

# 광학식 동작 포착 장비를 이용한 노이즈에 강건한 얼굴 애니메이션 제작\*

박상일  
세종대학교 디지털콘텐츠학과  
sipark@sejong.ac.kr

Noise-Robust Capturing and Animating Facial Expression  
by Using an Optical Motion Capture System

Sang Il Park  
Department of Digital Contents, Sejong University

## 요 약

본 논문은 얼굴의 표정과 몸 동작을 광학식 동작 포착장비를 활용하여 동시에 포착하는 경우에 있어 얼굴 부위 마커들에 대한 노이즈에 강건한 데이터 처리 방법에 대해 다룬다. 일반적인 얼굴 표정만 포착하는 경우와 달리, 몸의 움직임과 동시에 포착할 경우 포착용 카메라가 멀리 있어 얼굴에 붙인 마커들의 궤적 데이터는 특별한 처리를 요한다. 특히 궤적의 포식화, 빈 곳 메우기, 노이즈 제거의 과정이 필수적이며, 이러한 과정을 위해 본 논문에서는 지역좌표에 기반을 둔 궤적 데이터 처리 방법을 제안한다. 지역 좌표는 강체변형에 불변한 특징이 있으며, 얼굴 모양의 국지적인 변화를 의미하여, 궤적 데이터처리에 효과적으로 활용 될 수 있음을 보였다. 또한 제안한 방법을 활용하여 애니메이션을 제작해 실제 제작 환경에 적용 가능성을 보였다.

## ABSTRACT

In this paper, we present a practical method for generating facial animation by using an optical motion capture system. In our setup, we assumed a situation of capturing the body motion and the facial expression simultaneously, which degrades the quality of the captured marker data. To overcome this problem, we provide an integrated framework based on the local coordinate system of each marker for labeling the marker data, hole-filling and removing noises. We justify the method by applying it to generate a short animated film.

**Keyword** : 3D animation, character animation, facial animation

접수일자 : 2010년 01월 05일 일차수정 : 2010년 09월 02일 심사결과 : 2010년 09월 24일

\* 본 논문은 서울시 산학연 협력사업(10557M0214835)의 지원에 의해 연구되었습니다.

## 1. 서론

게임의 사실성과 몰입감을 높이기 위해서는 사실적인 캐릭터의 동작이나 표정 생성이 필수적이다. 이는 게임 사용자가 평상시 사람의 동작이나 얼굴에 익숙해져 있어, 게임 상 가상캐릭터의 움직임의 약간의 이상함이라도 금방 느낄 수 있기 때문이다. 최근 광학식 동작 포착 장비가 대중화되면서 간편하게 사실적인 캐릭터의 동작을 생성하는데 많이 사용되고 있다. 배우의 몸의 주요 관절 부위에 전부 약 40~60 여개의 마커를 붙이고, 이를 포착하여 강체 뼈대의 움직임을 가정하여 그 자세를 재구성하는 방식이다. 배우의 얼굴 표정을 포착하는 데 또한 광학식 동작 포착 장비가 주로 사용되는데, 이때는 몸과 같은 강체 뼈대라는 개념이 없으므로 얼굴에 모두 약 100여개의 마커를 2~3cm 간격으로 촘촘히 부착하여 각 점의 움직임을 포착하고, 이로부터 각 마커에 해당되는 작은 부분 영역들을 비강체 변형을 가정하여 변형하게 된다.

광학식 동작 포착 장비로 얼굴 표정을 포착할 시에는 마커를 좁은 영역에 많이 부착하므로 광학식 동작 포착의 단점인 마커간의 간섭이나 혼동이 많이 발생한다. 따라서 일반적으로 다수의 카메라를 얼굴을 향하도록 배치하고, 배우의 얼굴 외의 상반신을 포함한 다른 움직임을 제한하고 항상 카메라를 바라보도록 하여 포착되는 데이터의 정확도를 높이는 방식을 사용한다. 그러나 이러한 방식은 캐릭터의 얼굴뿐만 아니라 몸의 애니메이션까지 만들어야 할 경우에, 배우로부터 몸동작과 얼굴 표정을 따로 포착해야 하므로, 두 별개의 데이터 간의 동기화에 문제가 생기며, 또한 배우의 입장에서 두 번의 동작을 각기 몸과 얼굴표정에 대해 나누어서 연기해야 하므로 동작의 자연스러움이 제한될 수 있는 문제가 있어왔다.

