



비디오 기반 증강 현실

박상철¹⁾, 황분우²⁾, 이성환³⁾

목 차

1. 서 론
2. 증강 현실의 사례 연구
3. 비디오 기반 증강 현실 구현
4. 실험 및 결과
5. 결 론

1. 서 론

우리 일상 생활에서 TV나 영화 등의 미디어를 통해 영상을 볼 때 실제의 물체와 가상의 물체를 구별하기가 점점 어려워지고 있다. 가까운 미래에는 이들 가상의 물체들이 실제 환경에 따라 더욱 사실적으로 나타나게 될 것이며, 이는 영상 매체의 적용 가능성이 다양한 분야로의 확장됨을 예견하는 것이다. 이를 가능하게 하는 것이 증강 현실 기술이다.

증강 현실이란 가상 현실 적용 분야의 하나로 인식되고 있으며 실제 영상에 가상의 다른 영상을 자연스럽게 합성하는 기술을 말한다. 이는 화면에 나타나는 모든 물체와 영상을 생성하여 보여주는 기존의 가상 현실 기술인 VR(Virtual Reality)와는 차별성이 있다. 즉, 증강 현실이란 실제 영상과 가상 현실의 중간 형태의 기술인 것이다. 뿐만 아니라 영상에서 특정 객체를 인식하여 해당 객체에 관한 보다 많은 정보를 사용자가 원하는 시점

에 제공할 수도 있다.

현재 증강 현실 기술이 활발히 적용되고 있는 분야들은 의료 분야로서 기존의 초음파 영상만으로 수술하는 기법에서 벗어나 HMD(Head Mounted Display)를 이용하여 환자의 수술 부위를 3차원으로 보면서 정확히 수술할 수 있는 기술들이 적용되고 있으며 이는 의대생의 실습에도 유용하게 사용되고 있다[1].

또한 영화 분야에서는 빌딩의 간판이나 옥외 게시판 등을 자연스럽게 협찬사의 상표나 간판으로 교체하여 광고 효과를 높여 준다든지, 실제로 구입하기 힘든 소품이나 물건 등을 3차원 가상 객체로 자연스럽게 합성하여 관객들에게 보여 줌으로써 광고 효과의 극대화와 제작비 절감의 효과를 누릴 수 있게 되었다.

TV나 케이블 방송의 스포츠 중계에서도 증강 현실 기술은 다양하게 적용되고 있다. 특히 생방송 중에 카메라 동작 파라미터 정보를 추출할 수 있는 특수 하드웨어를 이용해서 경기장 바닥에 로고를 생성하는 기술이나 Virtual Wall 같은 기술 등이 있으며, 스포츠 중계에서의 이러한 서비스는 이미 국내에서도 시도되고 있고, 유럽이나 미국에서는 가장 보편적인 TV 시청자의 영상 서비스로

1) 고려대학교 컴퓨터학과 박사 수료
 2) (주)버추얼미디어 기술연구소장
 3) 고려대학교 인공시각연구센터 연구소장

인식되고 있다[2].

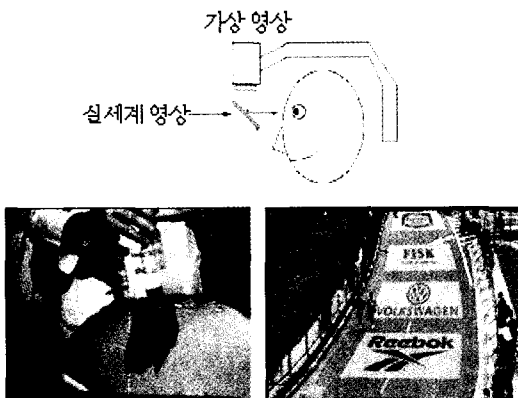
그 외에도 교육용 시뮬레이션 분야나 영상의 특정 객체 인식을 통한 부가 서비스 분야에도 증강 현실의 기술이 다양하게 적용되고 있다.

