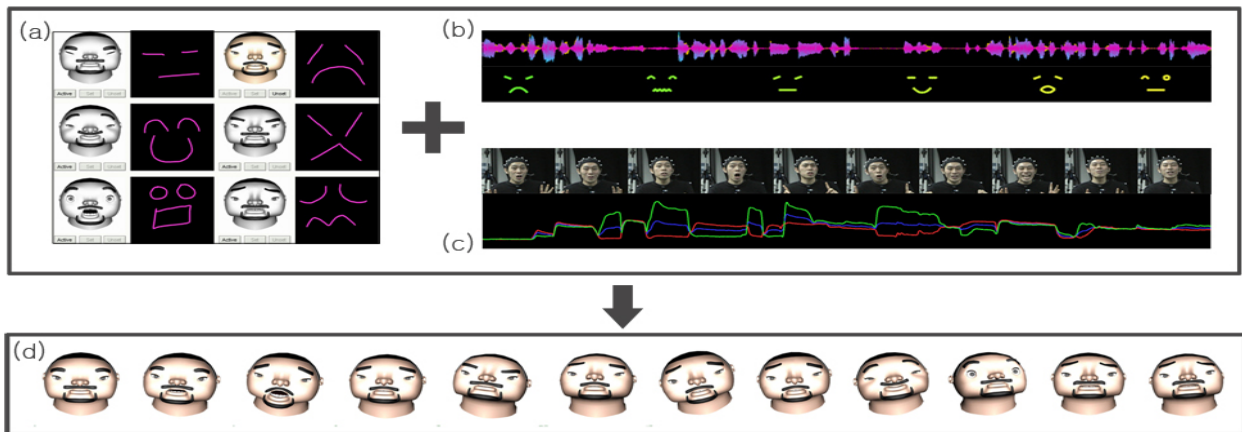


스케치 인터페이스를 이용한 데이터 기반 얼굴 애니메이션

주은정^o 안소민 이재희
 서울대학교 컴퓨터공학부
 {ejjoo, itscluster, jehee}@mrl.snu.ac.kr

Data-driven Facial Animation Using Sketch Interface

Eunjung Ju^o Sohmin Ahn Jehee Lee
 School of Computer Science and Engineering, Seoul National University



[그림 1] 전체적인 시스템 구조 (a) 사용자가 서로 다른 표정의 얼굴 모델을 자신만의 스트로크로 정의한다. (b) 음성 데이터를 기반으로 하는 타임라인의 원하는 위치에 키-프레임을 스케치한다. (c) 얼굴 캡처 데이터로부터 추출한 전이 정보를 이용하여 키-프레임 사이를 보간한다. (d) 이러한 과정들을 통해 감정을 표출하며 말하는 얼굴 애니메이션이 생성 된다.

1. 서론

자연스러운 얼굴 애니메이션 생성은 캐릭터 애니메이션 분야에서 중요한 문제이다. 지금까지 얼굴 애니메이션은 3차원 모델링 프로그램을 이용한 전문 애니메이터들의 수작업을 통해 생성되거나, 필요한 움직임 데이터를 직접 동작 캡처함으로써 만들어 왔다. 그러나 이러한 방식은 일반 사용자가 쉽게 접근 할 수 없으며 많은 시간과 비용을 요구한다는 단점이 있다. 본 연구에서는 실제에 가깝고 자연스러운 얼굴 애니메이션을 만들기 위해, 누구나 쉽게 사용할 수 있는 직관적인 방식의 스케치 인터페이스를 이용하고자 한다. 이를 통해 키-프레임을 생성하는 시스템을 구축하고, 얼굴 캡처를 통하여 얻은 데이터로부터 추출한 얼굴 표정간의 전이 정보를 이용하여 키-프레임을 보간하는 방식을 제안한다. 본 시스템은 전문 애니메이터가 아닌 일반 사용자도 쉽고 빠르게 다양한 감정을 표출하며, 동시에 말하는 얼굴 애니메이션을 만들 수 있도록 한다.

