

실시간의 참여적 엔터테인먼트를 위한 아바타 모션의 비주얼라이제이션

남양희*

세종대학교 컴퓨터공학과

Daniel Thalmann

스위스 연방 공대(로잔), 스위스

A Study on the Visualization of Avatar Motion for the Realtime Participatory Entertainment

Yanghee Nam

Dept. of Computer Engineering, Sejong Univ.

Daniel Thalmann

Swiss Federal Institute of Technology(EPFL), Switzerland

요 약

가상현실 응용에서 가상 캐릭터의 행위를 효과적으로 조명하는 일은 영화에서의 카메라워크를 실현하는 것과 같다. 그러나, 참여적 엔터테인먼트에서는 아바타의 행위가 참여자의 실시간 행위에 의해 결정되므로, 카메라의 위치나 각도 등을 미리 설정할 수 없다. 본 논문에서는 이러한 즉흥 애니메이션의 상황에서 가상 캐릭터, 즉, 아바타의 행동을 조명하기 위한 실시간 상황 인식 기반 비주얼라이제이션 기법을 제안한다.

1. 서론

본 연구에 앞서, 스위스의 EPFL에서는 사용자의 즉흥적 참여행위를 실시간 모션 캡처를 통해 가상세계에 반영하는 참여적 실시간 엔터테인먼트 시스템을 개발하였다[그림1 참조][1]. 즉, 자발적 참여자가 가상 무대에서 행위를 하고 이를 실시간 모션 캡처 시스템에 의해 가상세계 내의 아바타 모션으로 복제, 재생성해내는 것이다. 이와 같은 참여적 애니메이션 시스템은 가상 노래방, 가상 연극무대, 가상 사이코

드라마 등을 비롯한 다수의 응용에 적용될 수 있는 것으로, 미래지향적 엔터테인먼트 시스템의 한 형태를 이룬다.

그러나, 이와 같은 실시간 참여에 의한 엔터테인먼트 시스템은, 기대되는 행동이나 장면들이 미리 결정되어 있는 게임 및 기타 엔터테인먼트 시스템과는 달리 가상 캐릭터, 즉 아바타의 행동이나 위치 등을 정확히 예측하기 힘들다. 따라서, 현재의 실시간 참여 시스템들은 가상공간을 비추는 카메라의 위치 등을 고정시켜 두거나, 수동적 오퍼레이터를 두어 가상

카메라의 위치 등을 제어하도록 하고 있다.



[그림1] 실시간 참여적 엔터테인먼트
(스위스Orbit 전시장 광경)

그러나, 고정된 카메라의 경우 참여자가 가상적 공간을 벗어나거나 주요 행위가 눈짓, 얼굴 표정 등의 작은 부분에서 일어나는 경우, 이들 행위를 전달하기가 힘들며, 수동으로 카메라를 조작하는 경우는 오퍼레이터가 참여자의 주요 행위를 관찰하면서 가상공간의 카메라 이동도 제어해야 하므로 종종 올바른 타이밍을 놓치거나 피로해질 수가 있고, 카메라의 효과적 설정 방법에 대한 전문가가 아니므로 드라마틱하고 효과적인 카메라워크를 실현하기 어렵다.

본 연구에서는 장면설정이 미리 이루어지지 않은 참여적 실시간 엔터테인먼트의 경우 주어지는 참여자의 행위 입력 분석 및 상호작용의 형태 분석에 의한 실시간 상황 인식을 기반으로 장면에 대한 지식을 추출하고, 이를 바탕으로 한 카메라의 자동 제어 기법을 제안한다.

