

자연 환경 비디오에서 영상 분석 기법을 이용한 혼합 현실

김형진^o 박종승* 이만재

한국정보통신대학교 디지털 미디어 연구소, 인천대학교 컴퓨터공학과*
myditto@icu.ac.kr

Mixed Reality using Image Analysis Technique in Natural Scene Video

Hyoungjin Kim^o Jong-seung Park* Manjai Lee

Digital Media Lab, Information and Communications University, University of Incheon*

요약

CG 기술의 발전으로 실사 이미지 위에 3d 애니메이션을 적용하는 실감나는 영화를 많이 볼 수 있다. 그러나 아직까지도 실사 이미지와 3d 애니메이션 합성은 그래픽 아티스트의 수작업이 많이 필요하다. 본 논문에서는 자연 환경을 포함하는 실사 비디오에서 영상 분석 기법을 적용한 혼합 현실 시스템을 제안한다. 비디오 프레임에서 하늘, 바다 등 일정한 칼라 패턴을 가지는 영역을 분류하고 이 영역에 대해서 그래픽 객체를 삽입한다. 시스템의 각 단계에서 사용자 인터페이스를 제공하여 필요한 입력값을 제공한다. 본 연구에서의 기법은 간단한 그래픽 입력 과정 이외에는 대부분의 과정이 자동으로 처리되므로 혼합 현실이 방송, 영화 특수 효과 등의 작업에 편리하게 이용될 수 있다.

1. 서론

최근 성공한 대부분 할리우드 영화는 제작 공정의 50% 이상을 첨단 CG 기술을 활용하여 제작하고 있으며, 이를 통해 제작 경비도 획기적으로 절감하고 있다. 이러한 추세는 더욱 더 확대될 전망이며, 영화 방송 등의 매체가 디지털화됨에 따라 수작업을 많이 요구하던 아날로그 방식의 크로마킹 기법을 통한 영상 합성 방식이 자동화된 디지털 영상 합성 기술로 대체될 전망이다. 따라서 카메라 시점 일치 관련 기술의 고도화 및 고속화가 추가적으로 요구되며, 아직까지 수작업에 의존하는 로토스코핑의 자동화도 함께 요구된다. 이와 같은 로토스코핑의 자동화를 위해서는 영상 내 객체의 삽입, 삭제 등이 자유로워야 하며, 아울러 자연스러운 애니메이션 재현, 합성된 영상의 이질감을 없애고 현실감을 높일 수 있는 실사 조영 구현 기술과 색감 일치 기술은 고품질의 영상 콘텐츠의 요구가 높아질수록 그 필요성은 더욱 더 높아진다.

본 연구에서는 이러한 여러사항을 고려하여 실사 이미지 위에 자연스러운 애니메이션을 구현하는 데 목적이 있다.

2. 관련 연구

비디오 시퀀스에서 특정 패턴을 자동으로 인식하고 그 장면에 맞는 3차원 오브젝트를 자연스런 애니메이션으로 연출하는데 가장 큰 어려움은 우리가 원하는 애니메이션의 특정 패턴을 미리 예측하기 쉽지 않다는 것이다. 패턴의 예측이 가능하다고 하더라도 자연스러운 애니메이션을 구현하기가 쉽지 않다. 구현이 어려운 이유는 비디오

장면의 내용을 의미적으로 해석할 수 없다는 데 기인한다. 동작의 패턴을 인식하기 위해서는 상당히 큰 컴퓨팅 자원을 소모해야 한다. 이 문제를 해결하는데 image differencing[1] 같은 간단한 방법은 적용되지 않는다. 왜냐하면 우리가 원하는 패턴이 아주 작은 픽셀의 사이즈를 갖는다면 콘트라스트가 아주 작으면 찾기가 힘들어지기 때문이다. 하나의 대안적인 방법은 영상으로부터 클러터를 제거함으로써 신호 대 잡음비를 향상시키려고 시도하는 것이다. 이 방식은 상관된 클러터(correlated clutter)의 제거를 위하여 2-D 적응 래티스 알고리즘(2-D adaptive lattice algorithm)을 사용하여 Ffrench et al.[2]에 의해서 기술되었다. 주로 시뮬레이션된 영상들과의 비교들은 더 우수한 결과들이 얻어지는 반면에 계산 시간과 알고리즘 복잡도에서의 증가된 비용이 상당하다는 것을 보여주었다.

