

데이터 기반 애니메이션과 기계학습

서울대학교 | 이 제 희

1. 서론

살아있는 생명체의 움직임, 특히 인체 동작은 물리적 법칙, 생체 구조, 생태 환경, 각 개체의 심리, 개성, 분위기, 스타일 등 다양한 요소에 의해 결정된다. 이러한 복합적인 요인들로 인해 컴퓨터를 이용하여 생명체의 움직임을 생성하는 문제는 대단히 어려우면서도 학문적 연구 의욕을 자극하고 있다.

최근 컴퓨터 그래픽스 분야에서 각광 받고 있는 데이터 기반 애니메이션은 실제 생명체의 움직임을 녹화하여 대량의 데이터베이스를 구축하고, 이로부터 새로운 움직임 데이터를 생성한다. 데이터 기반 애니메이션 기술이 개발되기 이전부터 널리 사용되던 물리 기반 시뮬레이션은 뉴턴의 운동 방정식으로, 절차적 애니메이션은 프로그래머가 정의한 몇몇 규칙으로 운동의 원리를 규정했다. 이에 반해 데이터 기반 기술은 모션 데이터베이스로부터 대상 개체가 움직이는 원리를 유추하여 실제 사람과 같은 자연스러운 움직임을 재현한다. 모션 데이터베이스가 충분히 많은 데이터를 저장하고 있다면 이로부터 자유롭게 원하는 새로운 동작을 만들어 낼 수 있다.

데이터 기반 접근 방식은 이미 다른 분야에서도 그 효용성이 입증된 바 있다. 예를 들면, 1970년대 유행하던 전자 음향 합성기(sound synthesizer)는 파동 방정식(wave equation)에 기반하여 원하는 소리를 합성하고자 했으나, 1980년대 이후 샘플링된 소리를 혼합하거나 합성하는 데이터 기반 기술이 주류를 이루고 있다. 그래픽스 렌더링 분야에서도 비슷한 예를 볼 수 있다. 1980년대까지 컴퓨터에서 3차원 영상을 얻기 위해서는 광학(optics) 연구에서 비롯된 렌더링 방정식(rendering equation)을 실용적으로 단순화한 방법들이 대부분이었으나, 1990년대 이후 대량의 사진이나 비디오로부터 임의의 시점과 조명 조건에서의 영상을 얻고자 하는 이미지 기반 렌더링(image-based rendering) 연구가 활발히 진행되고 있다.

데이터 기반 접근 방식은 기본적으로 대량의 데이터를 효과적으로 분석, 저장하고, 이를 이용하여 움직임에 대한 일반화된 모델을 얻거나 혹은 변화된 조건에서 새로운 데이터를 합성하는 방법을 필요로 한다. 이러한 데이터 기반 접근 방법은 학습 데이터로부터 모델을 유추하는 기계학습의 목적과 부합한다. 데이터 기반 접근 방법과 기계학습의 유사성은 이미 개발된 다양한 기계학습 기술들을 애니메이션 문제에 적용하고자 하는 시도로 이어지고 있다.

기계학습 기술을 생명체의 움직임 데이터에 적용하기 위해서는 몇 가지 넘어야 할 과제가 있다. 먼저, 동작 데이터는 많은 자유도를 가지므로 소위 “차원의 저주”라 불리는 차원이 올라감에 따라 계산 시간이나 저장 공간이 기하급수적으로 증가하는 문제를 직면하고 있다. 일반적으로 컴퓨터 그래픽스에서 사용하는 인체 모델은 50에서 100개의 자유도를 갖고, 생체공학에서 사용하는 보다 정교한 근골격(musculoskeletal) 모델은 200 이상의 자유도를 갖는다.