2. 관련연구

3차원 가상현실의 초기작업들은 객체나 가상세계가 특정인의 시점에서 그려졌던 반면, 최근에는 지능적인 카메라 제어 방법에 관한 시도들이 행해지고 있다. Blinn은 카메라의 설정에 관해 중요한 개념을 제시하였는데, 카메라가 어디 놓일 것인가를 설정하기에 앞서 '무엇이 그 프레임 안에 보여져야 하는가'를 파악하고 이를 기준으로 카메라 설정이 이루어져야 한다는 것이었다[2]. Blinn 이후 대부분의 카메라 설정 관련 연구들은 그러한 맥락에서 이루어졌으며, Karp의 ESPLANADE 시스템은 촬영 언어(film idiom)를 통해 카메라 설정을 지식베이스화하였고[3], CINEMA 시스템을 만든 Drucker는 일반적인 카메라

제어에 대한 로우레벨 라이브러리를 구축한 후 애니메이터가 장면의 제약조건들로부터 상위 개념의 카메라 모듈을 생성할 수 있도록 하였다[4]. Christianson 역시 비슷한 연구를 하였으나, 샷(shot) 간의 전이를 위한 분야 의존적인(domain-specific) 전략을 선연적 방법으로 기술하였다[5].

이러한 방법들은 실시간 엔터테인먼트가 아닌 사전 제작되는 애니메이션 시스템을 대상으로 한 것이므로, 애니메이터가 제작될 장면을 미리 알거나 본 후, 카메라 제어 전략을 정의하는 방법들이다.

실시간 카메라 제어 방법을 고려했던 것은 Bares의 연구가 거의 유일하다고 볼 수 있는데, 사용자 선호도(preference) 입력에 기초한 Bares의 샷 설정방법들은 자동차 운행과 같이 고행의 단일몸체형(single-body) 물체가 주요 객체가 되는 경우에 적용 가능한 것으로, 본 연구에서 고려하는 바와 같이 모션의 중요 부위가 동적으로 변화하는 관절형의 아바타 모션 비주얼라이제이션에는 부적합하다[6].

3. 실시간 상황인식 기반 아바타 비주얼라이제이션

3.1 활성 아바타 및 활성 영역 탐지(detection)

카메라 설정은 현재 장면에서 어느 부분이 중요하게 보여져야 하는가에 달렸는데, 실시간 퍼포먼스인 경우는 참여자의 행위로부터 이를 탐지해내야 한다. 참여자가 여럿일 때는 현재 장면의 중심이 되는 아바타(들)가 누구인가를 탐지해야 하며, 또한 그(들)의 행위의 중심이 어디에 있는가를 탐지하여야 한다. 그렇게 탐지된 중심 아바타(들) 및 중심 영역이 카메라에 의해 가시화되어야 하는 영역이 된다.

중심 인물 및 영역의 탐지를 위해 사용된 방법은 시나리오가 주어지지 않은 상황이므로 철저하게 그 모션에 대한 실시간 분석에 근거하는 것이다.

참여자의 퍼포먼스에 앞서 사전에 주어져야 하는 데이터는 다음과 같다.

- ◆ 참여자의 역할 비율(role factor) - 참여자들의 행위를 대신하는 아바타들 및 가상 캐릭터들은 각

각 역할 비율을 가지며, 역할 비율 별로 행위의 중요성(significance)을 결정하는 임계치가 달리 설정된다. 이를 이용하면, 어떤 캐릭터는 중심인물로서 참여할 수가 있고, 또 다른 캐릭터는 무대위의 댄서(back dancer)와 같이 부차적 역할을 맡을 수 있다. 중심인물의 경우 미세한 동작이라도 의미를 가질 수 있으므로 가시화 대상이 되지만, 엑스트라형 캐릭터의 경우는 크고 예외적인 동작을 하지 않는 한 카메라의 가시화 영역에서 제외될 수 있다.

- ◆ 아바타 몸체의 부위별 모션 임계치 - 인간의 의미있는 모션은 부위별로 그 정도가 다를 수 있다. 예컨대, 얼굴 표정과 같은 것은 미세한 변화로도 의미를 가지게 되는 반면, 상대적으로 다리 자세의 변화 등은 큰 의미가 없더라도 정량적으로 큰 움직임을 발생시키게 된다.

