

텍스트로부터의 자동 디지털 영상제작을 위한 카메라 자동 제어*

장세민⁰ 박종철
한국과학기술원 전산학 전공 및 첨단정보기술연구소
{smjang⁰, park}@nlp.kaist.ac.kr

Automatic Camera Control for Automated Digital Cinematography from Text

Se-Min Jang⁰ Jong C. Park
Computer Science Division and AITrc, KAIST

요 약

영화를 제작하는 과정에 필수적으로 사용되고 있는 대본(臺本)에는 필요한 부분마다 영상기법이 명시되어 있어서 실제 장면을 구현하는 과정에 원작자가 의도하는 상황을 비교적 정확하게 재현하는 것이 가능하다. 이에 비하여 교통사고 사건보고서나 동화 등을 기반으로 디지털 영상을 자동으로 제작하려는 경우 이러한 영상기법이 명시되어 있지 않다. 그러므로 자연언어로 기술된 자료로부터 디지털 영상을 자동으로 제작하기 위해서는 작가의 의도를 파악하여 적절한 영상기법을 추출하는 방안이 있어야 한다. 본 논문의 선행 연구에서는 동화를 대상으로 하는 애니메이션 자동 생성을 위해서 시간 관리, 참조 해결, 위치 설정, 세부 명령 결정 및 다수 캐릭터 제어 등의 요소 기술이 필요하다는 것을 보이고 특히 시간 관리 중에서 적절한 장면전환이 필요한 경우를 자동으로 파악하는 방안을 제시하였다. 본 논문에서는 결합범주론법을 사용하여 동화 문장에 나타나는 작가의 의도를 분석하고, 이에 부합하는 다양한 카메라 운용기법을 자동으로 파악하여 적용한 디지털 영상 제작 방안을 제시하고 구현한 시스템을 보인다.

1. 서론

텍스트 애니메이션은 자연언어를 통한 애니메이션 자동 생성에 대한 분야로 현재 군사용 시뮬레이션, 가상 캐릭터 제어, 자동차 사고 시뮬레이션 등의 다양한 연구가 진행되고 있다. 이는 자연언어가 표현에 대한 별도의 훈련 과정 없이도 사용자가 생각하는 바를 자연스럽게 간결하게 표현할 수 있게 한다는 장점이 있으므로 자연언어 인터페이스의 개발이 활발하게 이루어지고 있기 때문이다. 본 연구에서는 자연언어문장으로 이루어진 동화의 내용을 카메라 기법이 적용된 애니메이션으로 생성하는 시스템에 대하여 다룬다.

기존의 텍스트 형태의 동화에 시각 및 청각적 효과를 결합한 애니메이션을 멀티동화라고 한다. [1]에서는 텍스트 형태의 동화를 입력 받아 멀티동화를 자동으로 생성할 때 해결해야 하는 문제점으로 시간 관리, 참조 해결, 위치 결정, 세부 명령 결정 및 다수 캐릭터 제어 등을 논의하고 있으며, [4]에서는 캐릭터 및 동작과 사건에 대한 참조현상의 해결방안에 대하여 논의하고 있다. 또한 [2]에서는 동화 상에 나타나는 사건들의 시간순서를 파악하여 사건 간의 시간관계를 작가의 의도에 부합하는 장면전환기법으로

표현하는 방법에 대하여 논의하고 있다.

그런데 애니메이션에서 극적인 효과를 담당하는 부분은 캐릭터나 객체 등의 그림보다는 카메라 기법이나 음향, 빛과 색채 등이다. 이런 부분들이 애니메이션을 단순히 움직이는 그림이 아닌 흥미진진한 이야기를 전달하는 매체로 만드는 것이다. 본 연구에서는 자연언어 동화로부터 문장 성분을 분석하여 문장에서 드러나는 작가의 의도에 따라 상황에 적절한 카메라 기법을 적용함으로써 효과적인 애니메이션을 생성하는 방안에 대하여 논의한다.

2. 관련연구

본 절에서는 자연언어문장을 애니메이션으로 생성하는 텍스트 애니메이션에 관한 연구와 디지털 영상기법에 대한 연구에 대하여 살펴본다.

