영상기반 기술

한국전자통신연구원 김성예 · 구본기 · 최병태

1. 서 론

컴퓨터 그래픽스의 궁극적인 목표는 사실감(Photorealism)의 추구이다. 60년대 중반부터 70년대에는 3차원 좌표로 입력된 물체의 형상을 선으로 연결하여 생성된 기하학적 모델의 표시에 중점을 둔 선구조 시대였다. 그러나 입체(Solid)에 대한 요구가 제기되면서 70년대 초부터 말까지 래스터 기술이 개발됨에 따라 영상의 은선/은면 제거 알고리즘의 발전이 가속화되었다. 또한 물체와 빛과의 관계를 고려한 쉐이딩과 그림자 생성 기법들이 영상에 사실감을 더해 주었다. 그 후, 80년대에 T. Whitted에 의해 전역조명 개념이 제안되면서 광선추적법이나 래디오시티와 같은 기법들에 의해 극도로 사실적인 영상의 생성이 가능하게 되었다. 하지만 여전히 복잡한 물체를 모델링하는 것은 많은 노력과 기술이 필요하고 정교한 조명 효과를 가지는 영상을 렌더링하는 것은 많은 시간을 요구한다. 지금까지의 전통적인 컴퓨터 그래픽스와는 달리 영상기반의 표현기법들은 입력으로 2차원 영상을 사용한다. 이러한 접근 방식은 확실히 모델링과 렌더링에 있어서 더 적은 비용으로 사실감 있는 표현을 이루어 낼 수 있다. 그러나 입력 영상에서 많은 유용한 정보를 얻을 수 있기는 하지만 단순한 영상기반 표현기법만으로는 정적인 특성의 제약을 벗어날 수 없다. 따라서 2차원 입력 영상에서 필요한 정보를 추출하고 픽셀의 재투영 등의 컴퓨터 비전 기술과 복원된 영상 정보를 이용하여 렌더링을 수행하기 위한 컴퓨터 그래픽스 기술의 접목으로 영상기반 기술이 대두되었다. 영상기반 기술은 일반적으로 영상기반렌더링, 영상기반모델링 그리고 영상기반라이팅 분야로 나뉘어 지며, 2차원 영상들을 입력으로 하여 영상의 복잡도와 관계없이 새로운 시점에서의 영상을 생성할 수 있다는 특징을 가지고 있다. 이러한 영상기반 기술이 주목을 받는 이유는 전통적인 컴퓨터 그래픽스와는 달리 정교한 모델링이 필요하지 않고 실시간으로 렌더링이 가능하기 때문이다. 한 장의 영상을 이용하여 텍스처매핑과 같은 단순한 영상기반 표현기법만을 사용한다면 전통적인 컴퓨터 그래픽스의 모델기반 렌더링 기법을 사용할 때 가능했었던 것들에 제한이 발생한다. 예를 들면, 시점의 이동에 따른 변화와 조명 정보의 변화, 물체의 추가 · 제거에 의한 변화가 불가능해진다. 하지만 영상기반 기술의 진보는 이러한 모든 제한을 가능하게 하였다. 영상기반렌더링은 새로운 시점에서의 영상 생성을, 영상기반모델링은 2차원 영상으로부터 3차원 정보를 추출하여 복잡한 물체의 모델링을 쉽게 하도록 하였다. 특히 사진이 제공하는 정보를 그대로 사용함으로써 실세계에서 표현되는 복잡한 그림자와 조명 효과를 그대로 사용할 수 있는 장점이 있다. 이에 더하여 영상기반라이팅은 영상에서 조명 정보를 추출하여 조작이 가능하게 함으로써 실세계 영상에 기하모델을 추가하는 것을 가능하게 하였다. 뿐만 아니라 2차원 영상과 전통적인 컴퓨터 그래픽스 기법인 3차원 기하학적 정보와의 혼합으로 더욱 다양한 영상기반 기술들이 연구되고 있다. 국외에서는 이미 80년대부터 텍스처매핑을 기반으로 한 영상기반 기술들에 대한 연구가 진행되었으며 이를 기반으로 한 소프트웨어가 상품화되고 있는 단계이다.