위와 같은 두가지 사전 데이터와 실시간에 관찰되는 아바타들의 모션 분석을 통한 가시화의 중요 대상 파악은 다음과 같이 이루어진다.

- ◆ 활성화 아바타 탐지 - 역할 비율에 따른 위 정도의 임계치와 아바타들의 현재 모션 데이터의 정도를 비교하여 활성화 캐릭터들을 탐지한다.
- ◆ 활성화 영역 탐지 - 몸체의 부위별 임계치에 대해 활성화 아바타들의 모션이 어느 부위에서 일어나는지 파악한다. 이에 따른 비주얼라이제이션 대상 영역 및 상황은 다음과 같이 분류된다.

활성 아바타 + 배경 : 활성화 아바타(들)이 가상 공간 혹은 무대를 가로지르는 등 크게 이동하는 경우이다.(롱 테이크 샷 영역)

활성 아바타의 몸 전체 (풀 샷 영역)

활성 아바타의 얼굴 (클로즈업 영역)

활성 아바타의 상반신 (상반신 영역)

활성 아바타의 무릎 위 상반신 (무릎 컷)

활성 아바타의 상위 몸체 + 특정 객체 또는 배경물체 : 활성화 아바타가 어떤 물체를 대상으로 상호작용하고 있는 경우, 카메라는 아바타의 행위 부분뿐 아니라, 대상이 되는 물체 역시 가시화 해주어야 한다.

이와 같은 분류는 영화 제작자들의 촬영 기법에 근거한 영역 분류에 따르면서 실시간에 탐지 가능한 상황을 통해 확장한 개념으로서, 시나리오에 대한 사전 지식이 없을 때에는 모션이 일어나는 영역이 주요 행위의 힌트가 되며 및 행위의 상호작용성 여부가 중요한 상황 자료가 됨을 제시한 것이다.

3.2 샷 모델링

샷(shot)은 일정 시간에 걸쳐 카메라 파라미터가 일정한 특징을 가지고 유지되는 걸 의미한다. 본 연구에서는 활성화 아바타 및 활성화 영역에 대한 실시간 정보를 기반으로 샷이 동적으로 결정되는데, 각 샷은 다음과 같이 앞 절에서 기술한 활성화 영역과 카메라 파라미터로 기술된다.

- 샷이 포함해야 하는 영역 및 객체(shot body)
- 카메라 설정 유형 : 카메라의 위치, 방향 뿐 아니라 샷 전이(transition) 유형, 트래킹이나 확대(zoom-in) 등의 카메라 모션도 포함한다.

4. 실시간 행위 탐지(action spotting) 기법

3장에서 기술한 활성화 아바타 및 영역 추출은 실시간으로 주어지는 참여자의 모션 데이터 분석에 의해 구해진다. 본 장에서는 모션 데이터 분석 방법에 대해 구체적으로 기술한다.

실시간으로 여러 아바타의 모션을 분석하기 위해 매 순간 각 아바타들의 모든 관찰각 및 위치의 변화를 추적한다는 것은 불가능하다. 따라서, 본 연구에서는 3장에서 기술한 주요 영역 별로 전략적 추적지점(strategic tracking point)을 설정하고, 그 지점에 의미있는(임계치 이상의) 모션이 발생했을 때 해당 영역을 모션을 좀 더 자세히 분석하는 다음과 같은 단계적 분석 방법을 적용하였다.

단계0 (기본단계) : 기본적으로 항상 각 아바타의 위치 및 각 영역별 중요지점(critical point; 전략적 추적 지점)의 값을 추적한다. 각 영역별 추적 지점은 그

영역 내에서 모션이 일어났을 때는 항상 그 지점에도 역시 모션이 발생하게 되는 영역의 대표 지점이라고도 하겠다. 예컨대, 상반신 영역에서는 양 손의 중심점 위치가 전략적 지점이 되는데, 이는 모션을 포함하는 영역의 범위가 상반신일 때 손이 제 위치에 고정되어 있을 수는 없기 때문이다.

