

증강현실의 기술과 동향

(Technology and Trend of Augmented Reality)

서 용 덕, 김 종 성, 홍 기 상

포항공과대학교 전자전기공학과 영상처리연구소

I. 서 론

요즘의 영상 저작물에서 특수 효과를 접하는 것은 매우 평범한 일이 되었고, 이전에는 구현하기 어려웠던 다양한 장면들의 연출을 통해 영상을 보는 이들은 여러 가지 환상적인 경험을 하게 되었다. 영화, 텔레비전 프로그램, 애니메이션, 광고 등에서 특수효과가 차지하는 비중은 예전의 그것과는 비교할 수 없이 커져 있으며, 보는 이들에게 좀 더 강렬한 인상을 남기기 위해서 제작자들은 새로운 기술이나 방법을 찾아내지 않으면 안되는 때가 되었다. 특수 효과를 얻는 데는 여러 가지 방법이 있는데 그 중에서도 컴퓨터 그래픽과 실제 영상을 합성하여 제작하는 기법이 많이 쓰이고 있으며 그 기술 자체에 대한 관심도 커지고 있다. 단지 하나의 특수효과를 담은 영상을 연출하고 제작하는데 있어서도 많은 분야의 기술자들이 필요하다는 것은 잘 알려진 사실이다. 특히, 컴퓨터 그래픽을 이용한 특수효과 장면의 합성을 위해서는 기본적으로 고도의 영상 처리 기술 및 컴퓨터 비전 기술이 필요하지만, 그러한 영상 제작에 있어 실제로 어떤 방법이 사용되고 있는지에 대해서는 많이 알려져 있지 않은 것 같다. 이 글에서는 컴퓨터 그래픽을 이용한 특수효과를 얻는 여러 가지 방법 가운데 특히 증강현실과 관련된 것들에 대해서 살펴보고자 한다.

II. 증강현실 기술

실제 영상과 가상의 컴퓨터 그래픽을 합성하는 방법으로 두 가지를 생각해 볼 수 있다. 처음 한 가지는 컴퓨터 그래픽을 이용하여 동영상이나 어떤 장면을 합성할 때 미리 복원해 둔 실제 물체의 삼차원 모델을 사용하는 것이다. 컴퓨터 그래픽으로 삼차원 동영상을 제작하는 경우 영상제작에 동원되는 모든 환경은 컴퓨터 그래픽에서 다룰 수 있는 형태의 것들이어야 한다. 따라서, 그에 필요한 물체들의 그래픽 모델이 필요한데 현실감을 높이기 위해서 사용할 수 있는 한 가지 방법이 그 모델을 실제 영상으로부터 복원하는 것이다. 실제로 동영상은 가상의 그래픽 카메라가 움직이는데 따라 만들어지기 때문에 모든 작업들은 컴퓨터그래픽에 의한 동영상 제작과정과 일치하게 된다. II.1절에서는 몇 장의 실제 영상으로부터 복원하고자 하는 물체의 삼차원 구조를 계산하여 VRML 등의 형태로 표현하는 삼차원 모델 복원 방법에 대해서 설명한다.

또 다른 한 가지는 비디오카메라로 찍은 동영상에 컴퓨터 그래픽으로 만들어진 가상의 모델을 합성하여 넣는 것이다. 이렇게 제작된 동영상은 영상 대부분이 실제로 존재하는 물체들로 구성되어 되는데, 최종 결과물을 만들기 위해서는 원래의 동영상 중에서 특별한 부분을 컴퓨터 그래픽을 이용하여 특정 부분들의 형태나 색깔 등을 제작자의 의도에 따라 다르게 바꾸거나 원래는 존재하지 않는 어떤 물체를 마치 실제로 존재하는

것처럼 삽입하는 과정을 거쳐야 한다. 이를 위해서는, 복원된 삼차원 물체를 사용하는 앞의 방법과 달리 실제 카메라의 움직임에 대한 정보를 동영상으로부터 계산(복원)하고 그 움직임에 따라 컴퓨터 그래픽을 이용하여 가상의 물체에 대한 동영상을 따로 제작한 다음 원래의 동영상과 합치는 과정을 거쳐야 한다. 이 과정은 II. 2절에서 설명한다.

사실, 이러한 두 가지 분류는 실제 영상으로부터 어떤 정보를 더 원하는가에 따라 이루어진 것이다. 실제로, 어떤 물체의 삼차원 VRML 모델을 여러 장의 영상에서 얻고자하는 경우 그 영상들을 얻는데 사용된 카메라에 대한 위치, 방향, 초점거리 등의 정보를 알아야만 하는데, 이를 구하는 과정을 카메라 자동 보정(camera self-calibration)이라고 부르며, 요즘은 투영기하방법(projective geometric method)를 적용함으로써 카메라 보정에 필요한 사진 정보의 양을 획기적으로 줄이게 되었다. II. 1절과 II. 2절에서 자세히 설명하게 되겠지만, 카메라 보정 문제와 삼차원 모델 복원 문제는 서로 맞물려 있는 관계이기 때문에 실제 계산과정에서는 카메라 보정과 삼차원 모델을 동시에 계산하게 되어 있다. 다만, 최종 결과로 구하고자 하는 것이 어떤 것이냐에 따라 삼차원 모델 복원과 카메라 자동 보정이라는 두 가지 문제를 따로 다루게 된 것이다.