본 논문에서 우리는 객체 크기가 8 픽셀 정도로 작을 때에도 객체 검출을 위하여 강력한 방법을 제공하기 위하여 웨이블릿 필터들의 유용한 특성들[3]을 이용한다.

3. 그래픽 합성 시스템 구성

그래픽 합성 시스템의 전체적인 구조는 그림 1과 같다.

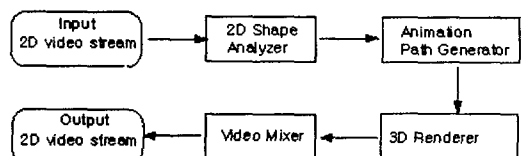


그림 1. 합성 시스템의 구조

그림 1에서와 같이 합성 시스템은 크게 4개의 세부 모듈로 구성된다. 비디오 프레임의 모양 분석을 위한 2차원 모양 분석기, 애니메이션 경로 생성기, 3차원 렌더러, 비디오 합성기의 4개로 구성된다. 각각의 단계를 순차적으로 거치게 된다.

사용자는 각 단계에서 간단한 그래픽 인터페이스 메뉴를 통해서 입력 데이터를 제공할 수 있다. 2차원 모양 분석기에서는 2차원 이미지들 자동으로 영역 분할하고 각 영역들을 색 정보를 사용하여 분류한다. 분류는 하늘, 산, 흙 등의 정형화할 수 있는 칼라값을 이용하여 장면을 단순히 몇 개의 영역으로만 분류할 수 있도록 한다. 사용자는 분류 결과를 수동으로 보정할 수 있다.

애니메이션 경로 생성기에서는 사용자가 경로의 제어점들을 마우스로 지정한다. 이들 제어점들을 연결하는 부드러운 곡선이 애니메이션의 경로가 된다. 3차원 렌더러에서는 가용 그래픽 모델들을 선택할 수 있도록 한다. 그래픽 모델들은 수시로 추가 및 삭제될 수 있다.

선택된 그래픽 모델은 자세 지정 메뉴로 초기 자세를 지정할 수 있게 되어있다. 일정 간격으로 그래픽 모델에 대해서 변환 행렬이 생성되어 각 애니메이션 패스에서의 구간에 대해서 보관된 변환 행렬이 적용되어 적절히 변환된 3차원 모델이 만들어진다.

변환된 3차원 모델은 렌더러에 의해서 렌더링된다. 렌더링된 이미지는 메모리의 가상 이미지 버퍼에 저장된다. 이렇게 얻어진 가상 이미지 버퍼의 이미지는 원본 비디오 프레임 위에 더해진다.

합성 시스템은 대부분의 AR 시스템과 같은 일반적인 구조이다. 본 연구에서의 합성 시스템은 자연적인 이미지 요소를 이용한다는 측면에서 다른 AR 시스템과 차별화된다. 본 연구에서는 하늘과 같은 특정한 자연적인 공간을 자동으로 인식하고 이 공간에 주어진 그래픽 모델을 삽입하는 방법을 제시하였다. 본 방법의 구현을 위해서는 하늘 영역의 자동 분류 기술이 필요하고 그래픽 모델의 외곽 테두리와 하늘에 해당하는 이미지 영역과의 자연스러운 접합 기술이 필요하다.

4. 영역의 자동 분류와 그래픽 접합 기술

본 시스템에서 입력은 비디오 스트림이다. 패턴 인식 기법(Wavelet filtering의 특성[2]을 이용한다.)을 사용해서 2차원 모양을 분석한다. 분석 결과로 그 장면에 맞는 3d 오브젝트 구조를 얻되, 사용자는 animation path를 마우스를 가지고 지정할 수 있다. animation path를 설정함으로써 동적인 애니메이션을 만들어낼 수 있다. 실사비디오와 미리 렌더링 된 3d 오브젝트를 합성함으로써 좀 더 실감나는 영상을 만들어 낼 수 있다.

