

모션 캡처의 과거, 현재, 그리고 미래

서울대학교 컴퓨터공학부 이제희

들어가는 말

애니메이션은 오늘날 가장 매력적이며 창조적인 미디어의 형태로 우리에게 가까이 다가서고 있다. 다른 미디어에 대해 애니메이션이 갖는 차별성은 움직임에 담겨있는데 있다. 애니메이터는 풍부한 감정과 개성, 역동성을 지닌 움직임을 창조하기 위해 노력하며, 많은 경우 실제 사람이나 동물, 혹은 자연 현상의 움직임을 관찰하고 이로부터 영감을 얻어 새로운 움직임을 만들어 낸다. 이러한 애니메이터의 작업은 많은 부분 수작업에 의존하며, 사실적인 동작을 만들어 내기 위해서는 오랜 기간의 훈련 과정을 거쳐야 하며 상당한 노력과 시간을 투자해야 한다.

컴퓨터를 이용한 애니메이션 제작은 실제 자연 속의 움직임에 대한 관찰을 보다 직접적이며 자동화된 애니메이션 제작 과정의 일부로 포괄할 수 있게 하였다. 움직이는 점의 3차원 공간 상의 위치를 측정할 수 있는 다양한 센서들이 개발되었고, 이러한 센서를 이용하여 실제 사람의 동작을 정교하게 컴퓨터 속에 재현하는 것이 가능해졌다. 모션 캡처(motion capture)는 넓은 의미에서 실세계에서 일어나는 생물, 혹은 무생물의 움직임을 분석 가능한 데이터로 변환하는 과정을 말한다. 근래에는 좁은 의미에서 실세계의 움직임을 녹화하고 디지털화하여 컴퓨터로 읽어 들이는 일련의 과정을 모션 캡처라 한다.

모션 캡처의 응용 분야는 애니메이션에 국한 되지 않는다. 실제로 의료 분야에서는 정형외과 혹은 재활의학의 측면에서 환자의 걸음걸이나 움직임이 정상적인지를 판단하기 위한 진단의 목적으로, 또한 수술을 통한 보정 이후의 결과를 예측하기 위한 수단으로 모션 캡처를 사용해 왔으며 지속적으로 연구하고 있다. 이 외에도 스포츠 과학, 사람-컴퓨터 인터페이스, 군사, 생체역학, 영화의 특수효과, 인간 공학, 심리학 등의 다양한 분야에서 활용되고 있다. 이 글은 모션 캡처가 발전해 온 과정과 현재의 기술 수준, 그리고 앞으로의 발전 방향을 고찰해 보고자 한다.

모션 캡처의 역사

모션 캡처의 역사는 말의 움직임에 관한 1800년대 말의 논란으로부터 시작되었다. 당시 풍경 사진을 주로 찍던 사진작가 에드워드 머이브리지는 (Eadweard Muybridge) 목장을 소유한 친구로부터 말이 장애물을 넘을 때 순간적이나마 네 발이 모두 공중에 떠있는 순간이 있는지에 관한 질문을 받았다. 말의 움직임이 너무 빨랐기 때문에 육안으로는 순간적인 말의 움직임을 정확히 관찰할 수 없었다. 이 문제를 해결하기 위해 1878년 6월 15일 머이브

릿지는 24대의 스틸 카메라를 일렬로 늘어놓고 일정한 시간 간격으로 셔터를 격발시켜 장애물을 넘는 말의 움직임을 일련의 사진에 담는데 성공한다. 이는 동작 분석을 위한 최초의 모션 캡처이자 이후 활동 사진 및 영화의 발명으로 이어지는 시발점이기도 했다.

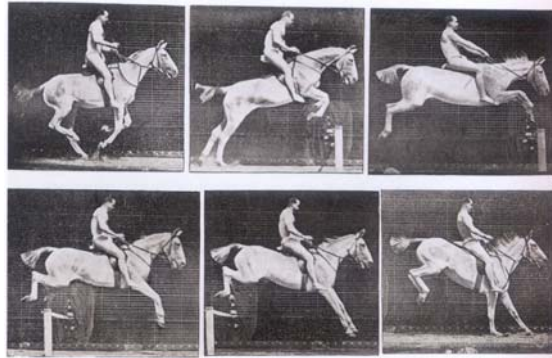


그림 1: 뮤브리지가 촬영한 말의 연속 사진
(<http://www.kingston.ac.uk/Muybridge>에서 발췌)