위 문제에 대한 가장 기본적인 해결방안은 몸 동작 포착을 위한 광학식 동작 포착 기기 설정을 하고 여기에 얼굴에 마커를 추가로 부착하여 몸의

동작과 얼굴의 표정을 동시에 포착하는 방법이다. 그러나 이 경우 몸 동작 포착의 경우는 별다른 문제가 없으나, 얼굴 표정 포착의 경우, 얼굴 포착만을 위한 기기 설정과는 달라 현실적인 문제들이 발생할 수 있다. 먼저 카메라가 얼굴 표정만 포착하는 상황보다 원거리에 위치하게 되며, 이 경우 마커 간의 거리를 좁히는 데 한계가 있게 되어 얼굴에 붙이는 마커의 수가 얼굴만을 포착할 때 보다 적게 되며, 또한 포착된 마커의 3차원 위치 정보에 간섭이나, 가림, 혼동, 노이즈 등의 문제점들이 두드러지게 발생하여 효과적인 데이터의 후처리 보정 작업이 반드시 필요하게 된다. 본 논문은 위의 두 가지 문제, 즉 광학식 포착 장비로부터 얻은 부정확한 마커 데이터를 보정하는 문제에 대한 현실적인 해결법을 제안하며, 또한 적은 수의 마커로부터 비교적 정밀한 변형을 만들어 내는 애니메이션 문제에 대한 방법론을 제시한다. 본 논문의 핵심 착안점은 마커의 궤적 데이터를 후처리함에 있어 전역좌표가 아닌 지역좌표를 활용하는 것으로 이를 통해 궤적 표식화, 빈 곳 메우기, 노이즈제거 등의 데이터 보정 작업을 통합적으로 수행한다. 적은 수의 마커로 정밀한 애니메이션을 제작하기 위해서 전체적인 비선형 변형을 선형 변형과 국부적인 비선형 단계로 나누어 비선형 변형의 범위를 줄여 효과적으로 처리함을 제안한다.

본 논문은 다음과 같은 순서로 구성되어 있다. 먼저 2장에서는 관련된 기존 연구들을 살펴보고, 이후 3장에서 본 방법의 개략적인 개관을 제시한다. 4장과 5장에서는 방법에 관하여 자세한 설명을 하며 6장에서는 실험결과를, 7장에서는 결론 및 향후 연구 방향을 제시하며 끝을 맺는다.

## 2. 관련 연구

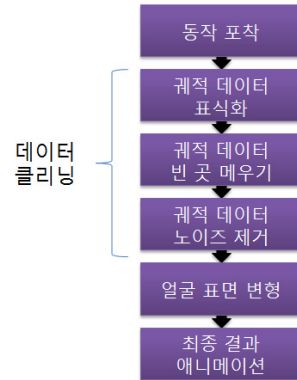
게임 제작 시 얼굴 애니메이션을 만드는 방법은 각기의 얼굴 표정들을 디자이너가 잘 디자인한 뒤 애니메이터가 수작업을 통해 표정을 섞는 방법을

취하고 있다. 이러한 방법은 시간이 많이 걸릴 뿐만 아니라 최종 결과 애니메이션이 애니메이터의 숙련도에 종속된다. 최근에는 음성을 입력으로 받아 이를 분석한 립싱크 방식에 따라 미리 정의된 변형들을 조합하는 방식을 취하기도 한다[1,2,3]. 그러나 이러한 방법들은 생성 가능한 얼굴표정이 미리 정의된 모형의 한도 내에서 제한되므로 실제 사람과 같은 다양한 표정을 만들어 내는 데에는 제약이 있었다. 이를 극복하기 위해 최근에는 광학식 동작 포착 장비를 이용하여 표정을 포착하고 이로부터 사실적인 애니메이션을 제작하는 과정을 많이 사용한다.

광학식 동작 포착 방식을 이용한 얼굴 애니메이션 제작은 마커 정보의 활용 방법에 따라 얼굴 리깅 기반 방법과, 얼굴 표면 변형 방법으로 나누어진다. 얼굴 리깅 방법은 얼굴 특징점에 부착된 마커의 위치와 이와 연관된 비교적 넓은 부위의 얼굴 표면의 변화를 미리 연관시켜 놓고, 마커 위치 정보를 제어 신호(control signal)로서 사용하는 방식으로, 적은 수의 마커로도 얼굴 전체의 표정을 생성할 수 있는 장점이 있으나, 결과 애니메이션의 품질이 미리 정의된 표정 변형 정보에 의존적이 되는 단점을 여전히 가지고 있다[4,5,6,7]. 미리 정의된 표정 변형 정보를 얻기 위해 Chung과 Bregler는 비디오를 분석하여 표정과 발음에 연관된 표면 정보를 분리하는 방법을 제안하였으며[4], Joshi와 Pyun등은 미리 정의된 변형 정보로부터 물리적으로 유의미한 보다 축약된 변형정보를 얻는 방법을 제안하였다[5,6]. Ju와 Lee는 미리 정의된 변형 정보에 불규칙한 상해 변형을 추가하는 방식으로 생성된 애니메이션의 사실성을 높였다[7].

이와는 달리, 마커를 표면 정보로서 직접 적용시키는 얼굴 표면 변형 방법은 동작 장비로부터 직접적으로 얼굴의 형태 변형을 포착하기 때문에 파라미터의 조정이나, 애니메이터의 수작업이 최소화되는 장점이 있다. 그러나 얼굴의 세밀한 변화를 포착하기 위해서는 마커를 상대적으로 많이 붙여야 하므로, 동작 포착 정보의 처리 시 데이터가 많이

필요한 단점이 있다[8,9].



[그림 1] 방법의 개관

위 방법들은 대부분 얼굴과 동작을 따로 포착할 경우의 얼굴 표정 포착 방식이나 최근 들어 몸 동작과 얼굴 표정을 동시에 포착하는 방법들이 영화 제작사들에 의해 개발되어 사용되고 있다. 영화 ‘베오울프’(2007)에서는 얼굴에 마커를 촘촘히 붙이고, 약 120여개의 포착용 카메라를 활용하여 얼굴 마커 포착을 수행하였으나 이와 같은 대단위 포착 장비의 활용은 우리나라의 제작 현실에 적용되기 힘들다. ‘크리스마스 캐롤’(2009)과 ‘아바타’(2009) 등과 같은 최신 영화에서는 얼굴 마커만을 포착하기 위한 추가 소형 카메라를 머리 부위에 부착하는 방법을 활용하고 있으며, 이를 통해 몸 동작과 얼굴의 세밀한 표정 변화를 동시에 포착하고 있으나, 아직은 일반적으로 활용되고 있지 않으며 추가 장비 구입의 비용이 든다. 본 연구에서 제안된 방식은 일반적인 동작 포착 장비의 활용만을 통한 얼굴 표정 동시 포착의 경우를 다루며, 이를 데이터처리 방법의 개선을 통해 이루고자 한다.