2. 기존연구

2.1 영상기반렌더링

영상기반렌더링 기법들이 추구하는 목표는 크게 새로운 시점에서의 영상 생성(Rendering)이다. 이를 위한 연구로서, Lipmann은 고정된 시점의 다양한 방향에 대하여 영상을 획득하여 DB를 구축한 후, 기존의 영상을 Panning, Tilting, Zooming함으로써 새로운 시점의 영상을 찾는 시스템을 개발하였다. Chen[5]은 전통적인 3D 렌더링 대신 한 프레임의 영상으로부터 다음 프레임까지의 영상을 보간하여 새로운 시점에서의 영상을 생성하는 기법에 대해서 소개하였다. 또한 Chen[6]은 실세계로부터 얻어진 환경맵을 이용하여 고정된 위치의 다양한 시점에서의 영상을 생성하기 위한 QuickTime VR을 소개하였다. 이에 대하여 Szeliski[28]는 가상환경 구성을 위한 파노라마 영상의 생성에 대하여 소개하였다. Seitz[24]와 Dyer는 사영기하학의 원리를 이용하여 이미지 몰핑에서 발생하는 부자연스러운 왜곡 현상을 보완하는 View Morphing을 소개하였다. Shade[26]와 Gortler는 새로운 시점에서의 영상을 생성할 때, 결과 영상에 발생하는 갭을 제거하기 위하여 Layered Depth Image를 이용하여 깊이 정보가 크게 변화하는 환경에 적용하였다. McMillan[19]은 Plenoptic Modeling을 소개하고 Plenoptic 함수의 샘플링, 재구성, 재샘플링을 기반으로 하는 영상기반렌더링 시스템을 소개하였다. 이후, 이를 기반으로 하는 여러 연구가 뒤따랐는데, Gortler[12] 등의 Lumigraph와 Levoy[17] 등의 LightField는 참조영상을 빛에 대한 4차원 함수의 2차원 슬라이스들의 집합으로서 해석함으로서 참조영상으로부터 3차원 모델 정보를 요구하지 않고 새로운 시점에서의 래디언스를 재샘플링하여 결과 영상을 생성하였다. Horry[13]는 참조 영상에 메쉬 기법을 적용하여 한 장의 영상만으로 애니메이션이 가능하게 하는 TIP을 소개하기도 하였다. 영상기반렌더링 기법 중에서 새로운 시점에서의 영상 생성을 목적으로 하면서도 보다 작게는 가상 환경의 실시간 탐색을 주요 목적으로 하는 가속화 기법들도 소개되었다. 이러한 영상기반 가속화 기법의 이슈는 실시간 렌더링을 위해서 시점과 가까운 물체는 기하모델로 먼 물체는 영상으로 대치하는 것이다. 이것은 사용되는 대치 영상의 생성 시점에 따라서 두 가지로 분류될 수 있다. 첫째, 시점에서 먼 기하모델을 와핑과 같은 기법을 이용하여 미리 영상으로 생성하고 탐색 동안에 기하모델과 영상을 스위칭 한다. Maciel[18]과 Aliaga[1]는 기하모델의 일부분을 평면에 투영시켜 얻은 샘플을 최종 영상에 텍스처 매핑하는 방식을 사용하였고 Xiong[31]은 기하모델을 대치하기 위한 영상으로 환경맵을 사용하였다. 둘째, 그래픽 하드웨어에 의해서 프레임 버퍼에 생성된 데이터를 재사용하는 방식으로 기하모델을 대치하기 위한 영상을 좀더 효율적으로 구성하여 동적인 영상 생성이 가능하도록 한다. 하지만 이를 위해서는 특별한 하드웨어뿐 아니라 렌더링 동안에 영상 데이터를 재사용하기 위한 통합된 메모리 구조가 요구된다. Torborg[29]와 Lengyel[16]는 최종 영상이 개별적으로 갱신된 영상의 레이어들로 구성되도록 하는데 새로운 하드웨어 구조를 사용하였다. Shade[25]와 Schaufler[22]는 환경의 일부분을 텍스처를 가지는 폴리곤으로 사용하고 시점이 크게 이동하는 경우에 텍스처를 갱신하는 기법을 사용하였다. Regan[20]은 시점의 변화에 대해서 상대적으로 변화가 적은 물체를 투영하여 환경맵을 구성하고 맵으로부터 시점까지의 거리에 따라서 환경맵을 갱신하였다. 이러한 방식을 포탈 텍스처라고 하며 건물 내부를 셀과 포탈로 구분하고, 포탈을 통해 보이는 부분을 기하학적 물체 대신 한 장의 영상으로 표현한다. Decoret와 Schaufler[11]는 위의 두 가지 분류를 통합하는 새로운 형태의 임포스터 기법을 제안하였다.