초창기의 모션 캡처는 컴퓨터의 도움 없이 수작업으로 이루어졌다. 대표적인 예가 1915년 맥스 플레이셔(Max Fleischer)에 의해 고안된 로토스코프(rotoscope)라는 방법으로 애니메이션 제작 시에 실제 사람의 움직임을 연속 사진에 담고 이를 그림판에 투영시켜 애니메이터로 하여금 윤곽선을 따라 그림을 그리도록 한다. 이러한 방식을 사용하면 매우 사실적인 움직임을 애니메이션으로 재현할 수 있어 애니메이션 제작에 폭 넓게 활용되었다. 디즈니의 초창기 애니메이션인 “백설공주”에서 로토스코프가 활용되었고, 최근에는 “아나스타샤”에서 실사 영화와 같은 느낌의 애니메이션을 제작하는데 사용되었다.

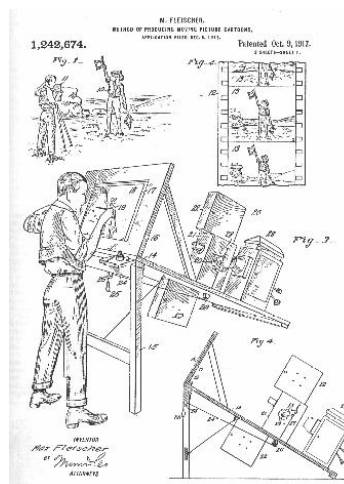


그림 2: 로토스코프는 애니메이터가 그림판에 비춰진 사진을 보며 윤곽선을 따라 그릴 수 있도록 고안되었다

모션 캡처는 이후 군사, 의료 목적으로 꾸준히 활용되어 왔으며, 본격적으로 컴퓨터 애니메이션에 활용되기 시작한 것은 1985년 로버트 아벨(Robert Abel)에 의해 제작된 텔레비전 광고 “브릴리언스(Brilliance)”에 등장하는 여성형 로봇의 애니메이션을 위해서였다. 은빛의 반짝이는 금속 재질과 사실적인 움직임으로 이 로봇 애니메이션은 당시 대단한 관심을 끌었다. 아벨은 로봇의 움직임을 얻기 위해 연기자에게 흰색의 레오타드를 입히고 18개의 주요 관절 부위를 검정색으로 표시한 뒤 연기자의 움직임을 여러 대의 비디오 카메라로 녹화하여 검정색 표시의 3차원 좌표를 계산하였다. 아벨이 시도한 방식은 최초의 광학식(optical) 모션 캡처로 기록되어있다.

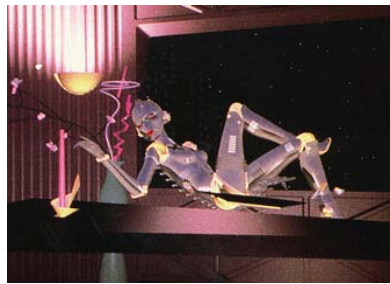


그림 3: 광고 “브릴리언스”의 섹시 로봇 (<http://www.accad.ohio-state.edu/~waynec/history/tree/abel.html>에서 발췌)

1988년 실리콘 그래픽스(Silicon graphics, Inc)와 드그라프-워먼(deGraf-Wahrman, Inc)에 의해 개발된 “마이크 말하는 얼굴(Mike the talking head)”은 얼굴 표정 캡처를 이용하여 실시간 얼굴 애니메이션이 가능함을 보여주었다. 실제 사람 마이크 그리블(Mike Gribble)이 각각의 음소를 발음하는 얼굴을 3차원 스캔하여 다면체(polygon)으로 이루어진 얼굴 모델을 얻고 음소에 해당되는 표정들을 보간(interpolate)하여 사실적인 얼굴 애니메이션을 얻었다.



그림 4: 마이크 말하는 얼굴 (<http://mambo.ucsc.edu/psl/mike.html>에서 발췌)

이후 모션 캡처 기술은 영화 산업을 중심으로 광범위하게 퍼져 나갔으며 최근 개봉된 수없이 많은 영화의 특수 효과를 위해 사용되었다. 모션 캡처가 사용된 영화를 예는 Titanic (1997), Batman and Robin (1997), The Mummy (1999), Gladiator (2000), The

Patriot (2000), The Mummy Returns (2001), Star Wars Episode 1 - the Phantom Menace (2001), Final Fantasy: The Spirits Within (2001), The Enemy at the Gates (2001), Pearl Harbor (2001) 등이 있다. 영화 특수효과 제작에 있어서 모션 캡처의 역할은 점차 증대되고 있으며, 이러한 추세는 당분간 계속될 것으로 보인다.