2. 관련 연구

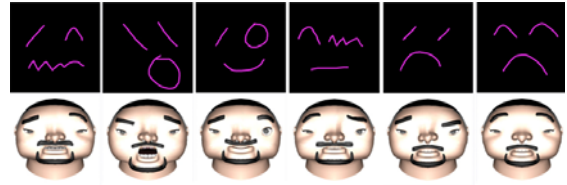
얼굴 애니메이션은 이미지나 비디오 같은 2차원 데이터 기반, 실제 사람 얼굴의 근육 모델 모방, 동작 캡처 혹은 스캔 데이터를 이용하여 3차원 얼굴 애니메이션을 제어하는 등의 다양한 형태를 기반으로 개발되어 왔다. 또한 애니메이션 생성의 목적에 있어서는 실제에 가까운 얼굴 표정을 표현하고자 하는 것[1]과 말하는 입모양에 주목하여 말하는 얼굴 애니메이션을 만들고자 하는 연구[2] 등이 있었다. 최근에 와서는 실제 사람처럼 감정 표현과 말하는 것 두 가지를 동시에 표현하는 연구[3]가 활발히 진행되고 있다. 이에 본 연구에서도 감정을 표현하면서 말하는 얼굴 애니메이션을 생성하는 것을 기본으로 한다. 또한 우리는 스케치 인터페이스 방식을 사용하여 일반 사용자도 쉽게 사용할 수 있는 시스템을 개발하고자 한다. 스케치는 모션 합성과 상호작용하는 캐릭터 애니메이션 생성 등에 있어서 널리 사용되어 왔다. 캐릭터의 걸거나 뛰는 경로를 스케치하거나[4,5], 모션의 2차원 자세를 스케치하여 3차원 자세를 만들어 넣으로써 애니메이션을 생성[6] 하도록 한 것이다. 이러한 연구들에서 보여지듯이 스케치 인터페이스는 사용자에게 있어 직관적인 입력 시스템으로써 비전문가 사용자들이 쉽게 애니메이션을 생성하는 것을 가능하게 해준다.

3.1 스케치 인터페이스를 이용한 키-프레임

우리는 키-프레임 보간법으로 캐릭터 합성을 할 때의 키-프레임 얼굴 모델을 스케치 인터페이스를 이용하여 생성하고자 한다. 그러나 얼굴 모델을 만드는 데에 있어서 개략적인 2차원 선의 입력만으로 3차원의 모델을 구성해 내는 것은 과소결정 문제로 이 문제를 해결하기 위해 목표 얼굴 모델에 대한 사전에 가정된 기본 지식이 필요하다. 본 연구에서는 사전 지식으로 전문 애니메이터들이 만든 다양한 표정의 얼굴 모델을 이용하며, 이러한 얼굴 모델을 몇 개의 영역으로 나누어서 저장하고 검색하여 사용한다. 이것은 사용자가 키-프레임을 스케치했을 때 얼굴의 각 영역별로 사용자의 각 스트로크에 맞는 모델들을 찾아 합성함으로써 새로운 얼굴 모델을 생성하게 됨을 의미한다.

정의된 형식 없이 자유롭게 그려진 그림은 해석의 범위가 광범위하기 때문에, 스케치 인터페이스를 이용하기 위해서는 입력 스트로크 형식의 정의가 필요하다. 우리는 표정을 얼굴의 주요한 부분을 상징적으로 표현하는 것으로 정의하고 양쪽

눈과 입을 각 한 개의 스트로크로 표현하기로 하였다. 이렇게 3개의 선으로 표현되는 얼굴의 표정에 대한 인식 역시 사람마다 다르다. 그래서 사용자는 이 시스템의 데이터베이스에 저장되어 있는 몇 가지 다른 표정을 가진 얼굴 모델에 대해서 각 모델을 3개의 선으로 표현해서 정의내리는 과정을 거치게 된다. 이후에 이 사용자가 키-프레임을 생성하고자 할 때 원하는 얼굴 표정을 미리 정의내린 스트로크로 표현하면, 시스템은 데이터베이스에서 가장 잘 매치되는 얼굴 모델을 찾아주게 된다.