2. 증강 현실의 사례 연구

현재 증강 현실은 다양한 분야로 발전하고 있으며, 이에 따라 많은 형태의 시스템과 응용 분야들이 있다. 일단 증강 현실 기술이 적용되어 보여지는 하드웨어 장비들 면에서 분류를 해 보면 HMD와 모니터 방식으로 분류된다.

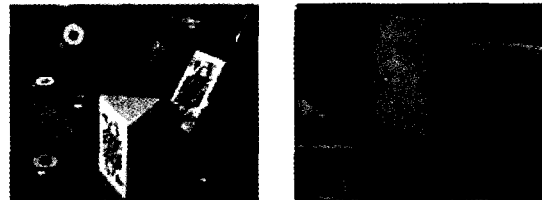
HMD 방식은 사용자의 눈에 아주 가까이 근접해 있는 디스플레이 장치와 헤드셋으로 이루어져 있으며 가상 현실 기술과 유사한 양쪽 눈에 디스플레이를 하는 장치와 한쪽 눈에 디스플레이 장치를 하는 방식이 있다. 이러한 방식은 의료용이나 교육용 등에 적용될 수 있다[1].

앞의 HMD의 방식과는 다른, 모니터를 통해 출력되는 방식이 있다. 즉, 다양한 형태의 디스플레이 화면을 통해 여러 사람이 동시에 볼 수 있는 방식이다. 모니터 방식의 디스플레이는 시청자를 위한 스포츠 중계나 영화 등에 적용될 수 있다[2].



(그림 1) HMD의 구조(상)와 적용 분야(하/왼쪽) 및 모니터 방식의 예(하/오른쪽)

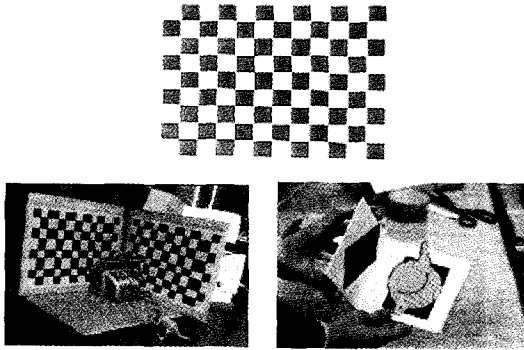
또 다른 분류 방법으로는 특정한 표식인 마커(marker)를 이용하는 방식과 마커를 이용하지 않고 실제 영상 자체의 특징을 이용하는 방식이 있다. 마커를 사용하는 방식은 특징을 추출하기 쉬운 색이나 패턴을 이용한다. 특징이나 패턴을 추출 및 추적하기 쉽기 때문에 빠른 속도와 정확도를 유지할 수 있다[3][4]. 마커를 사용하지 않는 방식은 자연 영상에서 특징이 될만한 요소들을 추출하고 그 특징들을 추적하게 된다. 이 방식은 마커를 이용한 방식보다 자연스럽게 다양한 동영상에 활용 가능하지만 정확도를 보장할 수 없다[5][6].



(그림 2) 마커 방식(왼쪽)과 비 마커 방식(오른쪽)

다른 분류 방법에는 카메라의 Calibration을 이용하여 정확한 위치를 찾아내고 3차원의 가상 객체를 거리 정보와 함께 합성할 수 있는 Calibration 방식과, 임의의 동영상에서 거리 정보를 배제한 가상 영상 합성을 실시하는 Non-Calibration 방식이 있다.

우선 Calibration 방식은 규격화된 패턴을 이용하여 카메라의 파라미터를 조율한 다음 합성하고자 하는 가상의 객체를 3차원 거리 정보를 주어 합성하는 방식이다. 카메라의 내부, 외부 파라미터를 비교적 정확히 계산 가능하기 때문에 가상의 객체를 다른 물체들 사이에도 합성할 수 있다[7]. Non-Calibration 방식은 카메라 내부, 외부 파라미터들을 계산하지 않고 영상 자체의 직선의 정보와 평면의 정보를 이용하여 간단한 3차원 정보를 추출할 수 있는 방식으로 일반적으로 생성된



(그림 3) Calibration 패턴(상)과 Calibration 방식(하/왼쪽) 및 Non-Calibration 방식(하/오른쪽)

동영상에 사용 가능하다[8].