왜 모션 캡처인가

움직임에 대한 인간의 지각은 때론 놀라울 정도로 정교하다. 특히, 움직임의 미세한 특징으로부터 감정의 기복이나 미묘한 개성의 차이 등을 읽어내는데 인간 지각은 대단히 뛰어나며 이는 애니메이션을 제작하는데 있어 어려움의 원인이 된다. 애니메이터에게 있어 인간 동작의 미세한 부분까지 창조해 내는 일은 대단히 어려우며 뛰어난 재능과 노력을 필요로 한다.

모션 캡처의 가장 중요한 장점은 실제 사람 동작의 미세한 특징들을 컴퓨터로 얻어내어 캐릭터에 적용할 수 있다는 것이다. 사람마다 지닌 제 각각의 스타일, 분위기, 감정 상태, 무게감, 물리적 특성 등 동작을 사실적으로 보이게 만드는 중요한 요소들이 녹화된 동작 데이터에 포함되어 있으며, 이러한 부분이 애니메이터의 작업을 도와 애니메이션 제작 기간을 단축하는데 도움을 줄 수 있다.

무엇을 캡처하는가

일반적으로 인체는 컴퓨터 내에서 트리 구조(tree structure)의 다관절체(articulated figure)로 표현된다. 전신 (full-body) 모션 캡처를 통해 우리가 얻고자 하는 데이터는 루트 (root)의 위치와 방향, 그리고 모든 관절각을 포함한다. (엉덩이에 해당하는 부분을 루트로 지정하는 경우가 많으나 반드시 그런 것은 아니며 사실상 신체 어떤 부위도 루트가 될 수 있다) 이와 더불어, 다관절체 연결 구조와 각 부위의 길이 등 캡처하고자 하는 대상 신체에 관한 정보를 필요로 한다. 이러한 정보는 특정 자세로부터 (일반적으로 양팔을 좌우 수평으로 벌리고 다리는 어깨 넓이로 벌린 자세) 반자동적으로 계산하기도 하고, 아주 높은 정밀도를 원하는 경우에는 버지니아 칼리퍼스(virginia calipus)와 줄자 등의 측정 도구를 이용하여 신체 부위 별로 수작업으로 측정하기도 한다.

관절체에 관한 정보 외에 피부의 변형에 관한 데이터를 같이 캡처하기도 한다. 이 경우 신체는 부드럽게 변형 가능한 물체로 간주되며 근육의 팽창 정도나 피부의 주름까지 시간에 따라 변화하는 세부적인 형상을 모두 녹화하여 저장한다. 상황에 따라서는 연기자가 입고 있는 옷의 움직임을 캡처하거나, 얼굴이나 손 등 특정 부위만을 확대하여 캡처하기도 한다.

모션 캡처의 또 다른 대상은 연기자가 들고 있는 기구나 장비, 예를 들면 공이나 총기, 막대기, 가방 등이 될 수 있다. 이와 더불어 연기자가 움직이는 환경도 캡처의 대상이 된다. 예를 들어, 정글짐에서의 동작을 녹화한다면 정글짐의 주요 부위에 표식을 부착하여 연기자

의 동작과 함께 캡처한다. 이 데이터는 정글짐을 컴퓨터로 모델링하고 이를 동작 데이터와 일치시키는데 활용될 수 있다.