2.2 영상기반모델링

영상기반모델링 기술은 한 장 이상의 2차원의 사진이나 영상으로부터 3차원 구조를 추출하는 일련의 과정을 일컫는다. 2차원 영상에서 3차원 정보를 추출하는 연구는 오래 전부터 컴퓨터 비전에서 삼각측량법을 이용한 고전적인 사진측량법을 통하여 수행되어져 왔다. 사진측량법의 원리는 삼각측량법에 근거를 이룬다. 삼각측량법은 카메라의 위치를 모르고 촬영한 적어도 2장 이상의 사진에서 특징점을 추출하고 이를 카메라의 상대적 거리로 계산하여 깊이를 추출하게 된다. 이러한 기술은 연속된 영상을 이용하여 그 구조를 추출한다는 의미에서 Structure from Motion이라고도 한다. Beardsley[3]는 동영상에서 사용자가 특징점을 찍어 그 상관관계를 유추하여 3차원 기하모델을 추출하는 연구를 수행하였다. Szeliski[27]는 컴퓨터 비전 분야에서 다루던 실루엣을 360° 회전하며 얻은

다중영상에 적용하여 물체의 형상을 복원하는 연구를 수행하였다. 기존의 Structure from Motion 방법이 영상사이에 특징점을 사용자가 정해 주었던 반면에 스테레오영상을 이용한 형상복원은 영상간의 정합관계를 자동으로 계산하여 3차원 형상을 구성함으로써 효과적으로 3차원 정보를 추출하였다. Debevec과 Naimark는 고안된 스테레오 카메라를 이용하여 캐나다 국립공원의 오솔길의 3차원 기하 모델을 추출하여 가상환경을 구축하였다[14]. 결국, 스테레오를 이용한 정합은 두 특징점간의 대응관계를 결정짓는 정합오류 문제를 얼마나 최소화시키는가에 따라 그 정확도가 결정된다. Debevec[7]은 항공사진을 이용하여 버클리대의 건축물을 모델링하였다. Debevec은 사용자가 미리 정한 프리미티브를 사진위에 겹쳐 그 경계선을 맞추는 것만으로 3차원 기하 구조를 해석하였다. 이러한 방법은 건축물과 같이 직선이고 대칭적인 물체에 적용되고 있다. 원이나 구와 같은 곡선을 포함한 물체에 대한 영상기반모델링 기술은 현재 진행중이며 단순한 형태는 현재 상용화되고 있다. 이러한 프리미티브를 이용한 3차원 모델링은 현재 상용화되어 나오는 제품 기술의 근간을 이룬다.