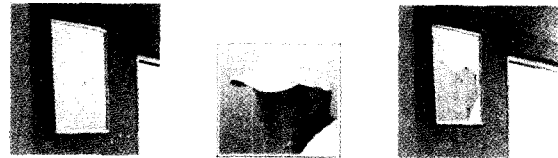
그 외의 분류에 의한 방법들은 방송에서 스포츠 중계 방송시 특수 하드웨어를 이용하여 카메라 움직임 파라미터를 즉각적으로 계산하여 영상 합성 서비스를 실시할 수 있는 시스템[9]과 이미 제작된 동영상들에서 특징 추출 및 추적을 통해 새로운 영상 객체들을 합성함으로써 다양하게 재사용할 수 있는 후처리 방식이 있다.

3. 비디오 기반 증강 현실 구현

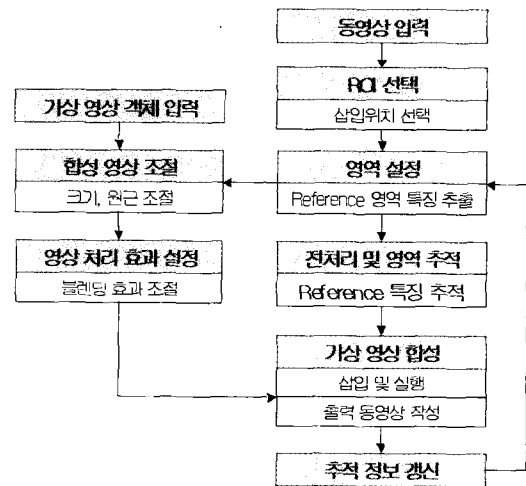
고려대학교 컴퓨터 비전 및 패턴 인식 연구실과 (주)버추얼미디어에서는 후처리 방식인 비디오 기반의 증강 현실에 관해 연구하고 있다. 이를 위해 필요한 기술들은 특징이 될 수 있는 요소들을 추출하는 기술과 그 특징들을 추적할 수 있는 기술들이 바탕이 된다. 또한 합성을 위해서는 원근 변형 모델을 사용하고 있으며 2차원 가상 객체 합성을 자연스럽게 하기 위해 영상처리 기술들을 사용하고 있다.

가상의 이미지를 실 영상에 적용하기 위해서는 먼저 가상 영상을 합성하기 위한 위치가 지정되어야 하고 지정된 영역에 대한 영상의 움직임 정보가 산출되어야 한다. 또한 산출된 움직임 정보를

통해서 동영상 내에서 재생 중에 자연스러운 합성과 함께 합성된 영상을 출력할 수 있어야 하며, 변형된 좌표에 대해서도 처리가 가능한 시스템이어야 한다. 실례로 스포츠 영상에서는 규격화된 필드가 존재하므로 그 필드의 정보를 얻어냄으로써 정확한 정보를 가지고 수행할 수 있으나 그 외의 드라마나 쇼프로그램과 같은 방송용 자연 영상 데이터에 대해서는 정확한 처리를 위해서 스포츠 영상보다 복잡한 추적 방법이 적용될 수밖에 없다.



(그림 4) 실영상(왼쪽)과 합성될 영상(중간) 및 합성 결과(오른쪽)



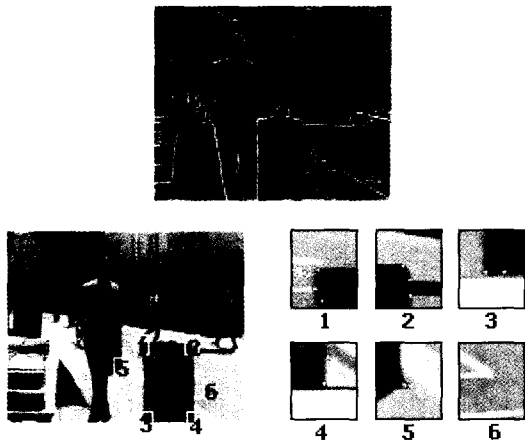
(그림 5) 추적 정보 구성도

위 (그림 5)에서의 대표적인 기술들에 관해서 알아본다.

