[1,2,3]과 정량화가 중요한 문제이다. 일반적으로, 정면 얼굴 움직임에 따른 3차원 애니메이션을 위해 제안된 기존의 방법들은 주로 얼굴에 마커(marker)와 같은 부착물을 사용하거나, 움직임을 측정하는 센서의 사용, 혹은 얼굴 고

정을 위한 장비등을 갖추고, 이러한 장비들로부터 얼굴 구성요소를 측정하여 애니메이션을 수행하였다. 그러나 이는 사용자에게 불편함을 주게 되며, 아울러 장비가 고가라는 단점 또한 갖고 있다.

따라서 본 논문에서는 부가적인 장비 없이 수행될 수 있는 얼굴 검출방법과 이를 바탕으로 하는 3차원 애니메이션으로의 매핑함수를 제안한다.

3. 얼굴 구성요소 검출

3.1 얼굴 구성요소 검출을 위한 전처리

얼굴 구성요소의 검출은 마우스를 이용하여 얼굴의 위치를 수동적으로 초기화 하는 것에서 시작한다. 얼굴영역으로 선택된 영역에는 기본적인 전처리 과정이 선행된다. 즉, 히스토그램 균일화[4], 영상 이진화를 거쳐 연결 요소 분할(Connected Component Labelling)[5]의 과정을 거친다.

3.2 특징 기반 템플릿 모델

영상의 전처리 과정을 거친 영역에 템플릿 모델을 이용하여 얼굴 구성요소의 후보 영역을 탐색하고 이를 이용하여 구성요소를 추출한다.

얼굴의 구성요소 검출을 위해 템플릿 모델을 적용하는 이유는 빠르고 정확한 구성 요소의 검출이 가능하기 때문이다. 본 논문에서 적용하는 원형 템플릿 모델[6]은 얼굴영역이 원형이라는 특징에 기초한다. 즉, 그림 1에서 볼 수 있듯이, 원형 템플릿 모델은 얼굴 후보 영역의 중심을 원점으로 동일한 넓이를 갖는 동심원으로 구성되어 있다. 이의 적용 과정을 보면 다음과 같다. 우선, 원형 템플릿 모델의 비례적 관계를 이용하여 후보 영역의 추정한다. 눈의 경우 후보 영역은 그림 3에서 중심선 L1을 기준으로 위쪽의 7번째와 8번째 원의 각각 좌우측으로 3번 구역을 중심으로 한다. 입의 경우 동일한 기준선 L1을 중심으로 아래쪽 7번째와 8번째 원의 중심 영역에 후보 영역을 둔다.

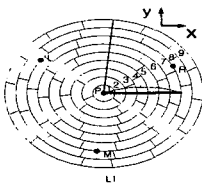


그림 1. 원형 템플릿 모델

같은 방법으로 영역내에서 가능한 후보영역을 모

두 찾아낸 다음, 연결요소 분할(Connected Components Labelling) 영역과 얼굴의 구성요소로 추정된 영역을 픽셀단위로 이동시키면서 아래 식(1)과 같이 그 값이 가장 적은 연결요소 분할 영역을 눈과 입의 영역으로 확정한다.

$$d = \frac{\sum_{i=1}^N L(i) - C(i)}{N} \quad \text{식 (1)}$$

C(i) : Circle Template에서 추정된 영역의 픽셀 값

L(i) : Labelling 영역 내부의 픽셀 값

N : 영역 내부의 넓이

그림 2는 입력된 영상을 전처리하고, 이에 원형 템플릿 모델을 처리한 결과를 보인 것이다.



그림 2. 원형 템플릿 모델의 매칭 과정 및 결과

4. 특징벡터 기반의 애니메이션 정합

본 장에서는 3장에서 추출된 얼굴의 구성정보를 기반으로 정면 얼굴의 움직임과 3차원 애니메이션 모델과의 정합시키는 방법을 기술한다.

4.1 3차원 얼굴 모델

본 논문에서 사용된 3차원 모델은 다각형(Polygon)으로 이루어진 파라미터 기본 모델[7]이다. 287개의 토폴로지(Topology)를 움직여 모델의 모양을 변화시키며, 모델의 간략화를 위해 좌우가 대칭적으로 구성되었다.

4.2 얼굴 구성요소의 정량화

카메라의 입력 영상만으로 3차원 모델의 얼굴과 사용자의 얼굴을 정합 시키기 위해서는 추출된 구성요소의 정량화가 선행되어야 한다. 이를 위해, 앞서 언급한 전처리 과정과 원형 템플릿을 이용하여 얼굴 구성정보의 크기값을 얻어내고 이를 정규화 하였다. 문제는 특정 사용자만을 위한 것이 아니므로 해당 사용자의 구성 정보값의 변위를 알아 내기가 힘들다는데 있다. 따라서 본 논문에서는 처음 입력되는 사

용자의 얼굴이 가장 기본적인 표정이라는 것을 전제로 하고, 이후, 계속적으로 추출되어지는 값들은 벡터공간에 정렬한다. 이 벡터 공간으로부터 파라미터를 계산하고 생성해 내는 정합함수를 구성하였다.