기계학습 기술이 성공적으로 적용되기 위해서는 문제 영역의 내재적, 선형적 구조를 최대한 활용해야 한다는 점은 널리 알려져 있다. 생명체의 움직임은 물리적, 기구적, 생체적, 생리적, 심리적, 사회적 요인 등 다양한 요소가 복합적으로 작용하고 있어 그 내재적 구조를 단순한 모델로 표현하기 어렵다. 특히, 몸체가 주변 물체나 바닥을 뚫고 들어가지 않아야 한다거나, 동작이 물리 법칙을 따르며 균형을 유지해야 한다는 등 반드시 만족되어야 하는 몇몇 필수 제약조건(hard constraints)은 일반적인 기계학습 기법에서 만족시키기 대단히 어렵기에 특화된 새로운 방법의 개발이 필요하다.

이 글에서는 최근 기계학습이 컴퓨터 애니메이션 분야에 성공적으로 적용된 예를 살펴보고 앞으로의 발전 방향에 대해 전망해 보고자 한다. 먼저, 애니메이션 분야에서 데이터 기반 접근 방법이 파생된 배경을 설명하고, 애니메이션 분야에서 활용되는 기계학습

기법의 예로서 마코프 모델(Markov model), 은닉 마코프 모델(hidden Markov model), 선형 동적 모델(linear dynamic model), 차원 감소(dimensionality reduction), 회귀 분석(regression), 강화 학습(reinforcement learning) 등의 적용 예를 알아본다.

2. 배경

영화 특수효과나 애니메이션에 등장하는 자연스럽고 매력적인 캐릭터의 움직임은 최근까지도 많은 부분 애니메이션의 수작업과 예술적 영감에 의존하여 제작되고 있다. 애니메이터들의 작업 방식은 과거 만화 영화를 만들던 셀 애니메이션 방식에서 크게 벗어나지 못하고 있다. 애니메이터가 몇몇 주요 장면(키프레임)에서 캐릭터의 자세를 지정하면 컴퓨터는 중간 프레임의 자세들을 부드럽게 보간하는 제한된 역할만을 수행한다. 더 높은 품질의 애니메이션을 제작하기 위해서는 더 촘촘하게 키프레임을 지정해야 하고, 극단적으로는 초당 30 프레임 전체에서 수작업으로 캐릭터의 자세를 미세 조정하는 경우도 드물지 않다.

생명체의 자연스러운 동작을 자동적으로 만들어 내기 위한 연구는 지난 30년간 컴퓨터 그래픽스에서 가장 중요한 문제 중 하나로 인식되어 왔다. 많은 연구들이 물리 시뮬레이션이나 질차와 규칙으로 인체 동작을 표현해 왔고, 상당한 진척을 이루기도 하였다. 이러한 접근 방법들이 많은 문제들에 대해 성공적인 결과를 보여주었음에도 불구하고, 간단한 수식 혹은 규칙만으로는 인체 동작이 갖는 미세한 특성과 다양한 변화, 정량화하기 어려운 미묘한 특질을 충실히 표현하기 어렵다는 공통적인 문제점이 있다.