3.1 Reference 영역의 특징 추출

우선 사용자의 입력으로 4개의 합성될 영상의

코너점을 입력받는다. 이 코너점들의 내부의 특징점들을 추출하여 기본 특징점으로 사용한다. 현재 프레임에서 요구되는 특징점을 추출하기 위해서 기본적으로 영상의 특징 정보로 범용적으로 쓰이는 코너 정보를 이용한다. 이 때, 추출된 특징점에 대하여 최대한의 노이즈를 제거하기 위하여 미분을 통한 영상에서 나타나는 전처리 영상의 첨예화를 가질 수 있는 도함수 필터(Derivative filters)를 적용하여 특징 정보에 대한 양을 최소화 하고자 하였다.



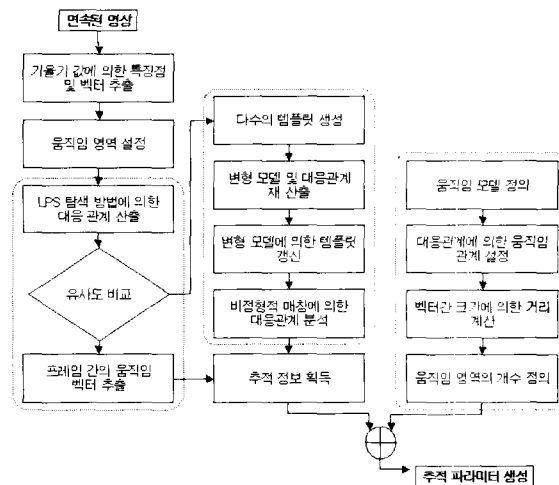
(그림 6) 특징 추출 방식

3.2 영역 추적

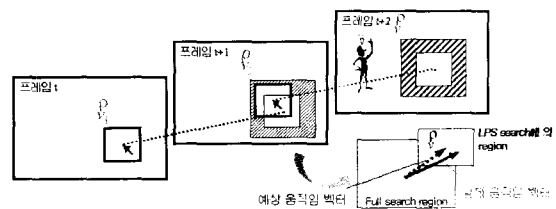
시작 프레임에서 결정된 특징점들을 이용하여 optical flow 기법에 적용하여 프레임 사이의 특징점의 대응관계를 추출한다. Lucas-Kanade 방법을 이용하였다. 특징점들 사이의 대응관계는 Normalized Correlation 방법을 이용하였다. 다음은 특징점들에 대한 추적을 위한 프레임 단위의 특징점 보정 및 추적에 필요한 특징점 추출 과정을 설명하고 있다.

추적 과정에 대한 보다 자세한 과정은 아래의 그림과 같으며 국부적 움직임 정보와 전역적 움직임 정보를 이용하였다. 또한 추적 정보를 통해 나타

난 동일 평면상의 특징점들을 이용하여 원근 변형 파라미터를 추출하였다. 선정된 관심 영역에 대하여 추출된 특징 벡터를 가지고 움직임 영역을 분석한다. 움직임 분석은 계산량을 고려하여 인접한 장면에 대해 예상되는 움직임 정보를 얻고 이웃하는 프레임에서 최종 움직임을 찾는다. 그리고 움직임 벡터에 대해 관심 영역 내의 LPS (Linear Prediction Search) 움직임 모델을 전역 탐색(FS: Full Search)에 적용한다. 즉, 예상되는 움직임의 벡터와 유사한 모델이 존재하지 않을 경우, 국부적 영역을 배제하고 전역 영상에 대한 움직임을 분석한다.