### 3. 방법 개관

본 방법은 광학식 동작포착장비로부터 얼굴에

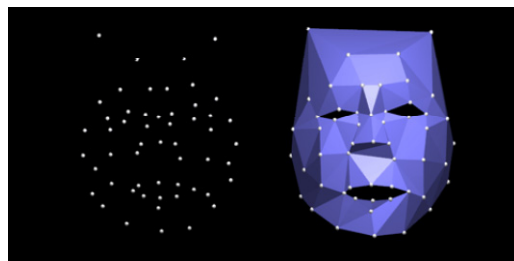


[그림 2] 얼굴 표면과 마커

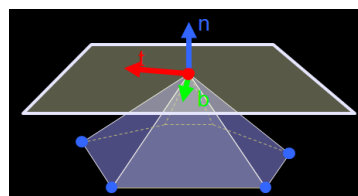
붙인 마커의 움직임 데이터를 얻는 과정과 이로부터 실제 변형을 생성하는 두 과정으로 이루어진다. 본 논문이 가정하고 있는 동작 포착 환경은 얼굴 표정 포착만을 위한 설정이 아닌, 몸 전체 동작 포착을 위한 장치 설정 상에서 얼굴에 마커를 추가로 붙이고 얼굴 표정을 몸동작과 동시에 포착하려는 상황을 가정한 것으로, 이 경우 동작과 얼굴의 동기화가 용이한 장점이 있다. 그러나 카메라가 일반적인 얼굴 포착만을 위한 설정보다 멀리에 위치하며, 얼굴을 정면으로 바라보고 있지 않은 경우가 많아 포착한 데이터가 정확하지 않으며, 얼굴에 붙일 수 있는 마커의 수가 상대적으로 제한 되게 되어 표면 변형 시에 이를 고려한 효과적인 변형 방법을 고안하여야 한다. [그림 1]은 본 방법의 수행 단계를 도식화해서 보여 준다. 먼저 광학식 장비를 통해 동작 포착을 수행하고, 얻어진 마커들의 궤적 데이터를 궤적 표식화, 빈 곳 메우기, 노이즈 제거의 세 단계를 통해 후보정(post-processing)한다. 후보정이 끝난 마커 데이터로부터 세밀한 얼굴 표면의 변형을 유추하여 만들고, 이를 MAYA와 같은 외부 렌더링 프로그램을 이용하여 가시화함으로써 최종 애니메이션을 제작한다.

#### 4. 데이터 포착 및 보정

광학식 동작 포착 장비로 얻어진 마커의 움직임



[그림 3] 마커와 이로부터 구성된 마커 표면 메시



[그림 4] 각 마커마다 정의된 지역좌표계

데이터는 일반적으로 가려짐이나 마커 간의 혼동으로 인해 이를 바로 애니메이션에 적용할 수 없으며, 데이터 보정 작업이 필수적이다. 일반적인 보정 작업은 크게 세 과정으로 이루어지는데, 먼저 포착 되던 마커가 가림(occlusion)에 의해 일시적인 포착 실패 후 다시 포착되었을 때 이전 마커와 동일한 마커임을 판단해야 하며 이러한 과정을 궤적 표식화(trajecctory labeling)라 한다. 두 번째로, 궤적 표식화 후 가림에 의해 포착되지 않은 시간 동안의 마커의 위치를 예측하여 온전한 궤적 데이터를 만들어야 한다. 이 과정을 빈 곳 메우기(hole filling)라 부른다. 마지막으로, 포착된 데이터에는 측정의 오차 및 마커 간 상호 간섭에 의해 부자연스런 떨림이나 갑작스런 위치 변화 등의 노이즈가 있으며, 이를 제거하여 깨끗한 신호로 복원해 주어야 하며, 이를 노이즈 제거(noise removal)라 한다.

#### 4.1 지역 좌표계 기반 위치 표현

본 연구에서는 마커의 위치를 전역좌표(global coordinate System)가 아닌 지역좌표(local coordinate system)에 기반을 두어 표현함으로써

위 세 문제, 즉 궤적 표식화와 빈 곳 메우기 및 노이즈 제거를 해결하는 방법을 제시한다. 기본적인 착안점은 마커의 경우 전역좌표보다는 주위 마커와의 상대 위치를 의미하는 지역좌표가 변형의 의미를 더 잘 반영하고 있다는 데 있다.

공간상의 한 공통의 원점을 기준으로 모든 마커의 위치를 표시하는 전역좌표와 달리, 지역좌표는 각 마커마다 그 마커를 원점으로 하는 좌표계를 정의하고, 주변 마커들의 위치를 자신의 좌표계로 표현 하는 것이다. 이 경우, 회전과 이동만을 포함하는 강체변형의 경우 지역 좌표가 변하지 않는다는 장점이 있다. 또한, 얼굴과 같은 비강체 변형에 대해서도 지역 좌표계는 값이 비교적 일정하다는 장점을 가질 수 있다. 예를 들어 전체적으로 얼굴을 관찰했을 때, 눈, 입, 턱 등의 부분에 대단위 변형이 일어나는 것을 볼 수 있으나, 국부적으로 한정하여 바라봤을 경우에는, 그다지 변형이 심하지 않을 것 또한 확인 할 수 있다. 이는 전체의 변형이 크다 해도 (즉, 전역좌표의 값의 변화가 커도), 국부적인 상대적인 좌표 값들은 크게 변하지 않음을 의미한다.