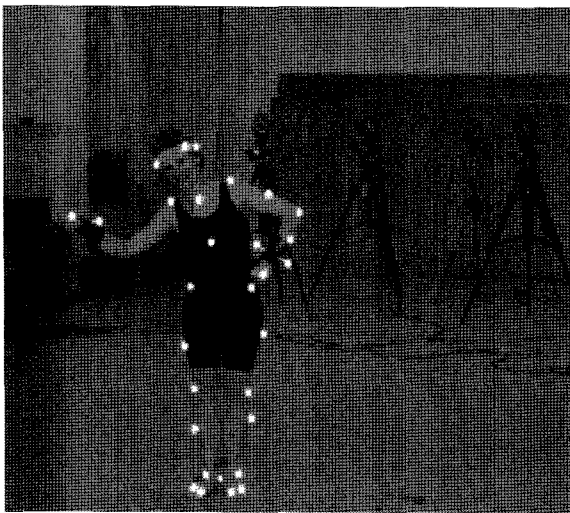


그림 1 광학식 모션 캡처

90년대 중반 이후 본격적으로 보급되기 시작한 모션 캡처 시스템은 데이터 기반 애니메이션 연구를 본격화하는 계기가 되었다. 초기에는 전자기 필드 트랜스미터(electromagnetic field transmitter)와 센서를 갖춘 전자기식 시스템이나 고속 고해상도 비디오 카메라와 반사 마커를 이용한 광학식(optical) 시스템이 주류를 이루었고(그림 1), 최근에는 가속도계(accelerometer), 자이로스코프(Gyroscope), 능동형 마커, 광학 섬유(optical fiber) 등 다양한 센서 기술을 적용한 모션 캡처 시스템이 속속 개발되고 있다. 수년 내에 10~20만원 수준에 구입하여 컴퓨터 마우스나 게임 콘솔 컨트롤러를 대체할 수 있는 저가형 모션 캡처 시스템이 상용화 될 수 있을 것으로 기대된다.

3. 기계학습 활용

3.1 모션 그래프(Motion Graph)

데이터 기반 애니메이션에서의 가장 기본적인 기술은 대량의 모션 데이터로부터 적절한 데이터를 부분적으로 선택하고 이를 새로운 순서로 연결하여 원래 데이터베이스에는 없는 새로운 동작 시퀀스를 생성하는 것이다. 이러한 작업을 효율적으로 수행하기 위해 모션 데이터를 방향성 그래프로 표현하고 저장한다[1, 2]. 모션 그래프의 정점은 대상체의 자세를 표현하고, 에지는 자세 간의 전이 가능성을 나타낸다. 모션 데이터가 입력으로 주어지면, 그 안의 모든 프레임들 중에 서로 자세와 속도가 비슷한 프레임 쌍을 찾아내고 이들 간의 전이 에지를 연결함으로써 모션 그래프를 구성할 수 있다(그림 2). 그래프 내의 정점과 에지를 지나는 임의의 경로는 입력 데이터의 조각들을 이어 붙인 애니메이션 시퀀스와 대응되어 효과적으로 새로운 움직임 데이터를 생성할 수 있다.

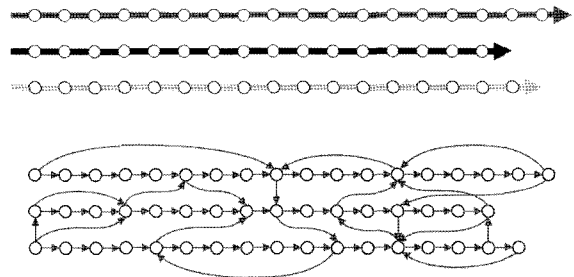


그림 2 모션 그래프 구성([1]에서 그림 발췌). (상단) 입력으로 주어진 모션 데이터는 모션 시퀀스들로 이루어져 있고, 각 모션 시퀀스들은 시간에 따라 선형적으로 변하는 모션 캡처 연기자의 자세들을 담고 있다. (하단) 모션 데이터에서 유사한 프레임 쌍들을 찾아내고 전이 에지를 연결하면 방향성 그래프로 프레임들 간의 연결 구조를 표현할 수 있다.

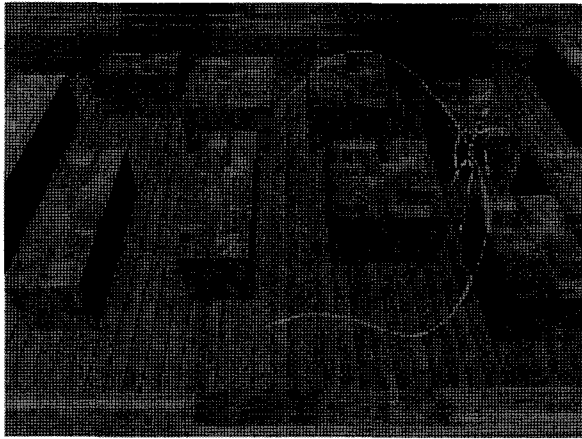


그림 3 모션 그래프와 스케치 인터페이스를 이용한 캐릭터 동작 제어([1]에서 그림 발체). 캐릭터는 주어진 모션 데이터베이스 내의 동작 데이터를 조합하여 사용자가 지시한 이동 경로를 따라 움직인다.