(그림 7) 움직임 모델을 이용한 추적

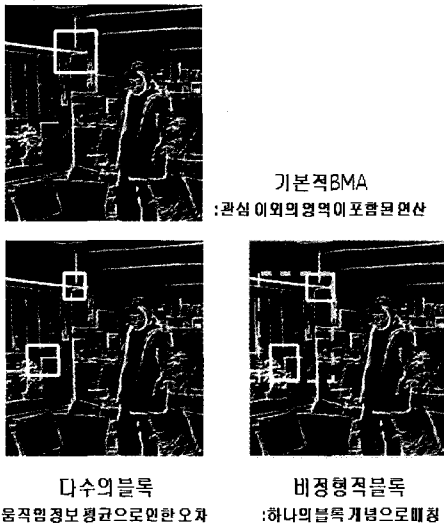


(그림 8) 움직임 분석 방법

추출된 움직임 정보에 의한 특징 벡터는 국부적 영역에 의한 특징 벡터이므로 흔들림 현상에 민감

하다. 이러한 특정 영역의 움직임에 대한 평균 오차를 줄이기 위해 다수의 비정형적 형판(Template)을 생성하여 매칭 알고리즘을 통해 다시 한번 전역 정보에 대한 움직임을 분석한다.

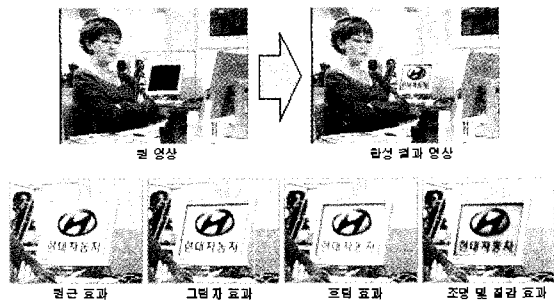
생성된 형판은 갱신이 될 경우, 누적되는 에러로 인하여 합성 영역에 대하여 가상 이미지가 이탈할 염려가 있다. 그러므로 생성된 형판은 고정시킨 상태에서 매칭 스코어를 통해 선별된 특징점으로 이용하고 위의 필터링을 통하여 흔들림에 대한 보정을 시도한다.



(그림 9) 비정형 블록을 이용한 매칭

3.3 가상 영상 객체 합성

영상을 합성하기 위해서 설정된 원 영상의 ROI 내에 가상의 영상을 합성하기 위해서, 그리고 단순 합성 후에 원 영상과의 자연스러운 조화를 위하여 특징 추적 과정의 정보를 이용하고 설정된 영역의 원근 변형 파라미터를 추출하여 합성될 가상의 객체에 원근 변형과 블렌딩, 블러링과 같은 여러 효과 필터를 거치게 된다.



(그림 10) 가상 객체의 자연스러운 합성 방법

위의 기술들 이외에도 추적 정보를 보다 세밀하게 하기 위해 sub-pixel을 통한 정확도 개선을 실시하였으며 방송용 동영상에서의 실험에서 발생하는 모션 블러를 제거하기 위해 프레임을 필드로 나누어 처리하였다.



(그림 11) 실험 결과 : TV 시트콤에서의 가상 영상 객체 합성



(그림 12) 실험 결과 : TV 드라마에서 가상 영상 객체 합성

4. 실험 및 결과

비디오 기반 증강 현실 기술이 실제로 사용 가능한 분야로는 현재 영화와 비디오 시장이 있다. 이를 위해서 방송용 동영상에 적용 가능한 시스템을 구성하기 위해 다음과 같은 사양의 하드웨어 및 소프트웨어 사양에서 실험하였다. 다양한 형식으로 입/출력이 가능하게 하기 위하여 NTSC (720×486)영상을 사용하였고, 확장 AVI 파일 형식(MJPEG: Motion JPEG)으로 저장하여 사용한다. DigiSuite LE Board를 통해 압축된 비디오는 적절한 속도로 저장되고 재생되기 위해 초당 40MB 정도의 전송 속도를 갖는 Ultra wide SCSI를 사용한다.