1. 삼차원 모델 복원

물체의 3차원 모델을 여러 장의 영상 또는 연속적인 영상열(image sequence)로부터 만들고자 하는 노력이 많이 있었다^[4]. 이러한 3차원 모델링 과정은 카메라 파라미터와 영상열 각 프레임의 카메라간의 상대적인 위치를 계산하고, 이러한 정보를 이용하여 영상열에 보이는 물체의 대략적인 3차원 구조를 계산하는 SFM(Structure From Motion) 단계와 복원하고자 하는 물체의 특징을 부각시키는 방법 등으로 전체 구조를 실제 물체와 유사하게 모델링하는 단계로 이루어져 있다. 최근에 자동으로 SFM을 적용하

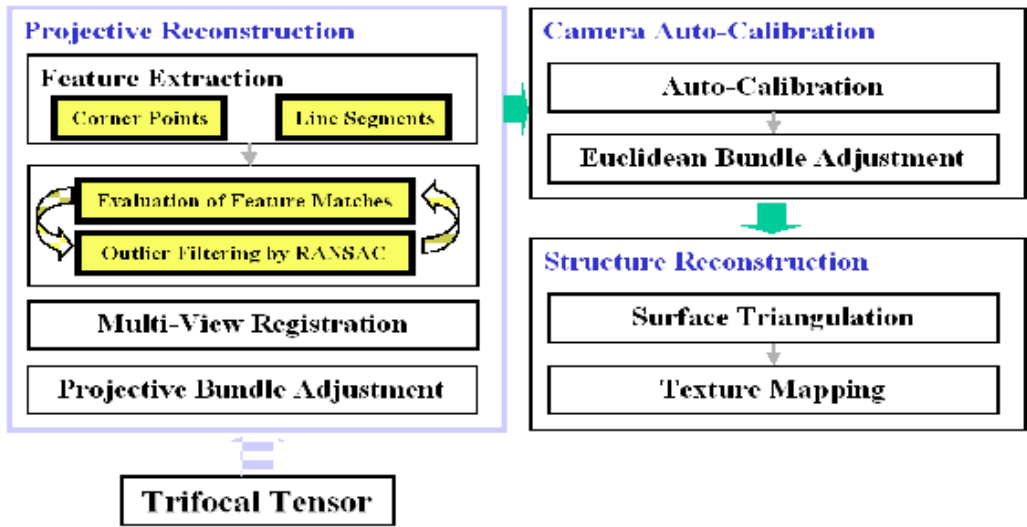
는 알고리즘에 대한 연구가 많이 진행되고 있으며, 모델링 과정에서도 사용자 개입(interaction)을 최소로 하는 알고리즘에 대한 연구가 진행중에 있다.

1) Trifocal tensor

임의의 3차원 구조를 보고 있는 세 개의 투시뷰(perspective view) 사이에 대수학적 연결 관계인 trilinearity가 있다^[2]. 이러한 관계를 수학적으로 모델링 한 것이 trifocal tensor이다. Trifocal tensor는 3차원 구조를 분석하는데 있어서, 두 개의 뷰에서 fundamental matrix가 하는 역할이 유사한 역할을 세 개의 뷰에서 한다. 이러한 trifocal tensor는 영상에서 특징점과 직선을 정합하고, 투영 카메라를 계산하며, 투영 구조를 복원하는데 이용된다. Trifocal tensor를 이용하여 영상열의 카메라를 복원하거나 특징점들을 연결하는 등의 SFM 시스템에 필요한 기능이 기존 에피폴라 기하(epipolar geometry)를 이용한 방법보다 더 정확한 결과를 얻을 수 있는 것으로 알려져 있다. 그래서 이러한 trifocal tensor를 세 개의 연속적인 영상에서 특징점으로 추출되는 점과 선을 동시에 사용하여 계산하는 등의 방법적인 면에 대한 많은 연구가 진행되었다^[3]. 대표적으로 세 개의 뷰에서 초기 정합된 특징점과 직선들로부터 RANSAC 알고리즘을 사용하여 trifocal tensor를 계산하는 방법이 있다. 이 때 필요한 fundamental matrix는 중간 과정에서 계산되는 trifocal tensor로부터 계산하여 사용한다. 최종 계산된 trifocal tensor는 위와 같이 투영 카메라의 계산에 사용되며, 이렇게 계산된 카메라들과 영상의 특징점을 사용하여 구조물의 3차원 투영 구조를 복원하였다.

2) Trifocal tensor 기반 SFM 시스템

Trifocal tensor를 이용한 SFM 시스템의 개요는 <그림 1>과 같다. SFM 시스템은 크게 다음과 같이 세 부분으로 나눌 수 있다.