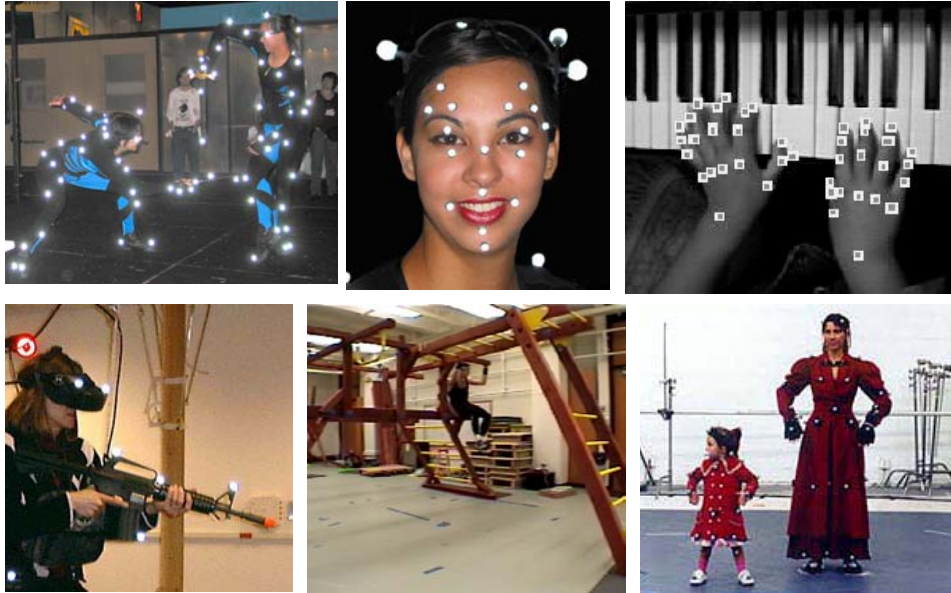


그림 5: (좌상) 전신 동작 캡처. (상중상) 얼굴 표정 캡처. (우상) 손 캡처. (좌하) 기구를 움직임을 같이 녹화하는 예. (하중상) 정글짐에서의 모션 캡처. (우하) 영화 “타이타닉”의 제작과정에서 옷을 입은 채 캡처하는 모습. 인체 모형을 부드럽게 변형 가능한 물체로 간주하고 녹화된 표식의 위치를 이용하여 모형을 변형한다. (좌상과 좌하, 상중상의 그림은 <http://www.MotionAnalysis.com>에서 발췌)

어떻게 캡처하는가

모션 캡처를 위해서 3차원 공간 상의 위치와 방향을 측정하기 위한 다양한 방법들이 개발되어 왔다. 대표적인 방법으로는 비디오 카메라를 이용한 광학식 (optical) 방법과 전자기장을 이용하는 전자기식 (electro-magnetic) 방법이 있다. 그 외에도 기계식과 음향을 이용한 방법, 그리고 광섬유(optical fiber)를 이용한 방식들이 연구되어왔다. 이러한 여러 방식들은 서로 다른 장단점을 가지며 응용 분야에 따라 적절한 방식이 선택된다.

광학식 시스템은 특정 주파수만 받아들이는 필터를 갖춘 여러 대의 비디오 카메라를 이용한다. 각 카메라의 전면에는 빨간색 가시 광선이나 적외선 광원을 갖추고 있어 이 빛이 연기자의 몸에 부착된 표식에 반사되어 돌아오면 이를 비디오 카메라로 녹화한다. 표식은 일반적으로 조그만 구형으로 빛을 반사하는 소재로 둘러싸여 있다. 따라서, 비디오 카메라에서 필터를 통해 얻는 이미지는 표식을 흰색으로 나머지는 모두 검정색으로 기록한다. 한 개의 표식을 두 대 이상의 카메라에서 동시에 관찰하면 표식의 3차원 좌표를 삼각 함수에

의해 계산할 수 있다. 다만, 표식이 카메라로부터 가려질 수 있기 때문에 일반적으로 적게는 6대, 많게는 24대 정도의 카메라를 이용하여 가려진 부위를 최소화한다.

광학식 모션 캡처는 일반적으로 후처리 과정을 통해 데이터를 손질하는 과정을 필요로 한다. 이는 모든 표식이 카메라에서는 다만 흰 점으로 보일 뿐 어느 표식이 어디에 부착된 것인지를 즉각적으로 판단할 수 있는 방법이 없기 때문이다. 따라서, 사용자가 초기 상태에서 각각의 흰 점이 어느 부위의 표식인지를 지정하면, 이후 점들의 이동 경로와 속도로부터 연속된 움직임을 따라 가는 방식으로 3차원 경로를 찾아낸다. 이와 같은 추적 (tracking) 알고리즘은 완전히 자동화되기 어려우며, 특히 표식이 가려졌다가 다시 화면에 나타날 때 오작동하기 쉽다. 따라서, 최근에는 표식들이 스스로 발광하도록 하는 능동적 광학식 (active optical) 시스템이 주목 받고 있으며 활발히 연구되고 있다.