[그림 2] 사용자의 스케치 입력에 의해 생성되는 얼굴 모델들은 양쪽 눈과 입이 서로 다른 감정을 표현하는 얼굴도 표현해준다.

우리의 시스템은 하나의 키-프레임을 생성할 때 사용자의 스케치 입력이 양쪽 눈과 입, 3개의 스트로크로 구성되었듯이 얼굴 모델을 세 부분으로 나누어서 사용한다. 이렇게 각 영역과 표정에 따른 분리된 모델들을 가지고 있음으로써, 사용자 입력인 세 개의 스트로크가 미리 정의된 얼굴 표정 중에서 서로 다른 표정을 표현하더라도 그에 맞는 얼굴 모델을 생성하는 것이 가능하게 된다. 이러한 방식을 통해 실제 가지고 있는 한 가지 표정만을 표현하는 얼굴 모델들을 이용하여 그 이상의 다양한 표정을 나타내는 새로운 모델들을 만들어 내는 효과를 얻을 수 있다.

3.2 얼굴캡처 데이터 기반 키-프레임 보간법

우리는 기본적인 애니메이션 생성에 있어서 혼합 모형 모델을 기반으로 하며 각 얼굴 모델을 세 개의 영역으로 나누어 이용한다. 혼합 모형 모델로 얼굴 모델을 만드는 것은 얼굴의 세 영역에 독립적으로 적용된다. 스케치 인터페이스를 이용하여 만들어진 키-프레임의 얼굴 모델들은 키-프레임 사이를 보간함으로써 하나의 완성된 애니메이션이 된다. 보간에 있어서 가장 일반적인 보간법인 선형 보간법은 키-프레임간의 간격이 넓을 경우 단조롭고 부자연스러운 애니메이션을 생성하는 문제점이 있기에 우리는 데이터 기반 보간법을 제안하여 이를 보완하고자 한다. 이는 사람의 얼굴 움직임을 캡처한 데이터로부터 실제 사람이 얼굴 표정을 바꿀 때의 전이 정보를 추출하여 이것을 키-모델간의 보간시에 가중치로 이용하는 것이다. 이로써 적은 개수의 키-프레임으로도 사실적이고도 자연스러운 애니메이션이 생성될 수 있다.

우리는 다양한 표정을 가진 얼굴 모델의 키-프레임 보간에 사용할, 실제 사람 얼굴 전면의 움직임에 대한 정보를 얻기 위해 마커 기반의 모션 캡처를 이용하였다. 사람의 얼굴 전반에 68개의 마커를, 머리 윗부분에 7개의 마커를 추가하여 붙임으로써 총 75개의 마커로 얼굴의 움직임을 캡처하였다. 이렇게 캡처된 데이터는 얼굴에 붙어 있었던 마커들의 3차원 좌표 형태로 저장되어서 매 프레임마다의 얼굴 전면의 절대적 위치를 표현해 준다. 이 데이터를 키-프레임의 보간에 이용하기 위해 서로 다른 표정으로 전이 할 때의 시간에 따른 변화 정보를 추출해야 한다. 현재 필요한 데이터는 시간에 따른 상대적 변화 정도이므로 각 데이터의 절대적인 수치가 아닌 기준 모델로부터의 상대적 값을 알아야 한다. 이를 위해 기준 모델을 무표정한 사람의 얼굴 상태로 정하고, 모든 프레임에 대해서 각 마커의 변화량을 무표정 상태의 마커위치와의 기하학적 거리로 정의하였다. 우리는 각 프레임마다 전체 마커 변화량의 평균값을 표정간 전이 정보 값으로 정의하였다. 이 데이터의 특징은 그 값이 시간의 변화에 따라 긴 구간 일정한 정도를 유지하다가 중간 중간 급격한 변화를 보이는 구간이 있다는 것이다. 이런 변화의 정도가 급격한 구간이 실제로 얼굴 표정의 전이가 일어난 시점이며, 우리는 이 시점의 데이터를 키-프레임 보간에 이용하였다. 전이 정보를 가져와서 적용시킬 목표 데이터가 키-프레임 형태이므로 전이 정보 역시 키-프레임에 기반하여 가공해야 한다. 우리는 변화량의 급격한 증감이 일어나는 시점, 즉 표정간의 전이가 일어나는 시점을 기준으로 데이터를 잘라 이를 비례 크기 조정하여 이것을 혼합 모형 모델의 혼합 가중치로 이용하였다. 이를 통해 데이터 기반 키-프레임 보간을 이용함으로써 자연스럽게 사실적인 애니메이션이 생성된다.