〈그림 1〉 Trifocal Tensor 기반 SFM 시스템 개요

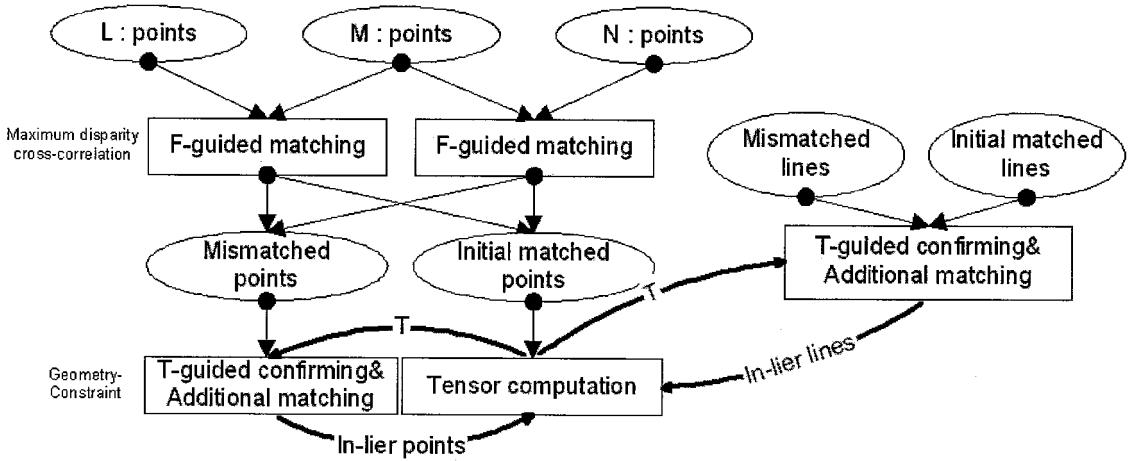
- 투영 기하 복원(Projective Reconstruction)
- 카메라 자동 보정(Camera Auto-Calibration)
- (유클리드) 구조 복원

투영 기하 복원 단계는 영상열에서 추출되는 특징점과 특징선들을 정확하게 연결하여, 영상열에서 관찰되는 2차원 특징들과 우리가 복원하여 모델을 만들고자 하는 3차원 구조와의 초기 관계를 계산한다. 이러한 관계는 실제 영상을 획득한 카메라의 파라미터들과 카메라간의 상대적인 관계를 파악함으로써 알 수 있다. 그러나 영상열에서 관찰되는 특징점들로부터 얻을 수 있는 카메라 정보는 실제세계의 3차원 유클리드 기하가 아닌 투영 기하이다. 이러한 투영 기하 정보를 유클리드 기하 정보로 변환하기 위해서 두 번째 단계인 카메라 자동 보정이 필요하다. 자동 보정은 2차원 영상 정보의 기하학적인 특징을 이용하여 투영 기하에서 유클리드 기하로 3차원 구조를 변환하는 동시에 실제세계에서의 카메라 변수(초점 거리, 카메라 중심, 축 비율, 비틀림 상수 등)와 카메라 간의 상대적인 위치를 정확하게 계산해 준다. 자동 보정 단계를 지나면 이제 우리가 원하는 모델의 3차원 기하 정보를 알 수 있으며, 최종적

으로 컴퓨터 그래픽스 기술을 이용하여 우리가 사용할 수 있는 그래픽 모델을 만들어 사용하게 된다.

3) 영상열의 특징점 연결 및 trifocal tensor 계산

SFM 시스템의 핵심 기술은 영상열에서 정확하게 카메라를 계산하는 것이다. 영상 열에서 카메라들을 계산을 위해서는 먼저 2차원 영상의 특징점이나 특징선들을 에러없이 연결할 수 있어야 한다. 이렇게 연결된 특징점과 특징선을 이용하여 투영 기하 아래의 초기 3차원 구조 및 카메라 위치를 계산할 수 있다. 이러한 과정은 Trifocal tensor를 계산하는 과정과 결합될 수 있으며, 세계의 뷰에서 전체적인 처리 과정은 〈그림 2〉와 같다. Trifocal tensor의 계산은 아주 정밀한 값을 요구하므로 잘못된 특징점이나 특징선의 연결이 들어가서는 안된다. 이러한 잘못된 연결(Outlier)를 제거하는 방법으로 LMedS(Least Median Square)나 RANSAC(Random Sampling Consensus) 기법이 사용된다. 이러한 기법들은 특정 모델 파라미터 값을 추정하는데 있어서 측정 데이터에 노이즈 또는 오류에 의해 잘못된 데이터가 포함되더라도 최적의 값을 계산해