2.3 영상기반라이팅

영상기반렌더링・모델링은 사진이나 영상을 이용하여 렌더링과 모델링이 가능하도록 하지만, 새로운 물체를 환경에 추가하거나 조명을 재계산하는 것은 불가능하다. 따라서 영상에 합성물체를 추가할 때 추가된 물체가 사실적으로 보이도록 하고 또한 추가된 물체로 인한 조명 효과의 재계산을 위하여 영상기반라이팅 기술이 제안되었다. 영상기반라이팅 기술은 빛의 반사를 표현하기 위하여 환경매핑을 사용한 Blinn[4]의 연구에서 근원을 찾을 수 있으며, 이에 기초한 반사매핑 기법은 80년대 초, Miller와 Chou에 의해서 실세계 영상에 합성 물체를 추가하는데 최초로 이용되었다. 최근, Debevec[8]은 합성 물체를 추가할 위치에서 입사 레디언스를 측정하고 전역조명 알고리즘을 이용하여 합성물체에 대한 조명 효과를 재계산하였다. 환경의 조명 자체를 변화시키는 것은 보다 복잡한 과정이 요구된다. Sato[21]는 점광원을 가지는 환경에서 촬영한 영상에서 물체의 반사특성이 어떻게 계산될 수 있는지에 대하여 소개하였고, Yizhou[30]는 임의의 광원 조건에서 반사특성을 어떻게 복구할 수 있는지에 대해서 소개하였다. Debevec[10]은 임의의 조명 조건과 임의의 시점에서 얼굴 모델을 렌더링하기 위하여 얼굴 모델의 반사특성을 얻어내는 기법을 제안하였다.

3. 본 론

3.1 영상기반렌더링 기술

◆ View Interpolation

일정한 간격을 두고 적은 시점변화에 따라 환경에 대한 영상을 생성해야 하는 경우, 거의 동일한 영상을 반복적으로 렌더링하게 되는데 이것은 환경의 복잡도가 렌더링 시간에 많은 영향을 주기 때문에 실시간 렌더링을 위해서 특별한 하드웨어의 지원이 요구되기도 한다. 이러한 문제를 해결하기 위해서 가시성 검사 결과에 따라 보이는 물체만 렌더링하기도 하였으나, 이러한 방법은 한계가 있다. View Interpolation은 새로운 영상을 생성하고자 하는 시점과 거의 비슷한 다른 시점에서의 영상을 이미지 모핑을 이용하여 중간 영상을 생성한다. 이때 미리 계산된 대응맵을 사용하여 모핑을 수행하는데 이 대응관계를 자동으로 계산하기 위해서 카메라 위치와 방향, 영상의 Range data를 이용한다. 이 기법은 3차원의 환경을 2차원 영상으로 대치할 수 있고, 렌더링 시간이 환경의 복잡도와 관계없다는 장점을 가지고 있다. 그러나 반드시 카메라 정보와 Range Data가 주어져야 하고 정적인 환경에 대해서만 적용이 가능하다는 단점을 가지고 있다.

◆ QuickTime VR

기존의 가상환경 탐색을 위한 기법들은 실시간 탐색을 가능하게 하기 위해서 특별한 하드웨어의 지원을 받아야만했다. Chen은 가상환경 탐색을 위한 영상기반 기법을 제안하였고 360° 실린더 형태의 파노라마 영상을 사용하는 QuickTime VR을 소개하였다. QuickTime VR은 영상기반 기술을 이용한 상업적 제품으로서 연속적인 카메라 팬잉과 줌잉 다른 영역으로의 점핑이 가능한 파노라마 무비 플레이어와 프레임 인덱싱을 이용하여 물체의 회전이 가능하도록 하는 물체 무비 플레이어를 제공한다. 또한 적절한 LOD기능도 제공한다. 그러나 여전히 정적인 환경에만 적용이 가능하고 다른 지역으로의 이동은 HotSpot 개념의 점핑만 가능하다.

또한 실린더 형태의 환경맵을 이용하기 때문에 제한된 환경의 탐색만 가능하다.

◆ Panoramic Image Mosaic

Szeliski는 파노라마 영상을 생성하기 위한 새로운 기법을 제안하였다. 이 기법은 입력 영상에 대한 제한이 없기 때문에 hand- held 카메라에 의해 촬영된 영상도 자유롭게 사용할 수 있다. Szeliski는 기존의 8개의 파라미터를 사용하는 Perspective Transform 대신, 3개의 파라미터만 가지는 회전 매트릭스를 사용함으로써 더 빠르면서도 쉽고 직관적으로 파노라마 영상을 생성할 수 있게 하였다. 그리고 생성된 파노라마 영상을 이용하여 효과적으로 환경맵을 구성하기 위한 방법도 제시하였다.