5. 결 론

지금까지 비디오 기반 증강 현실의 구현 방식과 실제로 적용 가능한 분야에 관해 알아보았다. 현재 동영상을 이용한 다양한 서비스나 멀티미디어의 분야가 확대되어 가고 있는 추세에서 비디오 기반 증강 현실의 적용 범위는 무한하다고 할 수 있다. 향후에는 한 객체 특징 정보를 추출하여 데이터베이스의 다양한 객체와 매칭을 통하여 인식해 내는 기술을 적용하여 사용자의 선택에 따라 필요한 정보를 대화형 방식으로 제공할 수 있는 기술을 (주)버추얼미디어와 고려대학교 컴퓨터비전 및 패턴 인식 연구실에서 개발하고 있다. 이를 적용할 경우 가상의 객체 합성 뿐만 아니라 동영상 시청자들의 요구를 즉각 반영한 부가 정보 서비스가 가능할 것이다.

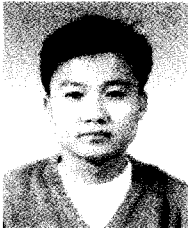
참고문헌

- [1] R. Azuma, "A Survey of Augmented Reality", PRESENCE: Teleoperators and Virtual Environments, Vol. 6, No. 4, pp. 355-385, August 1997.
- [2] www.symah-vision.fr
- [3] K. Dorfmüller, "Robust tracking for augmented reality using retroreflective markers", Computers & Graphics 23, pp. 795-800, 1999.
- [4] P. Liu, N. D. Georganas and P. Boulanger, "Designing real-time vision based augmented reality environments for 3D collaborative applications", IEEE Canadian Conference on Electrical and Computer Engineering, Vol. 2, pp. 715-720, May 2002.
- [5] U. Neumann and S. You, "Natural Feature Tracking for Augmented Reality", IEEE Trans. on Multimedia, Vol. 1, No. 1, pp. 53-64, March 1999.
- [6] D. Stricker and T. Kettenbach, "Real-time and Markerless Vision-Based Tracking for Outdoor Augmented Reality Applications", IEEE International Symposium on Augmented Reality 2001.
- [7] Y. D. Seo and K. S. Hong, "Auto-Calibration of Rotating and Zooming Camera", IAPR Workshop on Machine Vision Applications, Chiba, Japan, 1998
- [8] D. Stricker and T. Kettenbach, "Calibration-free Augmented Reality", IEEE Trans. on Visualization and Computer Graphics, Vol. 4, No. 1, pp.

1-20, 1998.
[9] G. Medioni, G. Guy, H. Rom and A. Fracois, "Real-time Billboard Substitution in a Video Stream", International

Workshop on Digital Communications multimedia communications, Ischia, Italy, pp. 71-84, 1998.

저자약력



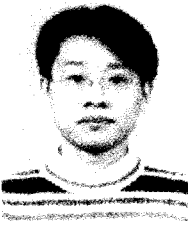
박상철

1996년 전남대학교 전산학과 (학사)
1998년 전남대학교 전산학과 (석사)
1999년-현재 고려대학교 컴퓨터학과 박사 수료
관심분야 : 증강 현실, 3D 비전, 그래픽스, 패턴 인식
이 메 일 : scpark@image.korea.ac.kr



이성환

1984년 서울대학교 계산통계학과 (학사)
1986년 KAIST 전산학과 (석사)
1989년 KAIST 전산학과 (박사)
1989년-1994년 충북대학교 컴퓨터학과 조교수
1995년-현재 고려대학교 컴퓨터학과 정교수
1997년-현재 고려대학교 인공지능연구센터 연구소장
2001년 미국 MIT AI Lab. 객원 교수
관심분야 : 컴퓨터 비전, 패턴 인식, 영상 처리
이 메 일 : swlee@image.korea.ac.kr



황본우

1995년 성균관대학교 전자공학과 (학사)
1997년 성균관대학교 전자공학과 (석사)
2002년 고려대학교 컴퓨터학과 (박사)
2002년-현재 (주)버추얼미디어 기술연구소장
관심분야 : 얼굴 복원, 증강 현실, 얼굴 인식, 패턴 인식
이 메 일 : bhwang@virtualmedia.co.kr