〈그림 2〉 세 개의 뷰에서의 Trifocal tensor 계산 및 특징점, 특징선의 연결



Trajectories of tracked points



Trajectories of tracked lines

〈그림 3〉 세 개의 뷰에서 연결된 특징점과 특징선의 궤적

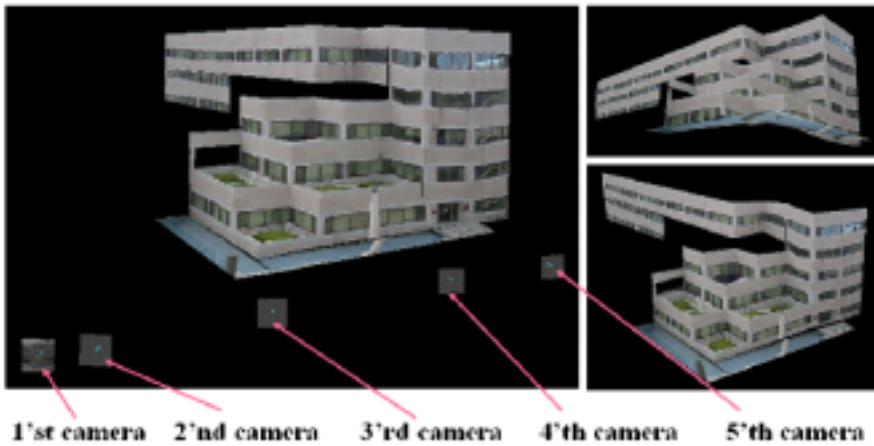
주는 추정 알고리즘이다. 이러한 과정을 거쳐 trifocal tensor를 정확히 계산한 후 카메라 복원단계로 넘어간다. 위 〈그림 3〉에는 실험 영상에서 추출되고 연결된 특징점과 특징선으로, trifocal tensor 계산 및 초기 3차원 투영 구조의 복원에 사용된다. 세 개의 뷰 단위로 연속된 영상열에서 계속적으로 계산된 trifocal tensor들과 특징점과 특징선들은 임의의 기준 좌표계를 중심으로 정렬하는 다중 뷰 정렬(Multi-View Registration) 처리를 통하여 통합된다. 이렇게 통합된 값들은 투영 통합 최적화(Projective

Bundle Adjustment)를 거쳐 투영 기하 아래에서 발생한 에러를 최소화 한다.

4) 카메라 계산 및 3차원 구조 복원

Trifocal tensor에서 3차원 복원에 필요한 투영 카메라와 초기 투영 구조를 계산해 낼 수 있다. 투영 기하에서의 카메라 P_k^P 와 2차원 영상 특징 x_{ki} 과 3차원 구조 X_i^P 의 관계는 다음과 같다.

$$\lambda_{ki}x_{ki} = P_k^P X_i^P, \\ P_k^P = [Q_k | q_k]$$



〈그림 4〉 복원된 카메라 위치 정보와 그래픽 모델

투영 기하를 유클리드 기하로 변환하여 실세계와 같은 기하 정보를 추출하는 데는 다음과 같은 관계식이 이용된다.

$$P_k^P T \cdot T^{-1} X_i^P = \lambda P_k^E X_i^E$$

그러므로 자동 보정은 변환에 필요한 매트릭스 T 를 추정하는 것과 같다고 할 수 있다^{[8][9]}.

자동 보정을 통하여 투영 기하에서 유클리드 기하로 변환하면 카메라 P_k^E 와 3차원 구조 X_i^E 는 다음과 같이 변환된다.

$$\lambda_{ki} X_{ki} = P_k^E X_i^E, \\ P_k^E = K_k [R_k | t_k]$$

이렇게 복원된 카메라 정보와 3차원 구조 정보에는 우리가 원하는 그래픽 모델을 만드는 데 필요한 모든 정보가 다 들어가 있다.

5) 유클리드 구조 복원

자동 보정 단계를 통하여 복원된 3차원 정보를 그래픽 모델로 만들기 위해서 먼저 3차원 데이터를 데이터 삼각법 (Triangulation)를 이용하여 다각형 메쉬 모델 (Polygonal mesh model)로 만들어야 한다. 그런 후에 텍스처 삽입을 통해서 모델의 현실성을 증가시킨다. 위 〈그림 4〉는 실험 영상에 대한 최종 그래픽 모델을 보여주고 있다^[7].