◆ View Morphing

View Interpolation에서 사용된 이미지 모핑기법은 영상내의 물체의 위치와 시점의 차이로 인해서 종종 원하지 않는 왜곡이 발생하는데 이것은 수동으로 수정이 어렵다. Seitz는 사영기하학의 원리를 이용하여 이러한 왜곡을 발생시키지 않는 확장된 이미지 모핑기법을 소개하고 이를 View Morphing이라 하였다.

이 기법은 두 장의 영상을 모핑하기 전에 와핑을 거친 뒤, 모핑을 수행한 영상에 다시 와핑을 수행하는데 3차원 정보를 필요로 하지 않기 때문에 사진이나 그림과 같은 영상에도 적용이 가능하다. 그림 1과 같이 서로 다른 시점에서 얻어진 두 장의 영상을 이용한 View Morphing 영상은 물리적으로 가상 카메라를 이용한 효과를 가지게 된다.

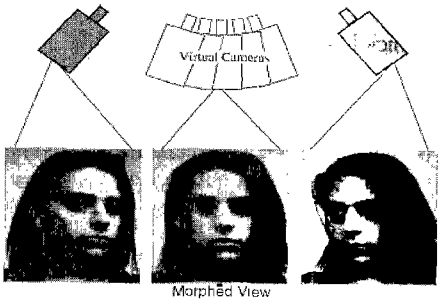


그림 1 View Morphing

◆ Plenoptic Modeling

McMillan과 Bishop은 영상기반렌더링 패러다임에 대한 문제를 명확히 하기 위하여 Plenoptic 함수를 사용하였다. Plenoptic 함수는 한 시점에서 보이는 환경에 대한 모든 정보를 7개의 파라미터로 나타낸다.

$$p = P(\theta, \phi, \lambda, V_x, V_y, V_z, t)$$

방향: θ, Φ / Wavelength: λ
관찰자위치: Vx,Vy,Vz / Time: t

McMillan은 Plenoptic 함수에 의한 이산적 샘플들의 집합이 주어질 때, Plenoptic 함수의 연속적인 표현을 위해 샘플링, 재구성, 재샘플링으로 구성되는 시스템을 소개하였다. 이 시스템은 샘플링 단계에서 카메라에 의해서 획득된 영상을 가지고 실린더 맵을 구성하고 재구성 단계에서 두 참조 영상과 두 참조영상 간의 Disparity Map을 이용하여 Epipolar geometry를 결정한다. 마지막으로 재샘플링 단계에서 새로운 시점에서의 영상을 생성한다.

◆ Light Field Rendering & Lumigraph

새로운 시점에서의 영상을 생성하기 위해서 제안되었던 많은 영상기반렌더링 기법들은 참조 영상들간의 대응관계나 깊이정보를 이용하여 참조영상을 와핑 하는 방식을 이용하였다. Levoy는 참조영상을 빛에 대한 4차원 함수의 2차원 슬라이스들의 집합으로 해석하고 이를 Light Field라 하였다. 이 기법은 깊이정보나 특징점 매칭과 같은 3차원 정보를 이용하지 않고 재샘플링에 의해서 새로운 시점에서의 영상을 생성할 수 있다. 그러나 Light Field를 구성하기 위해서는 많은 메모리가 요구되기 때문에 압축 기법이 사용되어야 하며 고정된 조명과 정적 환경에만 적용이 가능하다. 그림 2는 4차원(u,v,s,t) 함수로 구성된 Light Field를 나타낸다.