3차원 상에서 좌표계를 정의하기 위해서는 세 축이 정의되어야 한다. 마커 하나만으로는 축이 정의 되지 않으므로, 가까운 주변의 마커들과의 상대적인 위치 정보를 이용하는데, 주변의 마커가 2개 이상 주어진다면 이들 간의 위치 관계로부터 세 축을 일관되게 정의 할 수 있다. 한 마커와 그와 가까운 주변의 마커를 정의 하는 방법으로, 본 연구에서는 마커들을 정점으로 간주하여 삼각화를 수행한 삼각형 메시를 정의하고([그림 2,3]참조), 이로부터 주변 정보를 얻어오는 방법을 취한다. 삼각형 메시가 만들어 지면, 한 마커와 직접 예지로 연결된 다른 마커들을 주변 마커로 정의할 수 있다.

Lipman 등[10]이 제안한 것과 유사한 방법으로 그림 4에서 보이는 것과 같이 한 마커와 그들의 주변 마커로 이루어진 삼각형들을 분석하여 이들의 평균 법선 방향을 좌표계의 한 축  $\mathbf{n}$ , 임의의 주변 마커로의 방향 벡터와  $\mathbf{n}$  벡터와의 수직인 성분의

단위 벡터를  $\mathbf{b}$ , 그리고 이 두 벡터와 모두 수직인 접선 벡터  $\mathbf{t}$ 를 정의할 수 있으며, 이들을 세 축으로 하는 로컬 좌표계를 각 마커마다 만들 수 있다. 좌표계가 주어지면 한 마커  $i$ 에 대해 주변 마커들의 인덱스의 집합을  $N_i$ 라고 할 때, 주변 마커  $j \in N_i$ 의 전역좌표  $\mathbf{p}_j$ 를 마커  $i$ 의 지역 좌표계로 표현한 지역위치  $\tilde{\mathbf{p}}_{ij}$  들을 만들 수 있으며 이는 다음 수식을 통해 계산 될 수 있다.

$$\tilde{\mathbf{p}}_{ij} = \mathbb{R}_i(\mathbf{p}_j - \mathbf{p}_i), \quad (\text{식 1})$$

여기서,  $\mathbb{R}_i$ 는 임의의 전역 기준 좌표계와 마커  $i$ 로부터 만들어진 좌표계와의 방향 차이를 의미하는  $3 \times 3$  회전 행렬이다.

얼굴의 표정변화나 움직임에 따라 마커의 전역좌표가 변하듯 지역좌표도 변형에 따라 변하게 된다. (단, 회전/이동과 같은 강체변형만 있을 경우 불변한다.) 따라서 지역좌표의 값도 시간에 따른 함수이며, 통상적인 애니메이션 제작 시에 T-pose와 같은 기준 자세를 기반으로 모델을 만들고 이를 물체의 기준 모양으로 생각하는 것과 같이, 얼굴 또한 [그림 3]과 같은 기본 중립 표정 상태에서 기준지역좌표를 정할 수 있다. 본 논문에서는 특별한 언급이 없는 한 이러한 기준 중립 표정 자세에서 측정된 지역좌표를 얼굴의 지역좌표  $\tilde{\mathbf{p}}_{ij}$ 라 칭한다.

## 4.2 마커 궤적 데이터 표식화

한 마커의 지역좌표계로 주변 마커들의 좌표가 표현되면 이 좌표는 강체 변형에 불변하며, 비강체 변형에도 국부적으로 큰 변형이 일어나지 않는 한 값이 비교적 일정하다. 이러한 성질을 이용하여 본 논문에서는 마커 궤적에 대한 후처리를 수행한다. 먼저 궤적 표식화를 수행하는 방법에 대해 설명한다. 먼저, 본 방법은 첫 번째 프레임에서는 얼굴에 붙인 모든 마커가 포착되며, 각 마커에는 구별이

가능하도록 표식화가 되어있는 것을 가정한다. 이러한 가정은, 정상적인 동작포착은 카메라의 가까운 곳에서 모든 마커가 다 보일 수 있는 자세를 기본자세로 취하고 난 후 시작하기에, 일반적인 상황으로 무리가 없다고 볼 수 있다. 포착이 진행됨에 따라 각 마커의 움직임은 연속된 궤적 데이터로 표현되며, 포착 시 가려짐 등의 포착 실패 현상이 발생하면, 해당 마커의 궤적은 끝나게 된다. 가려짐 이후 다시 나타나게 되면, 새로운 궤적으로 인식되어 별개의 데이터로 처리된다. 이렇게 새롭게 포착된 궤적은 표식화가 되어 있지 않게 되며, 따라서 포착 후 궤적 데이터는 표식화가 되어 있는 데이터와 안 되어 있는 데이터로 분류될 수 있다.