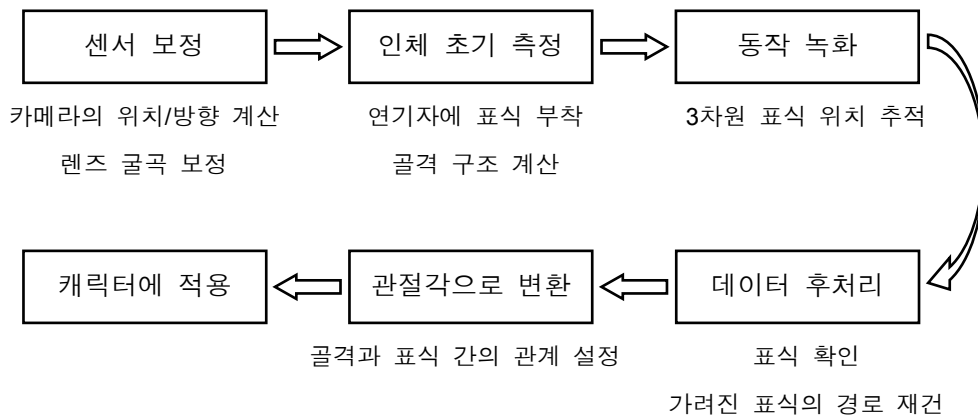


그림 6: 광학식 모션 캡처 과정

전자기식 시스템은 트랜스미터(transmitter)로 하여금 공간 상에 전자기장(electromagnetic field)을 형성하도록 하고 연기자의 몸에 부착된 센서들이 전자기장으로부터 스스로의 위치와 방향을 계산할 수 있도록 한다. 이 방식은 가려짐이나 표식 구분의 문제가 없어 실시간 시스템에 주로 활용된다. 다만, 연기자의 입장에서 배터리와 센서, 그리고 많은 전선을 몸에 두르고 있어야 하기 때문에 움직임에 제약이 많고 광학식 장비에 비해 정밀도가 떨어지며 캡처 영역이 협소한 단점을 가지고 있다. 근래에는 여러 개의 트랜스미터를 설치하여 캡처 영역을 넓히는 방안이 도입되고 있다. 전자기식 장비의 또 다른 문제점은 전자기장이 주변의 전도성 물질에 의한 간섭에 취약하다는 점이다. 대부분의 건물이 벽과 바닥, 천정에 철골을 지니고 있어 간섭을 피하려면 이들로부터 충분히 떨어지도록 캡처 장비를 설치해야 한다.

실시간 캡처를 위한 대안으로 기계식 시스템이 사용되기도 한다. 연기자는 기계 골격 (mechanical skeleton)을 몸에 묶어 고정하고 움직이면 기계 골격의 관절에 부착된 센서로부터 동작을 인식한다. 기계식은 대단히 정교하고 빨라서 초당 500 프레임 이상 녹화할 수

있으며 캡처 영역도 거의 제약을 받지 않는다. 그러나, 연기자의 움직임은 기계 골격으로 인해 전자기식의 경우보다 더 심하게 제약된다.

음향식 캡처 시스템(acoustic motion capture)은 딸각이는 소리를 내는 두 개 이상의 음원과 이를 감지하는 센서들로 구성된다. 연기자의 몸에 부착된 센서들은 소리가 전달되는 시차를 측정하여 음원까지의 거리를 계산하고, 다시 이로부터 3차원 좌표를 얻는다. 이 방식은 가려짐의 문제는 없으나, 캡처 영역이 좁고 센서의 개수가 제한되며 음향 센서를 위한 전자적 장비를 연기자가 짊어져야 하므로 근래에는 거의 사용되지 않는다.

광섬유(optical fiber)는 구부러지는 정도를 섬유 내에 빛이 투과되는 정도로부터 계산할 수 있다. 따라서, 광섬유를 인체의 관절 부위에 부착하면 관절각을 측정할 수 있다. 이러한 방식은 주로 데이터 글로브(data gloves) 등의 형태로 손의 움직임을 얻는데 사용되며, 최근에는 광섬유를 이용한 전신 모션 캡처도 시도되고 있다. 이 밖에도 표식이나 센서를 필요로 하지 않는 컴퓨터 비전을 통한 캡처 방식이 주로 얼굴 표정 인식을 위해 사용되고 있다. 전신 동작 캡처를 위한 비전 기술도 연구되고 있으나 기술적인 난해함으로 인해 실용화되지는 못하고 있다.