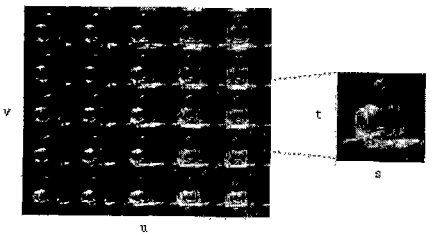
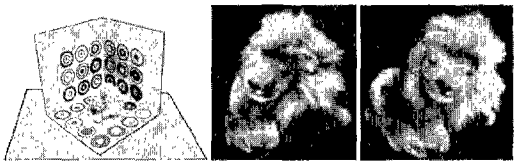


그림 2 LightField

Gortler는 Lumigraph라고 하는 4차원 함수를 구성하였다. 이것은 Light Field와 유사하게 한 시

점에서 보이는 모든 방향에 대한 빛의 흐름을 표현한다. Lumigraph를 구성하기 위해서 그림 3(a)와 같은 환경을 이용하여 물체를 캡처하는데 샘플링 정도에 따라 Light Field와 마찬가지로 많은 메모리를 요구하기 때문에 Gortler는 200~1000:1정도의 압축 기법을 사용하도록 하였다. 그림 3(b), (c)는 Lumigraph를 이용하여 서로 다른 시점에서의 영상을 렌더링한 결과이다.



(a) (b) (c)

그림 3 Lumigraph

◆ Tour into the Picture : TIP

2차원 영상을 이용하여 환경의 탐색을 가능하게 하기 위해서 여러 장의 참조 영상을 필요로 하는 QuickTime VR이나 Plenoptic modeling 등의 기법이 제안되었고, 또한 영상기반모델링 기술을 이용함으로써 직접 3차원 모델을 구성하는 기법도 제안되었다. 이와 다르게 Horry는 단지 한 장의 2차원 영상을 이용하여 탐색이 가능하도록 하는 기법을 제안하였다. 그림 4와 같이 2차원 영상을 분할하고 전경의 물체에 대해서는 폴리곤 모델로 배경은 5개의 사각 영역으로 구성한다. TIP에 의한 기법은 완벽한 3차원 환경의 구성없이 단지 한 장의 영상만을 이용하여 애니메이션이 가능하도록 하였지만 전경과 배경의 구분과 소실점의 선택이 수작업으로 이루어져야 하는 한계를 지니고 있다.

그림 4 Tour Into the Picture 구성

◆ Layered Depth Images

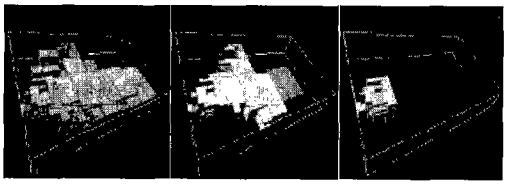
일반적인 텍스처매핑 기법은 텍스처맵과 최종 영상간의 차이로 인해 Pixelization이 발생하거나 하드웨어의 지원을 필요로 하는 한계를 가지고 있다. 이를 보완하기 위하여 스프라이트를 사용하거나 깊이 정보를 추가할 수 있다. 그러나 가시성의 변화 때문에 발생하는 갭이나 픽셀이 퍼져(spreading)보이는 문제는 여전히 발생하게 된다. Backward 매핑이나 여러 장의 참조 영상을 사용하는 등의 해결 방안이 있지만 여전히 문제가 있다. Shade는 이를 위해 두 가지 방법을 제시하였다. 첫째, 평평한 면에 대해서는 깊이를 가지는 스프라이트를 사용하고 둘째, 좀 더 일반적인 환경에 대해서는 LDI(Layered Depth Image)을 사용하는 것이다. LDI란, 한 카메라 위치에서 보여지는 환경에 대한 정보를 나타내며 하나의 시선에 대해서 깊이가 다른 다수의 픽셀 정보를 가진다. LDI를 구성하기 위해서는 여러 장의 깊이 정보를 가지는 영상을 이용하거나 변경된 광선추적법을 사용하고 또는 입력 영상으로부터 직접 Voxel Coloring을 사용하여 구성한다. LDI는 Z-버퍼를 요구하지 않고 LDI를 표현하기 위한 용량은 깊이 복잡도에 따라서 선형으로 증가한다. 결과적으로, LDI을 사용함으로써 PC상에서도 초당 많은 프레임을 렌더링할 수 있다.