[7]은 자동차 사고에 대하여 서술된 경위서로부터 애니메이션을 생성하는 시스템에 대한 연구로, 정적 객체(static object), 동적 객체(dynamic object), 충돌 객체(collision object)를 사용하여 사고에 관련된 물체와 사고 상황 등을 기술하고 이로부터 애니메이션을 생성하는 방법을 사용하고 있다. 그러나 이 연구에서는 사고배경 및

* 본 연구는 첨단정보기술 연구센터를 통하여 한국과학재단의 지원을 받았음

상황이 몇 가지 경우로 제한되어 있고, 사실적인 사고상황 전달만을 목적으로 하므로 극적 표현을 위한 영상기법에 대한 논의는 부재한다.

[5]에서는 3차원 가상환경에서 실시간으로 카메라를 자동 제어하는 방법에 대하여 논의하였다. 이 시스템은 사용자의 직접적인 상황 설정에 따라서 각 상황에 정의된 카메라 기법을 적용하여 애니메이션을 생성하는 방법을 사용하였으며, 상황 설정이 제한되어 있고 카메라 제어가 유연하게 적용되지 못한다는 문제가 있다.

[8]에서는 3차원 가상환경에서 사용자와 사용자에 의해 조종되는 캐릭터 간의 상호작용효과를 높이기 위한 자동 영상기법 시스템을 제안하였다. 그러나 이 시스템은 영상 기법의 적용공간이 제한되어 있으며, 상황의 변화에 따른 영상기법의 변화가 유연하지 못하다.

3. 카메라 자동 제어를 위한 문장성분 분석 및 처리

카메라와 피사체간의 거리, 상하각도, 방위, 그리고 카메라의 움직임이 조합되어 전체적인 카메라 기법을 이룬다. 카메라 기법은 잘 활용될 경우 자연언어 문장에서 캐릭터의 행동과 상태를 표현하는 문장성분인 동사와 부사, 형용사에 담긴 의미를 잘 전달할 수가 있다.

본 절에서는 동사, 부사, 형용사에 담긴 의미를 분석하여 적절한 카메라 표현방법을 제안하고, 결합범주문법을 이용하여 처리하는 방안을 보인다.

3.1. 문장성분의 분석과 카메라 제어기법

동사는 캐릭터의 행동을 나타내는 문장성분으로서, 애니메이션으로 표현해야 할 행동의 주요 사항을 포함하고 있다. 카메라 제어기법을 위해서 동사로부터 추출해야 하는 정보는 행동유형과 행동반경, 방향성이다. (<표 1> 참조)

표 1. 각 어휘별 행동 속성

| 어휘 | 읽다 | 가다 | 오다 |
|------|-----|----|----|
| 속성 | | | |
| 행동유형 | 읽기 | 걸기 | 걸기 |
| 행동반경 | 상반신 | 전신 | 전신 |
| 방향성 | 없음 | 후면 | 전면 |

각 행동은 동적인 속성이 다르며, 같은 행동도 표현방식에 따라서 전달하고자 하는 바가 다르다. 따라서 각 행동 유형별로 적절한 카메라 제어기법이 존재한다고 볼 수 있다. 각각의 행동을 식별하기 좋게 하며, 전달해야 하는 정보를 적절하게 제한함으로써 표현하고자 하는 바를 명료하게 전달하기 위해서는 행동의 반경에 따라서 화면에 포함되는 캐릭터의 범위가 달라져야 한다. 또한 행동의 방향성에 의해 상하좌우의 2차원 공간뿐 아니라 화면의 깊이

를 활용하여 3차원 공간으로 표현함으로써 객관적인 사실 전달에 그치지 않고 주관적인 사건 체험의 효과를 가질 수 있게 한다.

부사는 캐릭터의 행동이나 상태의 속성을 구체적으로 나타내는 문장성분으로서, 강조하여 나타내고자 하는 의도를 충실히 표현해야 한다. 예를 들어, ‘정신없이 달렸다’에서 달리는 모습만으로 캐릭터가 정신이 없다는 것을 표현하기는 어렵다. 그러나 캐릭터가 열심히 달리는 모습이 화면 가득히 보여진 후, 캐릭터의 등뒤에서 따라가며 캐릭터와 같은 시선으로 주위 사물들이 흔들리며 빠른 속도로 시야 밖으로 사라져가는 모습을 보여줌으로써 달리는 것 이외에는 관심이 없음을 표현할 수 있다.