모션 캡처와 애니메이션

녹화된 동작 데이터를 캐릭터에 적용하여 애니메이션을 제작하기 위해서는 녹화된 데이터를 그대로 사용하기 보다는 추가적인 과정을 거쳐 수정하고 변형해야 하는 것이 일반적이다. 동작 데이터를 수정해야 하는 이유는 여러 가지가 있다. 앞에서 언급한 바와 같이 어떠한 방식의 모션 캡처 시스템도 불완전하며 따라서 얻어지는 데이터도 측정 오류와 잡음으로 인해 손상되어 있다. 이런 문제가 때론 단순한 저주파 필터(low-pass filter)를 이용하여 잡음을 제거함으로써 해결되기도 하지만 많은 경우 구조적인 오류로 인해 더 복잡한 해결책을 요구한다.

동작 데이터가 완전하다 가정하더라도 데이터를 수정해야 할 이유는 여전히 남아있다. 모션 캡처 장비가 아직 대중화되지 않았고 모션 캡처를 위해서는 많은 비용이 소요되기 때문에 필요한 데이터를 그때그때 녹화하기 보다는 많은 데이터들을 미리 녹화해 두고 이를 재사용하기도 한다. 이 경우 캐릭터의 크기가 제각각 다르고 애니메이션 내의 환경도 달라질 수 있으므로 데이터를 주어진 캐릭터와 환경에 맞도록 변형해야 하며, 이를 모션 리타겟팅(motion retargeting)이라 한다.

애니메이터가 의도적으로 동작 데이터를 수정, 편집하기도 한다. 예를 들면, 발차기 동작을 더 높이 혹은 낮게 차도록 과장하거나 변형하기도 하고, 기쁜 동작과 슬픈 동작을 보간하여 다양한 스펙트럼의 감정 상태를 만들어내기도 한다. 동작의 의도, 감정 상태, 분위기, 과장 정도 등 다양한 축으로의 변형을 고려할 수 있으며, 따라서 애니메이터의 입장에서는 정교한 모션 캡처 시스템 못지 않게 동작 데이터를 자유롭게 편집하기 위한 편리한 소프트웨어 편집 도구를 필요로 한다. 이에 대한 연구가 컴퓨터 그래픽스 분야에서 활발히 이루어

지고 있다.

모션 캡처의 미래

모션 캡처 하드웨어에 대해 크게 두 가지 측면에서의 발전 방향을 기대할 수 있다. 첫째는 광학식 모션 캡처에 있어서 표식을 구분하고 가려지는 문제를 해결하기 위한 능동적 광학식 (active optical) 시스템의 개발이다. 가장 간단한 방법으로는 표식들이 시간차를 가지고 돌러가면서 발광하도록 하여 표식을 구분하는 방식이다. 이 경우 예견되는 어려움은 연기자의 몸에 전원과 발광을 위한 비교적 복잡한 전자적 장비를 부착해야 한다는 점이다. 이러한 장비는 연기자가 움직이는데 불편을 초래하며 부상의 요인이 되기도 한다. 또 다른 문제는 초당 녹화할 수 있는 프레임의 수가 줄어들 수 밖에 없다는 점이다. 최근에는 초당 1000 프레임까지 녹화할 수 있는 고속 카메라가 속속 등장하고 있지만 50개의 표식을 순서대로 발광시켜 캡처할 경우 불과 초당 20 프레임 밖에 얻을 수 없다. 그 밖에도 다양한 주파 대역의 전자파나 LED를 이용하는 방법들이 연구되고 있으나 아직까지는 캡처 가능한 영역이 좁거나 센서의 개수가 제한되는 등의 문제점을 가지고 있어 기존의 수동적 광학식 시스템을 대체하지 못하고 있다.

또 다른 발전 방향으로 저가의 대중적 모션 캡처 하드웨어 개발을 생각할 수 있다. 현재 광학식 장비는 매우 고가이며, 비교적 저렴한 전자기식 장비도 대중화되기에는 너무 비싸다. 장기적으로는 전자기식 장비의 경우 30만원 이하의 매우 저렴한 장비를 개발하는 것도 가능할 것이며, 궁극적으로는 현재 게임 콘솔에서 사용되는 조이스틱을 대체하는 게임 입력 장치의 역할을 하는 것도 가능하리라 기대된다. 모션 캡처의 대중화는 미래의 비디오 게임이 지금보다 한 단계 발전을 이루어 역동적이며 몰입감을 주는 매체로 변화해 가는데 일조할 것으로 기대된다.