단계1 : 단계0을 통해 임계치 이상의 모션이 존재하는 아바타 및 그 영역이 추출되면, 발생한 모션의 내용이 매순간 어떤 의미가 있는 자세를 이루는가를 분석한다. 이 단계의 주요 분석 대상은 포즈(posture)로서, 상반신 영역이 채택된 경우는 손 모양 인식 등이 수반되며, 얼굴 영역(Close-Up)이 샷의 대상 영역인 경우 표정인식 등이 수행된다.

단계2 : 이 단계에서는 한 순간의 자세를 파악하는 단계1과는 달리, 일정 시간에 걸쳐 계속적으로 만들어지는 시공간 모션 데이터를 분석하게 된다. 주로 단계1에 의해 분석된 각 순간의 자세들이 시간적으로 이어져서 어떤 의미있는 제스처를 나타내는지 하는 것은 이 마지막 단계에서의 연속동작 분석기에 의해 추출된다.

단계1과 단계2의 경우, 점진적으로 영역을 좁혀가며 깊게 분석한다는 다단계 분석 프레임웍은 유지하면서, 인식할 대상의 특징에 맞는 인식기가 플러그인 형태로 쓰일 수 있다.

5. 실험 : 디지털 아바타를 통한 즉흥 퍼포먼스

본 연구의 프레임웍은 실시간 모션 캡처를 바탕으로 한 즉흥 퍼포먼스에서 카메라의 자동 제어에 적용되었고, 본 연구의 실시간 활성 영역 추출 및 분석과, 이에 기반한 샷 결정 방법은 카메라가 고정되었을 때와는 달리 매우 드라마틱한 장면 효과를 연출하였다. 아래 그림은 결과 애니메이션의 여러가지 샷을 보여주는 것으로, [그림2]의 경우는 중요 아바타 중 하나가 무대에서 위치 이동을 할 때, 연속동작 분석을 통해 활성 아바타를 파악하고, 그 아바타의 동작이 걸어서 이동하는 것임을 인식하여 카메라가 트래킹하게 된 샷의 예를 보여준다. [그림3]의 경우는

두 사람이 모두 손짓을 써가며 대화를 하는 상황으로 두사람의 상반신 모두가 활성 영역으로 포착되어 비주얼라이제이션 되었다. [그림4]의 경우는 두 아바타 중 한명의 아바타만 말을 할 때 그 아바타의 얼굴을 클로즈업 한 예이다.



[그림2]

[그림3]

[그림4]

6. 결론

본 연구는 행위의 내용이 미리 결정되지 않은 참여적 엔터테인먼트의 경우, 참여자의 동작에 대한 실시간 분석을 기초로 한 카메라의 자동 제어 방법을 제시하였고, 즉흥 퍼포먼스 시스템의 아바타 비주얼라이제이션 컴포넌트를 구현하여 카메라 효과를 실험하였다.

[참고문헌]

- [1] S.Balcisoy et al., "An Interactive Interface for Directing Virtual Humans", In Proc. ISCI98, IOS Press, 1998.
- [2] J.Blinn, "Where Am I? What Am I Looking At", IEEE Computer Graphics and Applications, pp.76-81,1988.
- [3] P.Karp, S.Feiner, "Issues in the automated generation of animated presentations," In Proceedings of Graphics Interface'90, pp.39-48,1990.
- [4] S.M.Drucker et al. "CINEMA:A system for procedural camera movements," Symposium on Interactive 3D Graphics, vol.25, pp.67-70, 1992.
- [5] D.B.Christianson et al. "Declarative Camera Control for Automatic Cinematography," Proceedings of the AAAI-96, August 1996.
- [6] W.H.Bares, J.C.Lester, "Cinematic User Models for Automated Realtime Camera Control in Dynamic 3D Environments," Proceedings of the Sixth International Conference, UM97, Vienna 1997.