4.3 파라미터 정합

얼굴 구성정보의 정량화는 파라미터를 통해 애니메이션의 변화로 나타내어 줄 수 있다. 본 논문에서 사용하는 애니메이션 파라미터는 움직이는 토폴로지의 변위값, 해당 토폴로지의 변위값, 그리고 지연값으로 구성되어 있다. 이들 파라미터는 다음과 같은 방법으로 추출된다.

우선, 구성정보를 저장할 수 있는 벡터공간을 구성한다. 여기에 연속적인 구성요소 정보중에서 첫 번째로 정량화된 값을 해당 사용자 얼굴 구성 정보의 기준값으로 하고, 벡터공간의 중심에 위치시킨다. 이 기준값을 중심으로 가장 떨어진 벡터공간에 위치한 값을 토폴로지 변위의 최대/최소값으로 설정하며, 이후 계속적으로 추출되어 지는 값들은 기준값과의 비교를 통해 벡터 공간에 저장되고 정렬된다. 따라서, 토폴로지 변위의 최대/최소값에 해당하는 얼굴 구성정보의 정량화된 값은 계속적인 구성정보의 추출로 다양해진다. 이는 표현할 수 있는 토폴로지의 단계를 많아지게 하며, 이에 따라 3차원 모델의 얼굴 변화는 더욱 다양할 수 있다.

한편, 형성된 벡터 공간으로부터 현재 입력된 구성 정보가 위치하게 될 벡터공간의 위치를 계산하고, 동시에 해당 벡터공간의 변위에 상응하는 토폴로지의 변위값을 계산한다. 여기에, 해당 토폴로지의 변위와 지연시간을 덧붙여 파라미터 생성을 완료하고 이를 애니메이션을 수행하는 함수로 전송한다. 그림 3은 앞에서 설명한 정합함수의 알고리즘을 개략적으로 나타낸 것이다.

여기서 고려해야 할 사항이 있다. 얼굴 구성정보의 추출이 연속적이지 않을 경우, 즉, 구성정보의 추출에 불연속적인 시점이 발생하면 이를 파라미터에 어떻게 반영할 것인가 하는 문제이다. 제안된 알고리즘에서는 구성정보를 찾지 못한 불연속 점 양단의 파라미터를 계산하고, 보간법을 이용하여 이들 양쪽을 이어주는 파라미터를 추가로 생성하였다.

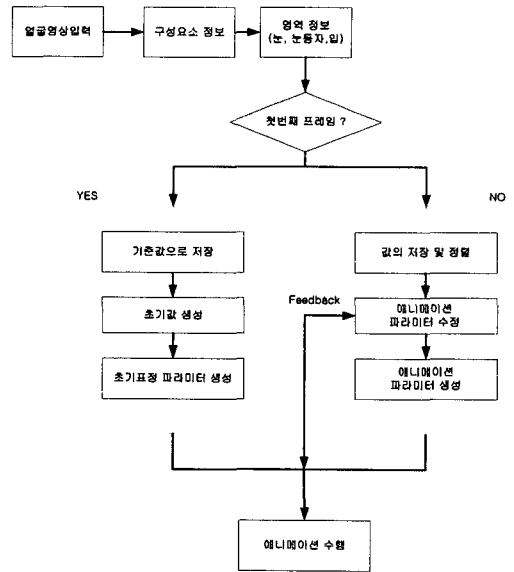


그림 3. 정합함수의 알고리즘 개념도

5. 실험 결과

본 논문에서 제안된 시스템의 실험영상은 320 × 240의 흑백 영상으로, 초당 20~30 프레임 비율로 피 실험자의 얼굴영상을 입력받았다. 피 실험자가 목적에 부합하는 영상을 생성하는 데에 제한 시간을 정하지 않았으며, 영상의 입력 시에 일반적인 조도의 실내조명을 배경으로, 무표정 상태에서 시작하여 눈, 입의 연속적인 움직임을 생성하도록 요구받았다. 피 실험자는 주로 20대 초반부터 30대 초반까지의 남녀로 구성되어 있다.

얼굴 구성요소의 정보와 애니메이션 파라미터 동작의 정확도를 확인하기 위해서 얼굴 구성요소를 정량화한 값과 이로 인해 생성된 파라미터의 변위값을 각각 정규화 한 후 MSE(Mean Square Error)를 이용하여 에러율을 계산하였다. 그림 4에서 추출된 구성정보와 그에 따른 파라미터값을 정규화하고, MSE를 이용하여 에러율을 검증한 결과의 한 예를 나타냈다. 그림에서 ●로 표시되어지는 선은 애니메이션 파라미터의 정규화된 값이며, ■는 얼굴 구성 정보를 정량화한 값을 다시 정규화한 값이다. ▲는 애니메이션 파라미터와 구성정보의 정규화된 값들의 MSE값을 나타낸다. 그림에서 보듯 생성되어 지는 애니메이션 파라미터는 전반적으로 얼굴 구성 정보의 변화량을 잘 추적해 가고 있다. 또한, 7, 9, 12번

째 프레임 등, 얼굴 구성정보의 추출이 연속적이지 않았을 경우의 애니메이션 파라미터들을 보면 전체적인 애니메이션의 연속성을 잃지 않고 있다는 것을 알 수 있다.