형용사는 캐릭터의 외형 및 심리 상태를 표현하는 것으로 많이 쓰이는 문장성분이다. 이와 같은 표현은 독자의 관심을 캐릭터 자체에 집중시키고자 사용되기 때문에, 이를 애니메이션에서도 강조하여 표현해야 한다. 예를 들어, ‘허름한 옷을 입은 사나이’에서 단지 허름한 옷을 입고 있는 모습을 보이기보다는 캐릭터를 중심으로 카메라가 한바퀴 돌며 보여주는 것이 그 옷의 상태를 보여주고자 하는 작가의 의도가 잘 전달된다.

이와 같이 적절한 카메라 기법을 적용함으로써 문장성분 고유의 속성과 이에 담긴 작가의 의도를 제대로 반영한 애니메이션을 생성할 수 있다.

3.2. 결합범주문법을 이용한 처리

결합범주문법은 범주문법에 결합자(combinator)가 추가된 형태로 소수의 축약 규칙에 의하여 구문분석이 이루어지는 어휘 문법으로, 각 어휘마다 문법, 의미, 담화 정보를 담은 범주가 할당되어 특별한 약정 없이 축약규칙만을 통해 문법정보 외에 의미정보나 담화정보까지 한번의 과정으로 유도할 수 있다는 장점을 지닌다[3, 6].

본 연구에서 영상기법을 추출하기위해 각 어휘에 할당하는 기본 범주는 np와 s로 다음과 같이 구성되어 있다.

np : [NP_ID, NP_FEATURE]
s : [NP, ACTION, SCALE, DIRECTION, FEATURE]

특정 객체를 나타내는 np는 객체 식별자인 NP_ID와 객체의 속성을 담고있는 NP_FEATURE로 이루어져 있다. 그리고 문장을 표현하는 s는 행동의 주체에 대한 정보가 담긴 NP와, 행동과 상태의 유형을 나타내는 ACTION, 행동의 반경을 표현하는 SCALE, 행동의 방향성을 나타내는 DIRECTION, 행동이나 상태의 구체적인 속성을 나타내는 FEATURE로 이루어져 있다.

<그림 1>은 결합범주문법을 이용한 구문분석의 결과를 보여준다.

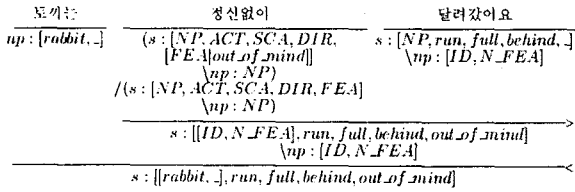


그림 1. 구문분석 결과

4. 구현 및 결과

전체적인 시스템의 구조는 <그림 2>와 같다. 먼저 자연 언어문장으로 이루어진 동화를 입력받아 CCG 파서를 통하여 구문분석을 한다. 구문분석의 결과를 스크립트 생성기에 입력하면 행동의 유형, 변경, 방향성 및 속성 정보를 조합하여 카메라 기법이 생성되고, 중간 결과 스크립트가 만들어진다. 이 중간 결과 스크립트는 애니메이션 스크립트 생성기를 통해 애니메이션 스크립트로 변환된다. 최종적으로 애니메이션 스크립트가 애니메이션 생성기에 입력되면 카메라 기법이 적용된 애니메이션이 생성된다.

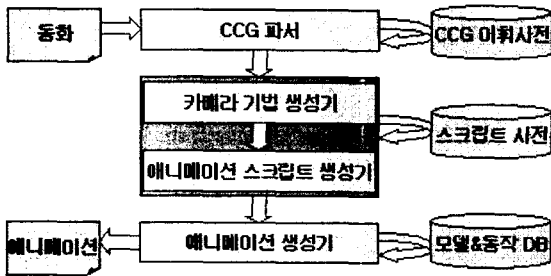


그림 2. 시스템 구조

<그림 3>은 <그림 1>의 예문에 대한 애니메이션 스크립트로, OBJLIST에는 화면에 등장할 캐릭터에 대한 정보가 있으며, ACTIONLIST에는 캐릭터의 동작 정보(객체 식별자, 동작, 좌표, 시간)와 카메라 제어 정보(움직임, 각도, 방위, 위치, 범위, 시간)가 담겨 있다.