본 논문은 한 마커의 궤적이 포착에 실패 했을 경우, 실패한 시간 동안 마커의 위치를 대략적으로 파악하여 이를 기반으로 같은 시간에 표식화가 되지 않은 다른 궤적들을 검사하고, 이 중 가장 높은 확률로 기존의 궤적에 연결될 가능성이 있는 궤적을 찾아 이를 기존의 궤적과 합쳐 주는 과정을 수행한다. 이때 핵심은 사라진 마커의 대략의 위치를 파악하는 것이다.

포착에 실패한 마커의 위치를 파악하기 위해 본 논문이 취하는 방식은 그 주변 마커들의 현재 위치로부터 사라진 마커의 위치를 추산하는 것이다. 주어진 한 프레임  $f$ 에서 포착에 실패한 마커  $i$ 에 대해 이 마커의 한 주변 마커  $j \in N_i$ 로부터 마커  $i$ 에 대한 미리 계산된 지역 좌표  $\tilde{\mathbf{p}}_{ji}^f$ 를 이용하여 다음과 같은 수식을 이용하여  $j$ 로부터 추산된 지역좌표  $\hat{\mathbf{p}}_{ji}^f$ 를 복원해 낼 수 있다.

$$\hat{\mathbf{p}}_{ji}^f = \mathbb{R}_j^{f-1} \cdot \tilde{\mathbf{p}}_{ji} + \mathbf{p}_j^f, \quad (\text{식 2})$$

여기서,  $\mathbb{R}_j^f$ 와  $\mathbf{p}_j^f$ 는 각각 프레임  $f$ 에서의 마커  $j$ 의 지역좌표계를 의미하는 회전 행렬과, 전역 위치를 의미한다.

한 마커에는 복수의 주변 마커들이 있으므로 이

들을 다 고려하여 다음과 같이 평균을 통해 프레임  $f$ 에서의 마커  $i$ 의 최종 추산 위치  $\hat{\mathbf{p}}_i^f$ 를 결정한다.

$$\hat{\mathbf{p}}_i^f = \frac{\sum_{j \in N_i^f} \hat{\mathbf{p}}_{ji}^f}{n_i^f}, \quad (\text{식 3})$$

여기서  $N_i^f$ 는  $i$ 의 주변마커 중 프레임  $f$ 에서 포착이 된 마커들의 집합을 의미하며,  $n_i^f$ 는 이들의 총 수를 의미한다.

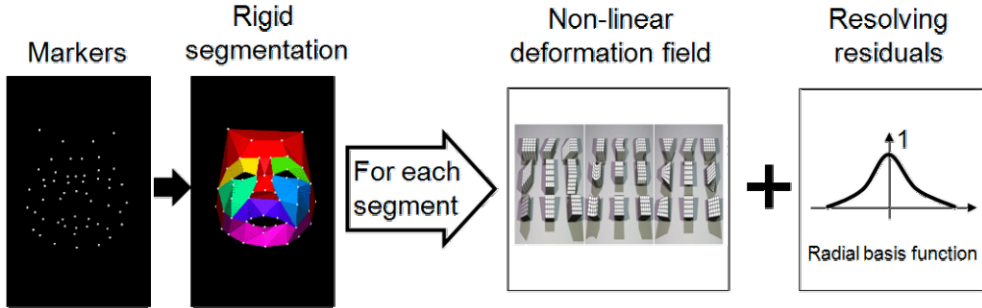
이렇게 계산된 추정 위치를 기반으로 프레임  $f$ 에서 포착은 되었으나, 표식화 되지 않은 궤적들을 검색하여 이들 중 미리 설정된 경계값(threshold)보다 가까이 있으며, 또한 그 궤적이 시간 상에서 마커  $i$ 의 궤적과 오버랩 되지 않는 것들 중 최소 거리를 가진 것을 골라 마커  $i$ 로 표식화 한다. 본 연구에 사용한 경계값으로 마커 간의 평균 거리의 50%를 사용하였다.

한 프레임에서는 복수개의 표식화 되지 않은 마커들이 존재하며 이때 그들 중 표식화 하는 순서에 따라 정확도가 달라 질 수 있으며, 일반적으로 주변마커가 많이 존재할수록 추정 위치의 정확도가 높아짐으로, 본 논문에서는 탐욕(greedy) 접근법에 기반을 두어, 현재 표식화 되지 않은 마커들 중 가장 주변 마커가 많이 존재하는 것부터 차례로 표식화 해 가는 방법을 취한다.

### 4.3 마커 궤적 데이터 빈 곳 메우기

표식화를 통해 포착은 되었으나 식별이 안 된 궤적들을 최대한 표식화 한 후 실제 포착에 실패하여 포착이 안 된 마커들의 위치를 추정하여 데이터를 채워 나간다. 본 논문은 이에 대해서도 지역좌표에 기반을 둔 방법을 제안한다.

지역좌표는 어느 정도 일정한 값을 유지하는 경향이 있으나, 변형에 따라 값이 변하게 된다. 따라서



[그림 5] 얼굴 표면 형상 변형의 과정: 강제변형과 2차 변형, 방사기반 변형으로 구성된다.