3.2 영상기반가속화 기술

◆ 포탈 텍스처(Portal Texture)

포탈 텍스처는 포탈을 통해 보이는 기하학적 모델을 영상으로 대치하여 렌더링 시간을 감소시킴으로써 실시간 탐색이 가능하도록 한다. Aliaga[2]가 소개한 포탈 텍스처 알고리즘은 건물 내부를 벽으로 구성된 공간(room)인 셀과, 셀을 통해 보이는 부분(문, 창문)인 포탈로 구분한다. 기존의 방법은 셀에서 포탈을 통해 보이는 모든 부분이 3차원 기하모델로 구성되어 있어 많은 렌더링 시간이 필요한 반면, 셀에서 포탈을 통해 보이는 기하모델을 2차원 영상으로 대치함으로써 실시간 탐색이 가능하다. 또한, 텍스처와 기하모델의 스위칭 시점에 대한 관계를 설정하여 사용자가 포탈에 가까이 다가가면 영상을 3차원 기하모델로 다시 대치하여 정확한 탐색이 가능하게 하였다.

그림 5(a)는 뷰-프러스텀 컬링에 의한 건물 내부 자료가 줄어든 모습이다. (b)는 셀과 포탈의 관계를 이용하여 현재 셀에 연결된 셀들의 집합들을 이용한 포탈 컬링을 수행한 결과이다. (c)는 포탈 텍스처를 이용하여 현재 셀에 있는 포탈을 포탈 텍스처로 대치한 결과이다. 그림 5에 의해 렌더링할 건물 내부 자료가 현저히 줄어든 것을 알 수 있다.



(a) (b) (c)

그림 5 포탈에 의한 기하모델 제거

◆ 임포스터(Imposter)

Decoret는 3차원 기하모델에 바운딩 박스를 씌운 뒤 물체가 시점에 가까워지면 기하모델로 표현하고 물체가 시점에서 멀어지면 바운딩 박스의 중점과 시점을 연결한 방향으로 물체를 평면에 투영하여 임포스터를 생성하였다. 임포스터는 바운딩 박스로 둘러싸인 3차원 물체를 평면 위에 투영한 영상으로써 물체의 기하학적 모델을 대신하기 위하여 사용된다. 이러한 임포스터를 이용한 렌더링 방법은 가상도시 모델을 위해 사용될 수 있다. Schaufler는 그림 6과 같이 깊이 영상을 이용하여 한 물체를 여러 개의 임포스터로 구성한 Layered Imposter를 소개하기도 하였다.

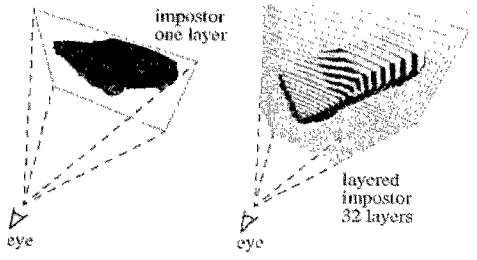


그림 6 Multi-Layered Imposter

3.3 영상기반모델링 기술

◆ Shape from Silhouettes

컴퓨터 그래픽스 및 컴퓨터 비전 분야에서 다중 영상으로부터 물체의 실루엣으로부터 윤곽선을 추출하여 물체의 형상을 복원하는 연구가 진행 되어 왔다. 다중의 영상을 이진화하여 얻은 실루엣 영상을 8진 트리를 사용하여 원추형으로 투영하고 서로 교차하는 지역을 획득하여 3차원 형상을 추출한다. 실루엣을 이용한 3차원 형상 측정은 구현하기 쉽고 빠른 실행속도를 보이는 반면에 제한된 해상도와 이진화에 따른 노이즈에 민감하다. 실루엣을 이용한 3차원 형상복원은 간단한 알고리즘이지만 3차원 구조에 대한 중요한 실마리를 제공하고 있다. Szeliski는 그림 7과 같이 물체의 360° 실루엣 영상을 획득하여 3차원 형상을 측정하였다.