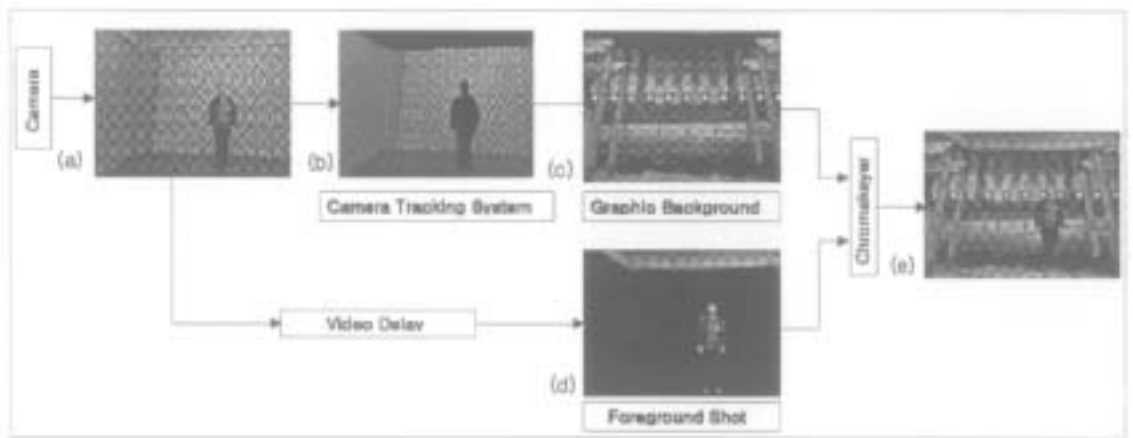
2. 카메라 보정에 의한 증강현실

1) 실시간 카메라 보정

기준좌표계로부터 카메라 정보 즉 카메라의 상대적인 위치와 방향 및 초점거리 등을 계산하기 위해서는 카메라자체에 하드웨어 센서를 부착하는 방법을 이용할 수도 있고, 카메라에서 주어지는 각 영상 프레임에 영상처리 알고리즘을 통해 분석하여 얻는 방법을 생각할 수도 있다. 이러한 두 가지 경우는 주로 실시간 증강현실 구현을 위해서 사용하는 방법인데, 스포츠 중계나 역사 다큐멘터리 등에서 각 방송사들은 전자의 경우를 더 많이 사용하는 것으로 알려져 있다. 그러나, 하드웨어 부착 방법은 카메라의 유동성이 떨어진다는 단점이 있다. 좀더 역동적인 영상구성을 위해서는 해결해야 할 과제들이 많이 있지만 결국 카메라 정보를 얻는데 있어 영상처리방법을 사용하는 것이 하나의 해결방법이 될 것으로 보인다.

영상처리에 의한 카메라 보정방법이 실시간으로 처리되기 위해서는 영상분석에 필요한 계산시간을 줄이고 영상으로부터 카메라 보정에 필요한 정보를 빠르고 쉽게 얻어내는 것이 중요한 관점이 된다.

〈그림 5〉는 영상처리에 의한 실시간 카메라 보정시스템과 그로부터 얻어진 카메라 정보를 SGI 그래픽스 컴퓨터와 연동하여 구현한 실시간 증강



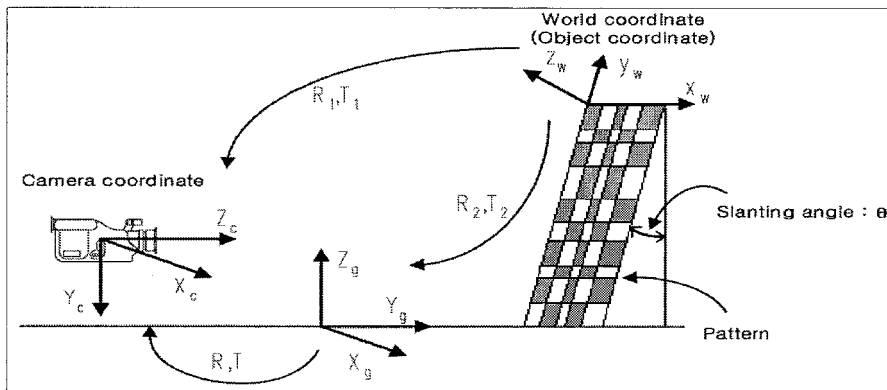
〈그림 5〉 영상처리에 의한 실시간 카메라 보정 방법을 이용하여 구현한 증강현실 시스템의 정보 흐름도.

현실 시스템의 전체적인 정보 흐름도이다. 카메라에서 실제로 얻어진 영상의 배경에는 카메라 보정을 위하여 미리 디자인하여 제작한 보정패턴이 설치되어 있는 것을 볼 수 있으며, 최종적으로 얻어진 영상에서 뒷배경은 컴퓨터 그래픽에 의해 합성된 화면으로 채워져 있다. 이러한 최종 결과를 얻기 위해서는 보정패턴에 고정된 좌표계와 카메라 좌표계 그리고 그래픽을 위한 좌표계 사이의 관계를 미리 설정해 두어야 하며, 카메라와 보정 패턴 사이의 상대적인 좌표변환 관계는 영상처리를 통하여 매 프레임마다 계산을 통해 얻게 된다.

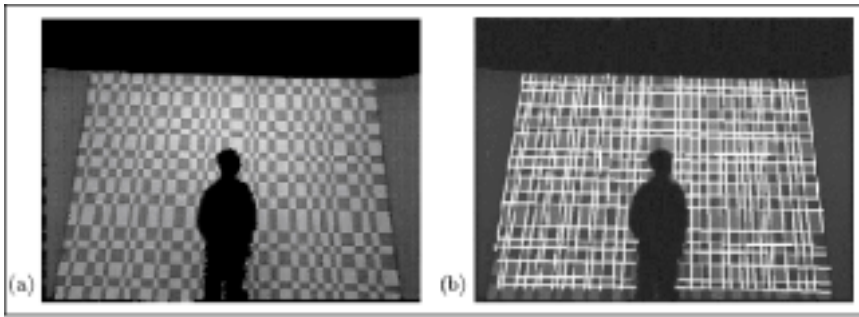
〈그림 6〉에서는 이러한 관계를 도시하였다. 즉,

그래픽 좌표계와 보정패턴 좌표계는 미리 결정되어 있어서 컴퓨터 그래픽에 의해 합성될 가상물체들의 상대적인 위치가 카메라보정단계 이전에 이미 알려져 있는 것이다. 따라서, 카메라로부터 얻은 영상을 분석하여 보정패턴을 인식하고 보정패턴의 삼차원 좌표와 그 좌표의 영상 위치를 알아낸 후 그 둘 사이의 관계를 이용하여 카메라 보정값을 얻게 된다.