각 예지에는 앞뒤의 프레임들이 얼마나 비슷한가에 따라 전이 적합성을 계산할 수 있고, 이는 다시 전이 확률로 전환될 수 있다. 각 예지에 전이 확률을 포함한 그래프는 마코프 모델(Markov model)을 표현하게 되므로, 기존의 마코프 모델로 할 수 있었던 많은 일들을 애니메이션 시퀀스에 대해서도 할 수 있다. 예를 들면, 빈 공간에서 캡처한 전후좌우 다양한 보행 동작으로부터 가상 환경의 캐릭터가 사용자의 지시에 따라 미로 속에서 충돌을 피하며 움직이도록 조정할 수 있다. 또한, 좁은 곳을 통과할 때는 가상 캐릭터가 자동적으로 옆 걸음을 선택하는 등 지능적 행동을 수행한다(그림 3).

3.2 은닉 마코프 모델과 선형 동적 모델

모션 데이터의 구조를 분석하고 새로운 동작 데이터의 합성을 위해 은닉 마코프 모델(hidden Markov model)이 사용된 바 있다. 예로서, 스타일 머신(style machine)은 가우시안 혼합 모델(Gaussian mixture model)로 모션 데이터의 분포를 근사하고, 추가적으로 유사한 동작들 내의 변화를 스타일 파라미터로 표현하였다[3].

동작 데이터가 갖은 시간성을 잘 표현하기 위해 가우시안 모델 대신 선형 동적 시스템(linear dynamic system)을 이용한 방법도 제시된 바 있다[4]. 동작의 큰 변화는 선형 동적 시스템 간의 전이를 통해 표현하고, 동작의 미세한 변화는 선형 동적 시스템의 파라미터를 조정하는 방식의 두 단계로 동작을 합성할 수 있다.

3.3 차원 감소(Dimensionality Reduction)

대량의 모션 데이터로부터 얻어진 모션 그래프나 은닉 마코프 모델 등은 다양한 종류의 동작들과 동

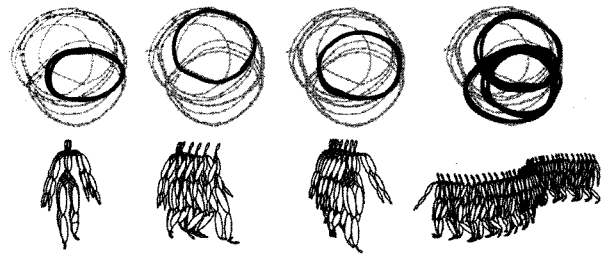


그림 4 (상단) 다양한 회전각도로 걷는 동작 데이터를 2차원 공간에 투영하였다. (하단) 애니메이터는 투영된 평면 위에 경로를 스케치하는 방식으로 원하는 동작을 생성할 수 있다([5]에서 그림 발체).

작들 간의 연결 구조를 표현할 수 있으며, 때로는 유사한 동작들의 집합을 매개 변수 모델로 표현하기도 한다. 애니메이터의 입장에서는 이러한 행위 모델 내의 수많은 종류의 동작들과 이들 간의 복잡한 구조를 직관적으로 이해하기 어려워 실제 사용하기에 제약이 따른다. 애니메이터의 관점에서는 편리한 사용자 인터페이스를 통해 직관적이고 직접적인 방식으로 행위 모델에 접근할 수 있어야 한다.

MDS(Multidimensional Scaling), Isomap, LLE(locally linear embedding) 등의 차원 감소 기법들은 고차원의 복잡한 구조체를 저차원 공간에 투영하는 방법을 제시한다. 차원 감소 기법은 모션 그래프를 낮은 차원에서(주로 2차원) 투영하여 연결 구조를 직접적으로 가시화할 수 있게 해준다(그림 4). 또한, 투영된 2차원 공간은 애니메이션 생성 및 편집을 위한 편리한 사용자 인터페이스 공간으로 활용될 수 있다[5].