구성정보의 추출에 있어 이와 같은 불연속점이 발생하는 것은 템플릿 모델이 갖는 특성 때문으로 생각된다. 템플릿 모델을 이용한 구성정보의 추출방안은 일반적으로 빠른 검출속도와 정확성을 갖고 있다. 그러나 얼굴 구성이 갖는 위치적 특징을 비례관계를 이용하여 접근한다는 특성 때문에, 조명이나 얼굴 위치 변화 등의 급격한 변화가 발생하는 경우, 변화된 영역내의 비례관계를 설정하지 못하고 기각되어 버리는 것이다.

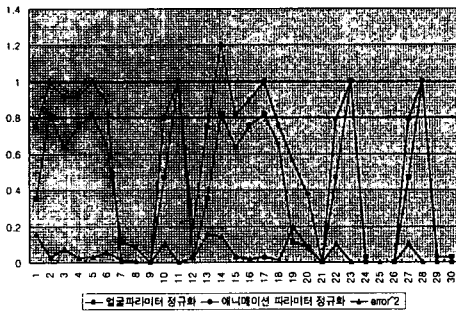


그림 4. 구성정보의 추출과 애니메이션 동작정도

그림 5와 6은 얼굴 구성 정보의 검출에 의한 3차원 애니메이션의 결과를 나타낸다.

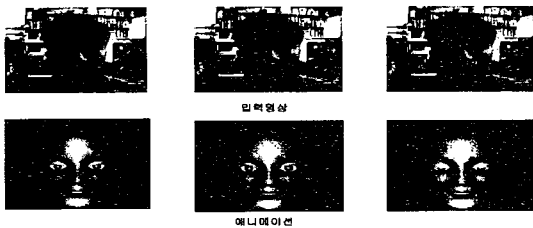


그림 5. 눈영역 검출에 의한 3차원 애니메이션 결과



그림 6. 입 영역 검출에 의한 3차원 애니메이션 결과

6. 결론 및 향후 연구방향

본 논문에서는 가상 인터페이스로서 범용으로 사용할 수 있는 실시간 비전기반 얼굴 애니메이션을 제안하였다. 이를 위해서 실시간으로 얼굴 구성요소를 추출하고, 정량화 하였다. 또한 정합함수를 이용하여 입력된 얼굴 정면 영상과 동기화 된 3차원 얼굴 모델의 애니메이션을 수행하였다. 실험결과 MSE 검증률 통한 전반적인 시스템 정확도는 91.01%이고, MSE값은 0.0896으로 나타났다. 제안된 시스템은 향후, 이동형 가상 인터페이스에 적용이 가능하며, 이외에도 인간의 감정 정보를 필요로 하는 의학이나 심리학, 인터페이스로서의 아바타 표정 생성 등 다양한 분야에 적용될 수 있다.

향후 연구 방향은, 수동적인 얼굴영역의 추출을 수정하여 이를 자동화하고, 입력영상의 조명과 환경 변화에도 잘 적용될 수 있도록 전처리 부분의 개선이 요구된다. 마지막으로 템플릿 모델의 수정과 국부적인 얼굴 구성 정보의 추가적인 이용으로 기각률을 줄이는 노력이 필요하다.

참고문헌

- [1] 김용주, 한태우, 이주호, 양현승 " MPEG-4 파라미터 기반의 계층구조를 갖는 얼굴 모델 생성 툴의 설계 및 구현 ", HCI 99~2000 pp 540-545, 2000
- [2] Marc Escher, Igor Pandzic, Nadia Magnenat Thalmann, "Facial Deformation for MPEG-4", courses notes of SIGGRAPH' 98, 1998
- [3] G. Abrantes and F. Pereira. "MPEG-4 facial animation technology: Survey, implementation, and Result", IEEE Trans. on Circuit and System for VideoTech, Vol9, pp290-305, Mar, 1997
- [4] H.Gharavi and M. Mills, "Block-Matching Motion Estimation," IEEE Trans. on Comm., vol. 38, pp. 950-953, 1990.
- [5] Y. Yacoob and L. Davis, "Computing spatio-temporal representation of human face," CVPR '94, June 21-23,1994, Seattle, pp. 70-75, USA.
- [6] 김동훈, "영역 정보 기반 템플릿 모델을 이용한 얼굴의 특징 추출", 인하대학교 전자계산공학과 석사 학위 논문, 2000.2
- [7] F.I.Parke, K.Waters,"Computer Facial animation" ,A.K .Peters. 1996