카메라 입력 영상에서 보정패턴을 포함한 뒷배경을 없애기 위해서는 크로마-키 방법을 사용한다. 이는 방송기술분야에서는 잘 알려진 것으로 뒤 배경의 파란색 부분을 다른 영상신호로 바꿔치기 하는 방법이다. 〈그림 5〉에서 보인 보정패



〈그림 6〉 카메라 좌표계, 그래픽 좌표계, 보정패턴 좌표계 사이의 관계



〈그림 7〉 실시간 증강현실 구현을 위한 영상처리의 보기. 보정패턴의 가로-세로 선들의 조합은 서로 다른 비조화비(또는 복비 : Cross Ratio) 값들을 가지도록 설계되어 있다. 가장 손쉬운 경우로는 영상의 가로 두 줄 세로 두 줄만 분석하면 보정패턴의 각 점에 해당하는 미리 정해진 삼차원 좌표를 구할 수 있다.

턴은 실제 구현에 있어 명도가 다른 동일한 색상의 파란색을 사용하여 제작되었기 때문에 영상처리를 통한 패턴인식에 필요한 만큼의 충분한 밝기 변화를 주면서도 크로마-키를 이용하여 배경 바꿔치기가 가능하였다.

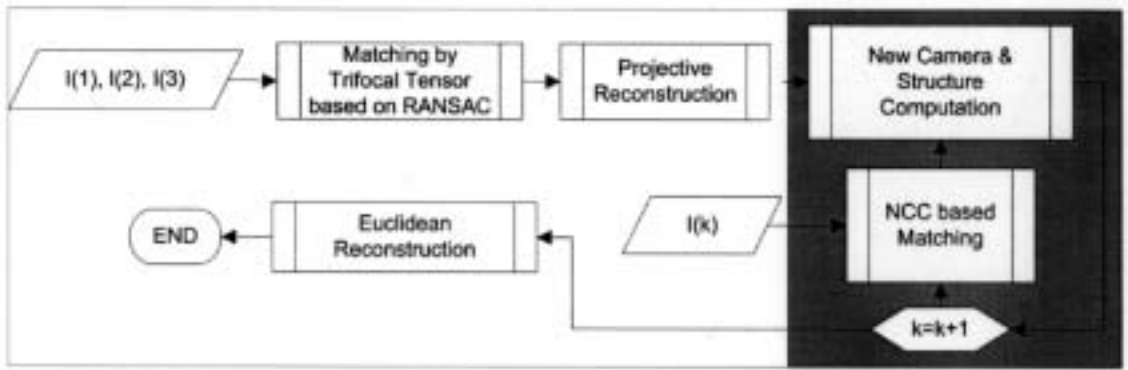
〈그림 7〉은 하나의 카메라 입력 영상에서 영상처리 방법을 통해 보정패턴을 구성하는 선들을 추출한 것을 보여주고 있다. 연속된 4개의 가로 선들 사이의 비조화비(Cross Ratio)는 모두 다른 값들을 가지도록 디자인하여, 영상값의 미분을 통해 얻어진 극대점들로부터 연속된 비조화비 값들을 비교함으로써 영상의 각 선과 실제 모델로 만들어진 선 사이의 대응관계를 손쉽게 찾을 수 있었다. 한편, 카메라의 렌즈에 의한 영상 굴곡이 있는 경우 이를 보상하는 과정을 한 번 더 거쳐야 하는데, 보정패턴의 형태를 이미 알고 있기 때문에 인식된 선들의 정보를 이용하면 카메라 렌즈의 굴곡률을 미리 계산하는 것이 가능하다. 영상처리 및 인식 과정과 보정값 계산 과정을 통해 얻어진 이 차원 영상 좌표값과 삼차원 패턴 좌표값 사이의 관계를 입력으로 하여 Tsai 알고리즘을 사용하면 카메라의 상대적인 위치와 방향에 대한 정보와 카메라의 초점거리에 대한 값을 얻을 수 있다^[1]. 컴퓨터 그래픽 합성에서는 카메라의 위치와 방향을 이용하여 가상 카메라의 위치 방향을 설정하고, 초점거리를 이용하여 가상 카메라의 스케일값을 조절함으로써 최종적으로 그래픽 영상을 합성하게 된다. 그러면, 앞에서 설

명한 바와 같이 크로마-키를 통해 〈그림 5〉의 오른쪽에 보인 것과 같은 결과를 얻게 된다^[5].

2) 카메라 자동 보정

카메라 자동보정의 방법론이 앞 절 II. 2. 1에서 언급한 방법과 가장 다른 점은 카메라 보정을 위한 패턴을 따로 디자인하여 사용하지 않는다는 것이다. 물론 이럴 경우 실시간 계산의 기능은 없게 된다. 〈그림 8〉은 카메라 자동보정 방법에 대한 간단한 흐름도를 보여주고 있다. 구현하는 방식에 따라 다소 차이가 생기긴 하겠지만, 일반적인 동영상을 분석하여 카메라 정보를 얻는 방식의 일반적인 알고리즘은 대체로 그림에서 보여주는 흐름을 따르고 있다. 앞 절 II. 1에서 실물체의 삼차원 모델을 복원하기 위해 사용한 방식을 여기서도 여전히 사용한다. 영상열의 처음 세 장으로부터 카메라정보를 계산하고 이를 전체 영상열로 확장한다는 것이 기본적인 아이디어가 되는데, 초기 계산값을 더 잘 구하기 위해서는 연속된 세 장의 영상이 아니라 영상에서 구한 정합점들이 일정 거리 이상 움직임이 있도록 조절하는 방법을 사용하기도 한다.