지역좌표 자체를 그대로 이용하여 값을 추산할 순 없다. 이런 이유로 Park과 Hodgins는 지역좌표에 대해 주성분분석(Principal Component Analysis)과 같은 통계적 방법을 취하였다[11]. 그러나 이러한 방법은 입력된 데이터에 따라 그 분석 값의 신뢰도 및 강건성이 많이 의존되는 문제가 있어, 소규모의 선형 변형에 적합하였다. 그러나 본 논문이 다루는 얼굴 표정은 주름과 턱 등에서 지역별로 변형의 정도가 몸의 다른 부분보다 많이 일어나므로 선형변형을 가정하진 위 방법은 적합하지 않으며, 보다 강건한 방법이 필요하다. 이를 위해 본 논문은 앞 4.2절의 (식 3)에서 계산한 추산 위치  $\hat{\mathbf{p}}_i^f$ 를 포착에 실패한 마커의 최종위치로 사용하는 방법을 제안한다. 이는 주변마커의 지역위치의 평균을 통해 구하므로 그 값이 강건한 특징이 있으며, 국부적인 변형에 대해 잘 표현되는 장점이 있다.

#### 4.4 마커 데이터 노이즈 제거

원거리에서 포착된 마커 데이터는 근원적인 광학식의 특성 상 노이즈를 많이 가지고 있으며, 또한 앞 절의 방법에 따라 사라진 마커를 복원하였을 경우에도 그 추산된 결과 값이 부드럽게 연결되지 않을 수 있어, 시간상 부드러운 궤적을 가질 수 있게 데이터를 보정하는 과정이 필수적이다. 기존의 전역 좌표에 기반을 둔 저역통과필터(low-pass filter)같은 방식을 사용할 수 있으나,

이는 물리적인 관점으로 보면 마커의 위치를 같은 자리에 머무르도록 하는 특성이 있어, 이를 이용하려는 전반적인 마커의 움직임이 더뎠으며, 표현을 재구성했을 경우 표면이 망가지는 단점이 있다. 예를 들어 회전/이동만을 하고 있는 강제 변형의 경우에도 포착한 마커들을 각각 저역통과필터 처리를 하게 되면, 마커 간의 거리가 달라지는 등 물체의 모양 자체가 변하는 문제들이 생긴다. 이러한 단점들을 극복하기 위해 본 논문에서는 마커의 지역좌표에 기반을 둔 필터링 방법을 제안한다.

핵심 아이디어는 현재 데이터로 가지고 있는 마커의 위치와 4.2절의 방법으로 추산한 위치 간의 차이를 지역변형의 값으로 보고, 이 값을 일반적인 신호처리 기법으로 후처리 하는 방식이다. 이러한 방식을 취할 경우 회전/이동과 같은 강제변형의 경우 지역좌표가 불변하기 때문에 물체의 원래의 모양을 최대한 유지하는 방향으로 필터링이 수행되게 된다.

이 방법을 수식화하면 다음과 같다. 프레임  $f$ 에서 한 마커의 위치  $\mathbf{p}_i^f$ 와 주변마커를 이용한 추산 위치  $\hat{\mathbf{p}}_i^f$ 가 주어지면, 이를 통해 다음과 같은 변위  $\mathbf{d}_i^f$ 를 정의할 수 있다.

$$\mathbf{d}_i^f = \mathbf{p}_i^f - \hat{\mathbf{p}}_i^f. \quad (\text{식 4})$$

이 위치를 지역통과필터와 같은 신호처리의 대상 값으로 다음과 같이 적용하여 새로운 변위  $\bar{\mathbf{d}}_i^f$ 를 구한다.

$$\bar{\mathbf{d}}_i^f = \frac{\sum_{t=-k}^k \mathbf{d}_i^{f+t}}{2k+1}, \quad (\text{식 5})$$

여기서,  $k$ 는 커널의 사이즈(본 논문에서는 5를 사용)이며, 노이즈 제거를 위해 평균필터(mean filter)를 적용한 수식이다. 이렇게 필터링 후 변위 값을 다시 추산 위치에 대해 주어 최종 필터링 된 마커의 위치를 계산한다.



[그림 6] 동작 및 얼굴 포착 장면

$$\bar{\mathbf{p}}_i^f = \hat{\mathbf{p}}_i^f + \bar{\mathbf{d}}_i^f. \quad (\text{식 6})$$

## 5. 얼굴 표면 형상 변형

포착된 마커로부터 변형된 얼굴의 표정을 만들기 위해 본 논문에서는 Park과 Hodgins의 방법을 적용하였다[11]. 이 방법은 마커들로부터 포착된 표면의 대략적인 모습을 세밀한 모델에 적용하는 방법으로 변형의 비선형성을 고려하여 얼굴 전체를

부분적인 강체들의 집합으로 근사하는 방법을 취한다([그림 5]참조).

이를 위해 먼저 마커들을 그 움직임이 유사한 것끼리 집합을 만들어 준다. 본 연구의 경우 얼굴을 이마, 좌/우 눈꺼풀, 좌/우 광대뼈, 윗입술, 턱의 7개의 영역으로 분할하였다. 이렇게 마커들을 분할하면 각 영역의 움직임과 변형은 각 할당된 마커의 움직임으로 표현된다. 각 영역 안에서의 비선형적 변형을 효과적으로 표현하기 위해 본 논문에서는 각 변형을 3가지 단계의 하위 변형으로 분할하여 이들의 합으로 전체를 표현한다. 하위 변형은 각각 강체변형, 2차(quadratic) 변형, 방사기반(radial basis) 변형이다.



[그림 7] 마커로부터 얼굴 애니메이션 생성 결과

강체변형은 한 영역 내에서 마커들의 움직임이 그 영역의 강체 운동을 표현한다고 가정하고, 이를 통한 영역을 대표하는 이동과 회전의 값을 구하는 것으로, 이는 일반적인 절대 방향(absolute orientation)문제로 수식화 될 수 있으며 그 해석해가 잘 알려져 있다[12].