장기적으로 볼 때 우리가 상상하는 이상적인 모션 캡처 시스템은 축구장 넓이의 공간에서 수 백 명의 사람들이 어우러져 움직이는 가운데, 개개인의 신체에는 표식이나 센서를 전혀 부착하지 않거나, 혹은 최소한의 표식 만을 부착하여 움직이는데 불편을 주지 않도록 하며, 전 과정이 자동화되어 실시간에 동작을 녹화할 수 있을 것이다. 이와 같은 이상적인 시스템이 개발된다면 월드컵이 진행되는 동안 열리는 모든 경기를 모션 캡처하여 개개 선수의 움직임을 분석하고 중계하는 것도 가능해 질 것이다. 이는 모션 캡처가 현재의 스튜디오에서 벗어나 우리의 실생활 속에 참여하며 영향을 줄 수 있음을 의미한다.

맺는 말

모션 캡처 데이터의 응용 기술에 관해서는 다양한 범주의 발전 방향을 예견할 수 있다. 특히, 애니메이션에의 응용을 위해서는 모션 캡처를 성공적으로 활용하지 못 했던 사람들의 견해를 통해 어떤 연구가 필요한 지 예측해 볼 수 있다.

“모션 캡처와 관련된 중요한 문제들은 중량의 배분, 무게감, 그리고 과장에 대한 것이다. 연기자들은 만화 캐릭터와 같이 과장된 동작을 연기할 수 없고, 신체 비례가 다른 캐릭터에 적용되었을 때 중량의 이동이 결코 좋게 보이지 않는다.” (에릭 다넬, “개미”의 공동감독)

“사람의 동작을 사람과 다른 신체 비례의 캐릭터에 적용하기 힘들다. 왜냐하면, 모션 캡처로부터 얻을 수 있는 가장 중요한 것이 무게 이동과 미세함, 그리고 인체의 균형을 유지하는 움직임이다. 만약 신체 비례가 바뀌면 모든 것을 버려야 할 것이다.” (리처드 추앙, PDI 부사장)

“우리가 모션 캡처를 사용하지 않은 이유는 간단하다. 사람으로부터 녹화한 동작 데이터는 우리가 찾고 있던 도마뱀 같은 움직임을 주지 못 했다. 모션 캡처는 고질라의 거대한 중량감도 반영하지 못 했다. 우리는 키프레임 실험에서 우리가 원하는 고질라의 동작이 거대한 중량감을 유지하면서도 우아하고 재빨라야 한다는 것을 알아냈다. 어떤 연기자도 우리가 원하는 결과를 주지 못 했다.” (카렌 콜리카스, “고질라”의 특수 효과 전문가)

이들은 공통적으로 애니메이션의 캐릭터는 현실적이기보다는 만화적으로 과장되어있고, 실제 사람과는 신체 비례의 측면에서 많이 다르다는 점을 지적하고 있다. 50에서 100 킬로그램의 몸무게를 갖는 연기자가 1 그램도 채 되지 않는 개미나 수십 톤에 이르는 공룡과 같이 움직이기는 물리적으로 불가능하다. 따라서, 실제 사람이 펼쳐 보일 수 없는 움직임은 모션 캡처를 통해 얻을 수 없으며 이 문제를 해결하기 위한 방법이 필요하다. 이 외에도 대량의 모션 데이터를 자동적으로 처리하고, 관리하고, 검색하여 비디오 게임과 같은 실시간 대화형 시스템에 활용하는데 관련된 많은 연구가 진행 중에 있다.

지금까지 모션 캡처의 전반에 대해 간략히 고찰해 보았다. 모션 캡처가 갖는 가장 중요한 장점은 현실 세계의 움직임이 갖는 사실감을 가상 세계로 끌어들이 수 있다는 점이다. 이는 모션 캡처가 갖는 중요한 단점이기도 하다. 결국, 녹화된 움직임은 현실 세계가 갖는 물리적, 생물학적 제약을 공유하며 상상력을 구속한다. 미래의 모션 캡처 응용 기술은 애니메이션이나 디자이너의 상상력을 어떻게 반영할 것인가에 초점을 맞추고 있으며, 이와 관련된 다양한 연구 주제가 지금 우리 앞에 펼쳐져 있다.