초기영상복원을 구한 후에는 영상열의 각 영상에 대해서 이전 시점의 영상 사이의 정합점들을 영상정합(Image based matching : Normalized Cross Correlation(NCC)를 이용) 방법을 통해 구하고 RANSAC 알고리즘을 기초로 새로운 영상에 대한 카메라행렬을 구하며, 새롭게 나타



〈그림 8〉 입력 영상열에서 카메라 자동보정 과정을 통해 카메라 정보를 얻는 과정

난 정합점들에 대해서는 삼차원 좌표를 구한다. 이 과정은 영상정합에 의한 결과 중에서 잘못 얻어진 정합점들을 제거하는 과정이 포함되어 있기 때문에 매우 중요한 단계가 되며 최종적으로 얻어지는 복원 결과에 큰 영향을 미친다.

앞 절에서도 언급했지만, 1992년부터 행해진 카메라 자동 복원에 대한 연구들은 이러한 카메라 행렬과 정합점들의 삼차원 좌표값들을 유클리드 공간 기하학(Euclidean Geometry)이 아니라 투영기하공간기하학(Projective Geometry)을 기초로 진행되었으며 결국 카메라 자동보정방법의 길을 열었다. 앞 절의 실시간 카메라 보정방법의 모든 과정이 유클리드공간에서 정의되는 값들을 근거로 하는 반면, 그림에서 보이는 대부분의 계산과정이 투영기하공간에서 이루어지며 최종 단계에서 투영기하공간에서 정의된 값들을 유클리드공간에서 정의되는 값으로 변환하게 되는데 이 부분 즉 Euclidean Reconstruction 부분이 카메라 자동 보정의 핵심이 되는 부분이라 할 수 있다. 한 가지 주의해야 할 점은 카메라 자동 보정 및 투영기하공간에서의 계산식들은 모두 비선형 방정식으로 되어있기 때문에 최소자승 오차법(Least square error method)으로 구해지는 값들이 원래 방정식을 제대로 따르지 못하는 경우가 많이 생긴다는 것이다. 따라서, 그림에서 도시하지는 않았지만, 비선형 최적화 과정이 항상 필요하며 유클리드공간으로의 변환과정의 최종 단계에서 그리고 투영기하공간에서 복원

값들을 구할 때 최적화 과정을 적절히 배치할 필요가 있다.

영상처리 과정을 포함한 카메라 자동 보정 과정을 통해 얻어지는 카메라 정보는 카메라의 위치와 방향 그리고 초점거리이다. 하지만, 카메라의 위치와 방향은 상대적으로 결정되는 값이기 때문에 컴퓨터 그래픽을 이용하여 가상의 물체를 삽입하기 위해서는 카메라 좌표계와 그래픽 좌표계를 연결해주는 작업이 필요하다. 이 작업은 영상의 정합점들의 삼차원 좌표들을 기준으로 할 수도 있고 〈그림 9〉와 같이 두 영상을 이용하여 할 수도 있다. 여기서 사용한 방법은 두 장의 영상에 그래픽 좌표계의 원점과 기준 축들의 위치를 사용자가 마우스를 이용하여 지정하는 방법을 사용하였다. 원점의 경우는 한 쪽 영상에 원점을 위한 한 점을 선택하면 다른 쪽 영상에 에피폴라 라인이 생기므로 그 라인 위의 한 점을 선택하는 방법을 적용하였고, 그 후 X-축과 Y-축의 위치를 설정하였다. 그림에서 보이는 정육면체는 그래픽 좌표계의 원점을 한 꼭지점으로 하고 각 변의 길이가 1인 단위 정육면체이다. 따라서, 이 정육면체를 기준으로 삽입하고자 하는 그래픽 물체의 위치와 방향을 설정하면 된다.

〈그림 10〉은 카메라 자동 보정으로 얻은 카메라 정보를 이용하여 가상의 건물을 삽입하여 제작한 동영상의 프레임들을 보여주고 있다. 영상을 기준으로 카메라 좌표계와 연결된 그래픽 좌표계의 원점과 축을 기준으로 삽입하고자 하는



〈그림 9〉 카메라 좌표계와 그래픽 좌표계를 연결하기 위한 작업. 카메라 좌표계 속에 그래픽 좌표계의 위치와 방향을 설정해야 하는데, 이미지를 이용한 방법을 사용한다.



〈그림 10〉 가상 건물을 삽입한 동영상의 첫 번째(왼쪽)와 서른 번째 프레임(오른쪽)

가상 물체의 위치와 방향을 설정한 후 OpenGL 그래픽 라이브러리를 사용하여 각 프레임마다의 그래픽 영상을 먼저 만든다. 그 후 입력 동영상의 각 프레임 위에 제작된 그래픽 동영상의 프레임을 겹쳐 없으면 그림에서 보인 바와 같은 결과물을 얻을 수 있다^[6].

III. 결 론

II절에서 설명한 방법들을 구현한 소프트웨어들 중에는 이미 상용화된 것들이 있다. 실시간 증강현실을 위한 방송 기술을 개발 하는 업체들의 이름은 이미 잘 알려져 있으며, 요즘의 텔레비전 프로그램이나 스포츠 중계에서 그 응용 사례들을

쉽게 접할 수 있다.