웨이블릿 필터는 일반적으로 신호 성분을 많이 포함하는 큰 계수와 신호 성분이 작은 크기의 계수로 나누어진다. 이러한 웨이블릿 계수의 특성은 적절한 임계값을 이용하여 웨이블릿 계수를 두 영역으로 분리하여 이진 마스크를 생성하고, 생성된 마스크의 정보를 이용하여 영상의 분리, 잡음 제거 등 다양한 분야에 쓰인다.

특정영역에서 검출할 때 한 프레임당 1초 이내에 찾아야 한다. 임의로 다중 검출 영역(영역의 사이즈 6x6)을 설정해서 하늘의 패턴을 찾아서 값을 매긴다. 여기서 빈도수(voting)를 보고 특정 패턴을 찾는다. 빈도수가 높은 영역이 3차원 좌표 계에서 원점이 된다. 더불어 사용자가 원하는 경로에 따라 미리 렌더링된 3d 오브젝트를 믹싱한다. 여기서 사용자가 마우스로 그린 경로 위에서 애니메이션을 구현한다

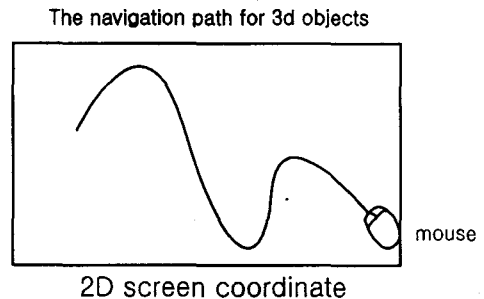


그림 2. 애니메이션 패스 지정을 위한 마우스를 사용한 사용자 인터페이스의 예시

그림 2은 사용자가 애니메이션 패스를 지정하는 예를 보인다. 지정된 애니메이션 패스는 그림 3와 같은 3차원 궤적으로 변환된다.

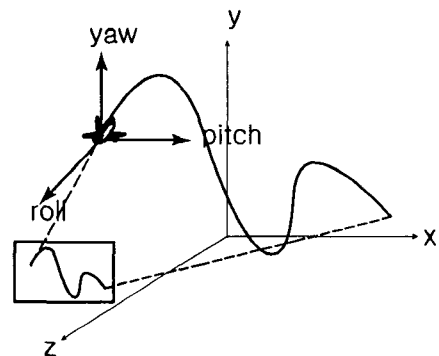


그림 3. 애니메이션 패스 생성을 위한 3차원 궤적

3차원 궤적은 2차원적인 패스로부터 임의로 생성해야 한다. 본 방법에서는 2차원 윈도우에서의 좌표값으로 roll과 pitch값을 결정하고, 좌표값의 스크린 높이값을 이용하여 임의로 yaw를 결정하는 식으로 이루어진다.

애니메이션 패스가 생성되면 해당하는 변환 행렬에 의해서 모델의 3차원 렌더링한다. 렌더링 결과 얻어진 렌더링 이미지는 실사 비디오 프레임과 합성된다. 합성 시 발생하는 부자연스러운 이음쇠를 제거해야 한다. 슈퍼 샘플링을 이용한 앤티 앨리어싱 기법을 사용하여 이음쇠 부분이 거의 분간할 수 없는 수준으로 된다.

비디오 스트림 위에 3d 오브젝트(예를 들면 비행기)가 움직이는 경로를 마우스로 그려주면, 그 경로는 3차원 projection space로 역변환 되어 3d 오브젝트는 그 위에서 움직인다. 여기서 고려할 사항은 실사의 비디오 영상과 3d 오브젝트가 얼마나 잘 블렌딩할 수 있는 지이다. 특히 3d 오브젝트의 외곽선 처리가 관건이다. 본 연구에서는 실시간을 고려해서 하드웨어적으로 anti aliasing이 지원되는 그래픽 카드를 이용하여 처리한다. 마지막으로 중요한 고려사항은 자연스러운 애니메이션을 위해 지정된 경로 위에서 보간을 통해 부드러운 움직임을 연출한다.