3.4 강화 학습(Reinforcement Learning)

모션 캡처 데이터를 이용한 실시간 애니메이션은 컴퓨터 게임이나 가상현실 등 대화형 응용 분야에서 효과적으로 사실적인 인간형 캐릭터를 생성하기 위한 주요 기술이다. 사람의 상위 단계 행위를 시뮬레이션 하기 위해서는 단순히 모션 데이터베이스에서 당장 필요한 데이터를 근시안적으로 선택하는 것으로는 부족하며, 행위의 종류에 따라 미래를 예측하고 검색하기 위한 상태 공간 검색(state space search) 및 행위 계획(behavior planning) 과정을 필요로 한다. 모션 그래프는 상태 공간 검색 과정에서 매 순간 선택 가능한 동작들을 제시하는 역할을 수행할 수 있다. 일반적으로 모션 그래프가 크고 많은 전이 예지를 가질수록 애니메이션 되는 캐릭터는 더 유연하고 풍부한 움직임 보일 수 있다. 반면, 모션 그래프의 연결 구조가 복잡해질수록 상태 공간 검색을 위한 계산 시간과 공간은 기하급수적으로 증가하게 된다.

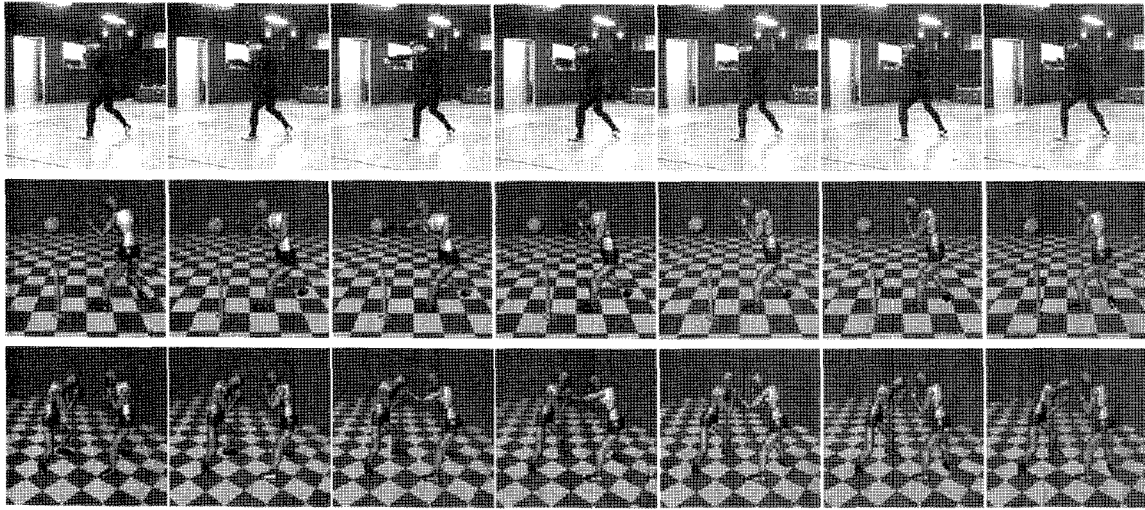


그림 5 강화 학습을 이용한 가상 권투 선수 동작 제어([6]에서 그림 발췌). (상단) 실제 권투 선수의 모션 캡처. (중단) 강화 학습을 이용해서 가상 권투 선수가 움직이는 목표를 추적하고 가격할 수 있도록 학습. (하단) 두 명의 가상 권투 선수가 학습된 행위 모델을 이용하여 가상의 권투 경기

강화 학습은 모든 상황에서 가능한 모든 동작들이 주어진 목표에 대해 얼마나 적합한가를 미리 테이블 형태로 계산해 줌으로써 캐릭터가 실시간 애니메이션 될 수 있도록 한다[6]. 예를 들면, 유명한 권투 선수가 shadow boxing하는 동작을 30분 정도 모션 캡처하고, 이 동작 데이터를 사용해서 스스로 목표를 찾고 추격해서 가격하는 가상의 권투 선수를 생성할 수 있다. 강화 학습을 이용하면 가상 권투 선수는 목표의 상대적 위치와 자신의 현재 자세로부터 최적의 다음 단계 동작을 학습된 테이블로부터 손쉽게 찾아낼 수 있다(그림 5).