2차 변형은 강체변형으로 표현되지 않는 나머지 변형에 대해 이를 2차원 함수들의 집합으로 표현하는 것으로 이를 통해 구부림, 비틀림, 부풀림 등의 변형을 근사할 수 있다. 방사기반변형은 분산데이터 보간의 가장 일반적인 방법으로 각 위치에서의 변형 값을 거리에 반비례하여 분산시키는 방법이다.

본 논문에서는 Park과 Hodgins의 방법에 따라 위 세 개의 변형을 차례대로 적용시켜 마커로부터 최종 애니메이션을 얻었다.



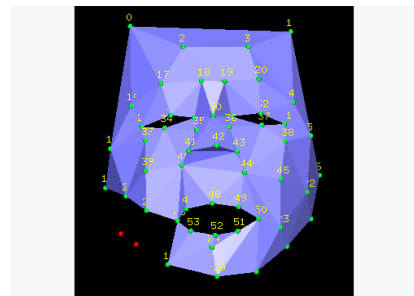
## 6. 실험 결과

본 연구의 실험을 위해 미국 카네기멜론 대학의 적외선 방식의 동작 포착 장비를 활용하여, 포착 카메라 12대를 사용하였으며, 2명의 연기가 서로 호흡을 맞춰 연기하는 시나리오를 목표로 동작 및 얼굴 표정을 동시에 포착하였다. 포착 프레임 빈도는 120 frame/sec 이며, [그림 6]은 동작 포착 당시의 장비 설정을 보여준다. 일반적인 얼굴 포착만을 위한 설정과는 달리 동작을 포착하기 위해 포착 카메라들이 모두 연기자와 멀리 떨어져 있다. 연기자의 얼굴에는 [그림 2]에서 보이는 것과 같이 표정을 포착하기 위해 56개의 작은 마커를 부착하였다. [그림 7]은 남성 연기자에 대해 포착된 마커로부터 재생성한 얼굴의 애니메이션의 모습을 보여준다. 마커의 움직임에 따라 얼굴 표면의 변형이 일어남을 확인할 수 있다. [그림 8]은 본 방법을 사용하여 만든 애니메이션의 한 장면을 보여준다. 최종 결과 동영상은 인터넷 사이트에서 확인해 볼 수 있으며 (<http://dasan.sejong.ac.kr/~sipark/face/>) 이를 통해 2명의 연기자에 대해 동작과 얼굴을 동시에 포착하여 자연스러운 애니메이션을 만드는 것이 가능했음을 보인다.

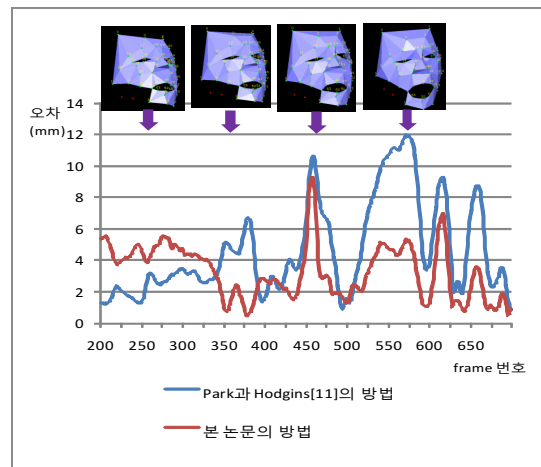
본 논문에서 제안한 마커 궤적 데이터의 빈 곳 메우기 방법의 효과성을 검증하기 위해 다음과 같은 정량적 실험을 수행하였다. 미리 잘 포착된 마커 궤적 데이터로부터 일부 구간동안 특정 마커의 정보를 삭제한 후 이를 제안된 방법으로 복원한 뒤 삭제 전의 원본 정보와의 차이를 비교하여 정확성을 측정하였으며, 특히 Park과 Hodgins의 PCA 기반 방법[11]과의 오차 결과 비교를 수행하였다. 실험에 사용한 데이터는 1902 프레임 길이 (약 15.85초)의 [그림 8]의 결과 애니메이션을 만드는 데 사용한 남성의 데이터이며, [그림 9]에서 보이는 것과 같이 턱 부분의 마커 2개를 임의로 선택하여 검증을 수행하였다. 이들 마커를 선택한 이유는 메시의 경계면에 위치하고 있어 참고할 주변 마커의 개수가 다른 부분에 비해 적어, 빈 곳 메우

기 방법의 오차가 가장 클 수 있는 부분이기 때문이다. 본 실험을 위해 선택한 구간은 200 프레임부터 700 프레임 구간이며, 이 구간을 선택한 이유는 이 구간 내에서 실험에 선택한 마커의 원본 데이터 자체가 빈 곳 메우기 방법이 필요 없게 잘 포착되어 있었기 때문이다.

[그림 10]은 빈 곳 메우기 방법의 오차 결과 데이터를 보여준다. 본 논문의 방법이 Park과 Hodgins[11]의 방법보다 보다 정확한 결과를 보여줌을 확인 할 수 있다. 특히 입을 크게 벌리는 등의 큰 변형이 일어나는 구간에 대해 원래의 정보를 더 잘 복원하는 것을 확인할 수 있다. [표 1]은 오차의 통계값으로 평균과 표준편차를 보여준다.