카메라 자동 보정 방법과 관련하여 일반화된 것들을 들자면, 영국 2d3사의 bouzou라는 이름의 소프트웨어와 프랑스 RealViz사의 match mover가 있는데, 이들 소프트웨어는 카메라 자동 보정 알고리즘과 최적화 알고리즘 등을 탑재하고 있으며 이미 Enemy at the gate, AI, Pearl Harbor 등과 같은 영화의 특수 효과 제작에 사용된 것으로 알려지고 있다. 위 소프트웨어들이 사용하는 알고리즘들이 어떤 것인지는 알려져 있지 않으나, 지난 10여 년간 발표된 카메라 자동 보정 및 삼차원 복원 논문들의 내용을 적절히 사용하고 있는 것으로 예상할 수 있다.

그러나, 잘 알려진 바와 같이, 영상 처리 및 컴퓨터 비전 (카메라 및 삼차원 복원) 알고리즘은 입력 데이터의 구성 조건에 따라 여러 가지 많은

변화를 가지게 된다. 한 가지 예로, 입력 영상이 거의 평면에 가깝다거나 카메라의 움직임이 직선 운동만을 한다는 등의 조건에서 제대로 동작하는 상용 소프트웨어들을 찾기는 어렵다. 또, 컴퓨터 비전 알고리즘을 적용하고자 하는 응용 사례들마다 각각 특성이 있기 때문에 일반적으로 사용할 수 있는 소프트웨어를 제작하는 것은 매우 힘든 일이라 할 수 있다.

현재 국가가 중점적으로 육성하고자 하는 기간 산업 중에 영상산업이 들어있다는 사실은 컴퓨터 비전이나 영상처리 관련 구성원들에게 큰 기대를 가지게 한다. 고급 영상물의 제작에는 예술의 측면이 강하지만, 그러한 예술적 구상을 실제 화면에 담아내기 위해서는 과학 기술의 도움이 절실히 필요하기 때문이다. 한국의 영상 산업 발달의 근간에 과학 기술 인력의 핵심적 역할을 기대할 때가 된 것이다.

참 고 문 헌

- [1] Roger Y. Tsai, A Versatile camera calibration technique for high accuracy 3-D matching vision metrology using off-the-shelf TV cameras and lenses, IEEE Journal of robotics and Automation, Vol. 3, No. 4, 1987.
- [2] Amnon Shashua, Trilinearity in Visual Recognition by alignment, ECCV, LNCS, Vol 800, pp.479-484, 1994
- [3] R.I.Hartley, Lines and Points in Three Views and the Trifocal Tensor, IJCV, 22(2) : 125-140, 1997
- [4] A.W.Fitzgibbon and A.Zisserman, Automatic camera recovery for closed or open image sequence, ECCV, Vol.1, pp. 311-326, 1998
- [5] 박 성우, 서 용덕, 홍 기상, Real-time camera calibration for virtual studio, Real-Time Imaging, No. 6, Dec. 2000.
- [6] 서 용덕, Anders Heyden, 홍 기상, Flexible Auto-Calibration and its application to augmented reality, IAPR Workshop on Machine Vision Applications, Tokyo, Japan, Nov. 2000.
- [7] 김 중성, Trifocal Tensor 기반 SFM과 EM 알고리즘을 이용한 평면 구조물 복원에 관한 연구, 석사 학위 논문, 포항공대, 2001
- [8] 서 용덕, 홍 기상, Calibration-Free Augmented Reality in Perspective. IEEE Transactions on Visualization and Computer Graphics 6(4) : 346-359, 2000
- [9] 서 용덕, 홍 기상, A theory and practice of self-calibration of a rotating and zooming camera from two views, IEE proceedings on Vision, Image and Signal Processing, Vol. 148(3), pp.166-172, June, 2001

저자 소개



徐龍德

1992년 2월 경북대학교 전자공학과 (학사), 1994년 2월 포항공과대학 대학원 (석사), 2000년 2월 포항공과대학 대학원 (박사), 1997년 7월~2000년 2월 : 포항공과대학 전자전기공학과 위촉 연구원, 1999년 4월~5월 : Lund Univ., Mathematical Imaging Group, 방문연구원, 2000년 5월~2001년 4월 : Cambridge Univ., Engineering Dept., SVR Lab., 방문연구원, 2001년 5월~현재 : 포항공대 전자전기공학과 BK21 연구원, <주관심 분야 : SFM, Auto-Calibration, AR>



金鐘成

2000년 2월 고려대학교 전과공학과 (학사), 2002년 2월 포항공과대학 대학원 (석사), 2002년 1월~3월 : Microsoft Research Center, Asia 방문학생, <주관심 분야 : SFM, 3D Modeling>



洪起祥

1977년 2월 서울대학교 전자공학과 (학사), 1979년 2월 한국과학기술원 (석사), 1984년 2월 한국과학기술원 (박사), 1984년 4월~1989년 6월 : 한국에너지 연구소 선임연구원, 1986년 7월~현재 : 포항공과대학 전자전기공학과 교수, 1988년 3월~1989년 8월 : Carnegie-Mellon Univ., 방문교수, <주관심 분야 : 영상처리, 컴퓨터 비전>