3.5 회귀 분석(Regression)

회귀 분석은 통계학으로부터 유래되어 주어진 데이터를 일반화하기 위해 사용되는 가장 일반적인 방법이다. 대표적으로 가우시안 프로세스(Gaussian process)의 여러 가지 변형들과 다양한 LWR(Locally Weighted Regression) 기법들이 애니메이션 분야에서 활용되고 있다. 데이터 기반 역운동학(inverse kinematics)과 동작 보간(motion interpolation) 등이 회귀 분석 응용의 좋은 예이다.

역운동학은 인체와 같은 관절체의 자세를 만들어 내기 위해 사용하는 기본적인 기술이다. 인체 몇몇 부위의 공간 상에서 위치를 사용자가 지정하면 역운동학은 조건에 부합하는 자세를 자동적으로 찾아준다. 일반적으로 역운동학 문제를 해결하기 위해 자코비안(Jacobian) 형렬과 수치적 최적화 기법이 자주 사용되었으나, 상황에 따라 사실적이지 않은 자세가 나오

는 문제점이 있었다. 데이터 기반 역운동학은 주어진 동작 데이터 내의 자세들을 분석하여 사용자가 지정한 제약조건을 만족시키면서도 가능한 자연스러운 자세를 만들어 낸다(그림 6).

회귀 분석 기법은 여러 동작 데이터를 보간하여 하나의 매개 변수 모델을 구성할 수 있다[8]. 예를 들면, 걷는 동작의 경우 다양한 속도와 회전각의 걷는 동작 데이터 집합을 보간하여 속도와 회전각을 매개 변수로 갖는 모델을 구성할 수 있다. 이 모델은 가상 캐릭터가 자유롭게 걷는 속도와 방향을 바꿀 수 있게 해준다.

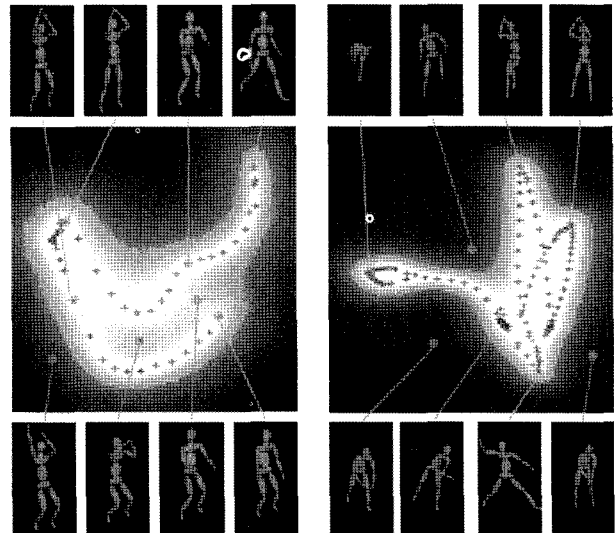


그림 6 스타일 기반 역운동학([7]에서 그림 발췌). 가우시안 프로세스의 일종인 SGPLVM(Scaled Gaussian Process Latent Variable Model)을 이용하여 자세 공간을 모델링하였다.

4. 향후 전망

데이터 기반 애니메이션은 컴퓨터 게임, 애니메이션 제작, 영화의 특수효과 등 다양한 그래픽스 응용 분야에서 빠르게 그 적용 범위를 넓혀가고 있다. 실제 응용 분야에서의 데이터 기반 기술은 인체의 사실적인 움직임을 적은 비용과 노력으로 빠르게 만들어 낼 수 있다.