[그림 9] 빈 곳 메우기 방법 검증을 위해 사용한 마커 (적색 표시)



[그림 10] 빈 곳 메우기 방법의 오차 분석 결과

본 논문의 방법이 더 잘 동작하는 이유는, Park과 Hodgins[11]의 방법은 몸의 변형을 구하는 것을 목적으로 폐곡선(closed-surface)으로 구성된 메시를 가정했으나, 얼굴애니메이션의 경우는 본 논문의 경우와 같이 개곡선(open-surface)으로 구성된 메시가 일반적으로 쓰이며, 이 경우 경계면의 마커들은 참고할 주변 마커의 개수가 적어져, PCA와 같은 주변 변형을 적극적으로 반영하는 방법은 오차가 증폭될 수 있으나, 본 연구에서 제안한 복원 방법은 이를 오차로 보아 첫 프레임의 모습을 최대한 유지하기 때문이다.

한 좀 더 전역적인 정보들, 예를 들면 얼굴 턱 왼쪽과 얼굴 턱 오른쪽의 대칭적 움직임, 이마 윗부분의 강체적인 움직임 등의 부가 정보를 데이터의 후보정에 활용할 경우 더욱 강건한 결과를 얻을 수 있을 것으로 예상되어 이에 대한 방안을 간구하고 있다. 또한, 얼굴 표면의 주름과 같은 세밀한 변형의 원거리 포착 방법에 대한 방안에도 대해서도 연구가 이뤄질 예정이다.

[표 1] 오차의 평균 및 표준 편차

	Park과 Hodgins[11]	본 논문의 방법
평균오차 (mm)	6.645	3.258
표준편차	3.010	1.719

## 참고문헌

- [1] Brand, M., "Voice Puppetry", In Proceedings of SIGGRAPH 99, pp. 21-28, 1992.
- [2] Deng, Z., Neumann, U., "efase: expressive facial animation synthesis and editing with phoneme-isomap controls", In Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 251-260, 2006
- [3] Kshirsagar, S., Magnenat-Thalmann, N., "Visyllable based speech animation", Computer Graphics Forum, vol. 22(3), pp 632-640
- [4] Chuang, E., Bregler, C., "Mood swings: expressive speech animation", ACM Transactions on Graphics, vol. 24(2), pp. 331-347, 2005
- [5] Joshi, P., Tien, W. C., Desbrun, M., Pighin, F., "Learning controls for blend shape based realistic facial animation", In Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 187-192, 2003
- [6] Pyun, H., Kim, Y., Chae, W., Kang, H. W., Shin, S. Y., "An example-based approach for facial expression cloning", In Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 167-176, 2003
- [7] Ju, E., Lee, J., "Expressive Facial Gestures from motion capture data", Computer Graphics Forum, vol. 27(2), pp. 381-388, 2008
- [8] Pighin, F., Hecker, J., Lischinski, D., Salesin, D., Szeliski, R., "Synthesizing Realistic Facial Expressions", In Proceedings of SIGGRAPH 98, pp. 75-84, 1998
- [9] Ma, W., Jones, A., Chiang, J., Hawkins, T., Frederiksen, S., Peers, P., Vukovic, M.,

## 7. 결론 및 토의

본 논문은 얼굴의 표정과 몸의 동작을 광학식 마커 포착 장비로 동시에 포착할 때 얼굴에 붙인 마커 데이터의 후보정 및 애니메이션 제작의 실제적인 문제들에 대한 해결 방안을 제시했다. 포착 데이터의 문제점으로는 마커의 가림이나 혼돈으로 인한 마커 궤적의 표식화의 문제, 마커의 포착 실패로 인한 데이터 빈 곳 메우기의 문제, 그리고, 마지막으로 데이터의 보정을 위한 노이즈 제거 문제를 다루었으며 이들의 해결 방법으로 한 마커와 그 마커의 주변마커로부터 정의되는 지역 좌표를 활용하여 모든 문제의 해결 방안을 제시하였다. 구현된 결과를 확인해 보기 위해 실제 포착 데이터에 적용하여 그 유용성을 확인하였으며 이로부터 애니메이션을 제작하여 보았다.

향후 연구 과제로는, 본 연구의 방법은 순전한 마커 하나와 그 주변 마커에 대한 정보만을 활용하여 신호처리 방법을 적용하고 있으나, 얼굴에 대

Ouhyoung, M., Debevec, P., "Facial performance synthesis using deformation-driven polynomial displacement maps", ACM Transactions on Graphics, vol. 27(5), article 121, 2008

- [10] Lipman, Y., Sorkine, O., Levin, D., Cohen-Or, D., "Linear Rotation-invariant Coordinates for Meshes", ACM Transactions on Graphics, Vol. 24(3), pp. 479-487, 2005
- [11] Park, S., Hodgins, J., "Capturing and Animating skin deformation in human motion", ACM Transactions on Graphics, vol. 25(3), pp. 881-889, 2005
- [12] Horn, B., "Closed-form solution of absolute orientation using unit quaternions", Journal of the Optical Society of America, vol. A(4), pp. 629-642, 1987



박 상 일 (Park, Sang Il)

1997 연세대학교 기계공학과 학사  
1999 KAIST 기계공학과 석사  
2004 KAIST 전산학과 박사  
2005-2007 미국 카네기멜론 대학 박사후 연구원  
2007 일본 국립총합연구소(AIST) 초빙연구원  
2007-현재 세종대학교 디지털콘텐츠학과 전임강사

관심분야 : 컴퓨터 그래픽스, 캐릭터 애니메이션

---