5. 실험 결과 및 결론

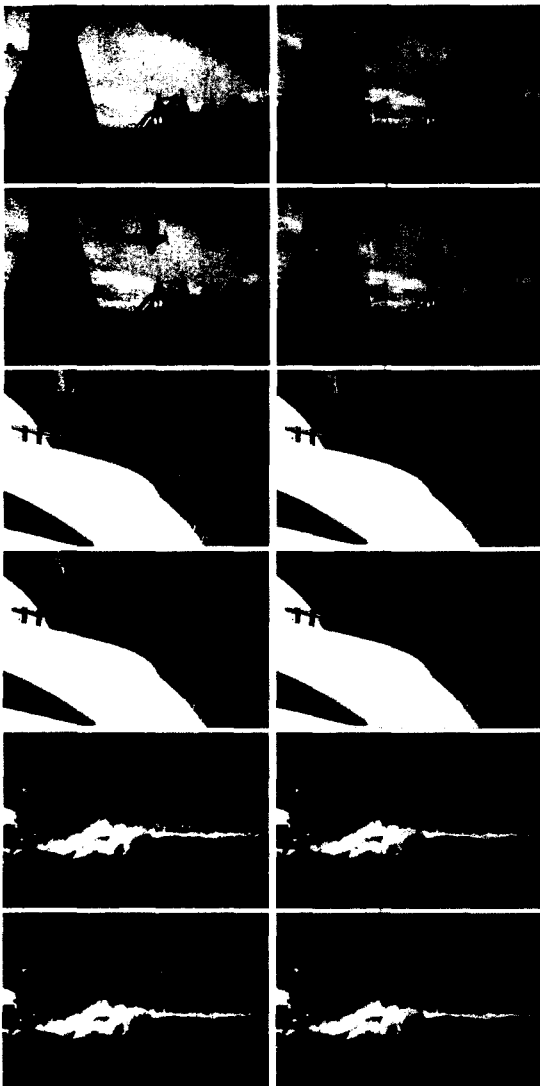


그림 4. 그래픽 오브젝트를 합성한 실험 예

ARToolKit은 비전 기반 AR 응용 시스템을 위한 라이브러리이다. 오픈소스이면서 멀티 플랫폼을 지원한다. University of Washington의 human interface technology lab[5]에서 다운받을 수 있다. ARToolKit을 이용하면 빠른 AR 응용 프로그램의 구현이 가능하다.

자연스러운 AR 시스템의 구현을 위해서는 인위적인 마커보다는 자연적인 환경에서의 특징점들을 추적해야한다. 자연적인 특징점들의 추적인 기술적으로 상당한 수준까지 연구가 되어있다 [6,7].

현재의 합성 시스템에서는 이미지 영역의 분할이 제대로 되지 않는 경우가 많이 발생하여 목표한 애니메이션이 이루어지지 않는 경우가 빈번히 발생한다. 또한 모델의 각도 변화를 자유롭게 지정할 수 없는 단점이 있다. 이의 해결을 위해서 정확한 영역 분류 방법 및 경로의 각 제어점에서의 각도 지정 방법의 개발이 필요하다.

참고 문헌

1. P.L Rosin and T. Ellis. Image difference threshold strategies and shadow detection. In Proceedings of the 6th British Machine Vision Conference, pages 347~356. BMVA Press, 1995.
2. P.A. Ffrench, J. R. Zeidler, and W. H. Ku. Enhanced detectability of small objects in correlated clutter using an improved 2-d adaptive lattice algorithm. IEEE Trans. on IP, 6(3):383~397, 1997.
3. C.K. Chui. An Introduction to Wavelets. Academic Press Inc, 1992.
4. D. Davis, P. Palmer, and Mirmehdi, " Detection and tracking of very small low contrast objects" , in 9th British Machine Vision Conference, September 1998
5. <http://www.hitl.washington.edu/>
6. Ulrich Neumann and Suya You, Natural Feature Tracking for Augmented Reality. IEEE Transactions on Multimedia, 1(1), pp. 53-64, 1999
7. V. Lepetit, L. Vacchetti, and P. Fua, Fully Automated and Stable Registration for Augmented Reality Applications, In International Symposium on Mixed and Augmented Reality, Tokyo, Japan, September 2003.