데이터 기반 애니메이션은 아직 많은 도전적인 과제를 남겨두고 있다. 특히, 아래 문제들은 실제 응용 분야에서 활용 가능성이 높으나 아직 해결되지 않고 있다.

물리 기반 제어 및 시뮬레이션(Physically based control and simulation): 현재의 인간형 로봇들은 기계적이며 전형적인 스타일의 움직임만이 가능할 뿐, 실제 사람과 같이 다양하고 자연스러운 스타일을 보여주지 못하고 있다. 동작 데이터를 이용해서 자동적으로 제어기를 만들어 낼 수 있다면 인간형 로봇들도 보다 자연스럽게 다양한 동작들을 수행할 수 있을 것이다.

다중 에이전트 상호작용(Multi-agent interaction): 현재까지 개발된 대부분의 데이터 기반 애니메이션 기술은 한 사람의 개별적인 움직임을 캡처하고 재활용하는데 초점을 두고 있었다. 여러 명의 사람들이 서로 협력하며 물리적으로 접촉하는 경우 움직임의 모델링이 훨씬 더 어려워지며 흥미로운 문제들을 발생시킨다. 예를 들면, 다중 연기자의 움직임과 상호작용의 연관 관계를 동시에 캡처해야 하며, 상호작용을 모델링하여 재활용할 수 있도록 해야 한다.

군중 행위 시뮬레이션(Crowd behavior simulation): 행위자의 숫자가 많아져서 군중을 이루면 사회적, 군중 심리적 요인들이 개별 행위자의 움직임을 결정하는데 더 중요하게 부각된다. 군중 행위의 캡처와 분석, 재구성은 모두 흥미로운 향후 연구 과제이다.

참고문헌

[1] Jehee Lee, Jinxiang Chai, Paul Reitsma, Jessica Hodgins, and Nancy Pollard, "Interactive control of avatars animated with human motion data," ACM Transactions on Graphics(Siggraph 2002), Vol. 21, No. 3, pp. 491-500, 2002.

- [2] Lucas Kovar, Michael Gleicher, and Frederic Pighin, "Motion graphs," ACM Transactions on Graphics(Siggraph 2002), Vol. 21, No. 3, pp. 559-568, 2002.
- [3] Matthew Brand and Aaron Hertzmann, "Style machines," Siggraph 2000, pp. 183-192, 2000.
- [4] Yan Li, Tianshu Wang, and Heung-Yeung Shum, "Motion textures: A two-level statistical model for character motion synthesis," ACM Transactions on Graphics(Siggraph 2002), Vol. 21, No. 3, pp. 465-472, 2002.
- [5] Hyun Joon Shin and Jehee Lee, "Motion analysis and synthesis in low-dimensional spaces," Computer Animation and Virtual Worlds, Vol. 17, issue 3-4, pp. 219-227, John Wiley & Sons, July, 2006.
- [6] Kang Hoon Lee and Jehee Lee, "Precomputing avatar behavior from human motion data," Graphical Models, Vol. 68, issue 2, pp. 158-174, Elsevier, March 2006.
- [7] Keith Grochow, Steven Martin, Aaron Hertzmann, and Zoran Popovic, "Style-based inverse kinematics," ACM Transactions on Graphics(Siggraph 2004), Vol. 23, Num. 3, pp. 522-531, 2004.
- [8] Tomohiko Mukai and Shigeru Kuriyama, "Geostatistical motion interpolation," ACM Transactions on Graphics(Siggraph 2005), Vol. 24, Num. 3, pp. 1062-1070, 2005.



이 제 희

1995.2 한국과학기술원 전산학 석사
2000.2 한국과학기술원 전산학 박사
2003.3~현재 서울대학교 컴퓨터공학부 조교수
관심분야: 컴퓨터 그래픽스, 애니메이션, 지능형 가상 캐릭터, 물리 기반 시뮬레이션
E-mail : jehee@cse.snu.ac.kr