4. 결론

[그림 2]은 사용자의 여러 가지 스케치에 대해 만들어진 얼굴 모델이다. 사용자의 세 개의 스트로크가 서로 다른 세 개의 표정을 각각 표현할 때에 본 시스템은 그에 맞는 새로운 얼굴 모델을 생성해주었다. 또한 키-프레임의 보간에 있어서는 캡처 데이터로부터 추출한 전이 정보를 이용함으로써 시간에 따른 혼합 가중치가 선형의 단조로운 형태에서 전이 정보에 따른 비선형의 곡선을 이루게 됨으로써, 더 자연스럽게 사실적인 얼굴 애니메이션을 생성하였다. 본 시스템을 이용하여 한편의 애니메이션을 만들기 위해 실행되는 모든 과정은 실시간으로 진행된다.

본 연구의 한계점으로는 스케치로부터 3차원의 모델을 만들어내는 과소결정 문제를 캐릭터 모델로 구성된 데이터베이스로부터 스케치에 가장 잘 맞는 모델을 검색해서 찾아오는 방식으로 해결하기 때문에, 생성되는 애니메이션의 질이 데이터베이스의 양에 제한을 받는다는 것이 있다. 이러한 한계점을 극복하기 위해 본 시스템에서처럼 개략적인 스케치로 찾아낸 감정을 표현한 얼굴 모델을 기본으로 하고, 몇 개의 스트로크를 추가하여 다른 섬세한 조정을 하는 스케치 인터페이스를 개발할 수 있을 것이다. 이를 통해 한정된 데이터베이스에서 나온 제한된 개수의 얼굴 모델을 사용자별로 섬세한 부분을 조절할 수 있게 함으로써 키-모델 생성의 자유도를 증가시킬 수 있을 것이다.

5. 참고 문헌

- [1] Liu, Z., Shan, Y., Zhang, Z., Expressive expression mapping with ratio images, ACM Transactions on Graphics (SIGGRAPH 2001), 271-276
- [2] Ezzat, T., Geiger, G. & Poggio, T., Trainable videorealistic speech animation, ACM Transactions on Graphics (SIGGRAPH 2002), 21, 388-398
- [3] Cao, Y., Tien, W. C., Faloutsos, P. & Pighin, F., Expressive speech-driven facial animation, ACM Transactions on Graphics (SIGGRAPH 2005), 24, 1283-1302
- [4] Kovar, L., Gleicher, M. & Pighin, Motion Graphs, ACM Transactions on Graphics (SIGGRAPH 2002), 473-482
- [5] Lee, J., Chai, J., Reitsma, P. S. A., Hodgins, J. K. & Pollard, N. S., Interactive control of avatars animated with human motion data, ACM Transactions on Graphics (SIGGRAPH 2002), 21, 491-500
- [6] Davis, J., Agrawala, M., Chuang, E., Popović, Z. & Salesin, D., A Sketching Interface for Articulated Figure Animation, Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2003, 320-328