<그림 4>는 C++와 MilkShape3D로 만들어진 애니메이션 생성기의 결과 애니메이션 장면이다.

```

<OBJLIST>
id(0) rabbit nil
</OBJLIST>
<ACTIONLIST>
id(0) run -300 0 50 10 105
cam static low quarter in knee id(0) 11 50
cam handheld eye back center shoulder id(0) 51 104
</ACTIONLIST>
    
```

그림 3. 애니메이션 스크립트

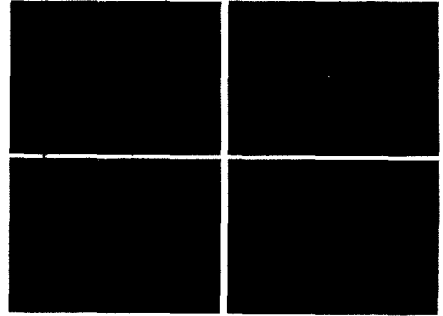


그림 4. 애니메이션 생성 결과

5. 결론 및 향후 과제

본 논문에서는 자연언어문장으로부터 애니메이션을 생성하는 텍스트 애니메이션에서 작가의 의도를 파악하여 애니메이션으로 재현하기 위한 방법에 대하여 논의하였다. 특히, 행동 및 상태를 표현하는 문장성분인 동사, 부사, 형용사에 담긴 정보를 활용하여 카메라 기법을 자동으로 제어함으로써 작가의 의도를 반영한 애니메이션 자동생성 방법을 제안하였다.

텍스트 애니메이션은 이야기 글을 애니메이션으로 생성하는 것을 목적으로 한다. 이야기 글은 한 문장이상으로 이루어져 있으며 문장의 흐름에 따라 이야기 전개방식이나 표현방식이 달라진다. 마찬가지로 애니메이션도 그 흐름에 따라 표현방법이 달라져야 한다. 앞으로 문맥에 따라 다르게 적용되는 카메라 기법에 대한 연구를 진행할 예정이다. 또한 카메라 기법 외에 음향, 색채, 조명 등의 영상 기법에 대하여 자동 생성하는 연구가 필요하다.

참고 문헌

- [1] 김현숙, 박종철, " 결합범주문법을 이용한 실시간 한국어 멀티동화 제작", 한글 및 한국어 정보처리학술대회 논문집 제 13권, pp.509~515, 2001.
- [2] 장세민, 박종철, " 디지털 영상기법을 위한 자연언어 멀티동화 시간 관계 분석 및 재현", 한글 및 한국어 정보처리학술대회 논문집 제 14권, pp.309~316, 2002.
- [3] 조형준, 박종철, " 한국어 병렬문의 통사, 의미, 문맥 분석을 위한 결합범주문법", 정보과학회논문지, pp. 448-462, 2000.
- [4] Kyung Wha Hong and Jong C. Park. Anaphora Resolution in Text Animation, Proceedings of The IASTED International Conference on Artificial Intelligence and Applications(AIA 2004), pp. 347-352, 2004.
- [5] Li-wei He, Michael F. Cohen and David H. Salesin, The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing, Proceedings of SIGGRAPH, pp.217~224, 1996.
- [6] Mark Steedman. The Syntactic Process. The MIT Press, 2000.
- [7] Sylvain Dupuy, Arjan Egges, Vincent Legendre and Pierre Nugues, Generating a 3D Simulation of a Car Accident from a Written Description in Natural Language : the CarSim System. Proceedings of ACL Workshop on Temporal and Spatial Information Processing, pp. 1-8, 2001.
- [8] William Michael Tomlinson, Jr., Interactivity and Emotion through Cinematography, MS. Thesis at the MIT., 1999.