그림 7 Shape from Silhouettes

◆ Traditional Photogrammetry

2차원 영상에서 3차원 정보를 추출하는 연구는 컴퓨터 비전에서 사진측량법의 한 분야로서 연구되어 왔다. 사진측량법의 원리는 삼각측량법을 기본으로 하며 삼각측량법은 2장 이상의 사진에서 공통으로 나타나는 특징점을 추출하여 상대적인 대응관계를 유추하면 카메라의 위치를 찾을 수 있고 이를 이용하여 3차원 정보를 추출한다. 이러한 기술은 항공사진에서 대응점을 표시함으로써 항공 관측으로부터 지형도를 추출하는 방법을 제공하기도 한다. Beardsley는 비디오로부터 특징점을 추적하여 3차원 위치와 카메라 위치를 찾는 Structure from Motion에 관한 연구를 수행하기도 하였다.

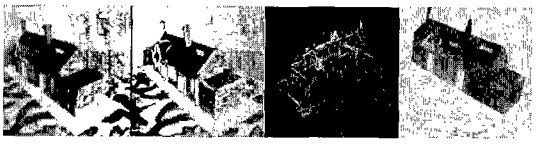


그림 8 Structure from motion

◆ Shape from Stereo

기존에 사용되던 사진측량법에 의한 3차원 구조 복원은 두 장의 영상간에 특징점들 사이의 관계를 사용자가 정해 주어야하는 단점이 있었다. 만약, 두 영상간의 특징점간의 대응관계가 자동으로 계산된다면 훨씬 효과적일 것이다. 어떤 거리에서 촬영된 두 영상이 있다면, 스테레오 정합을 이용하여 기준 영상에 대한 대응 영상간의 정합을 수행하여 그 문제를 해결 할 수 있다. 두 영상간에 특징점들간의 대응관계를 알면 삼각측량법을 사용하여 3차원 깊이 정보를 얻을 수 있다. 스테레오 정합 알고리즘은 인간의 시각구조를 잘 반영하고 있는 반면에 기

준 영상의 특징점에 대응하는 다른 영상에서의 유사한 특징점이 다수 존재함으로 인해 생기는 정합 오류 문제를 가지기 때문에 정확한 3차원 정보를 결정하기 어려운 단점이 있다. Debevec과 Naimark는 두 대의 16mm 카메라를 설치한 장치를 고안하고 1m 간격으로 스테레오 영상을 획득하여 캐나다 국립공원의 가상환경을 구축하는데 사용하였다. Kanade[15]는 50대의 카메라를 사용하여 스테레오 정합 알고리즘의 정합오류 문제를 최소화하였다.

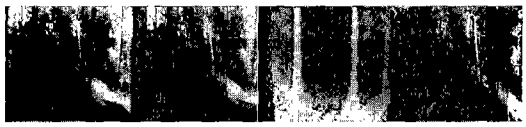


그림 9 Shape from Stereo(Immersion94)

◆ Photogrammetric Modeling with Facade

Debevec은 항공사진을 이용하여 건축물의 모델링과 렌더링을 수행하였다. 그 결과 그림 10과 같은 Facade는 최소한의 사진으로부터 실물과 동일한 3차원 건물을 구성하였다. 사용자들은 미리 정의된 프리미티브를 사진 위에서 이동하여 그 경계선을 맞추는 것만으로 3차원 기하 구조가 구성된다. 프리미티브를 이용한 3차원 모델링의 단순함은 건축물이 가지는 대칭성을 최대한 이용함으로써 기존의 방법들이 가지던 난해한 정합문제들을 해결하는 역할을 하였다. 또한, 건축물의 벽면과 같은 아주 정밀한 부분의 모델링을 위해서는 Model-based Stereo 기법을 사용하였다.

그림 10 건축물의 모델링 및 렌더링(Facade)

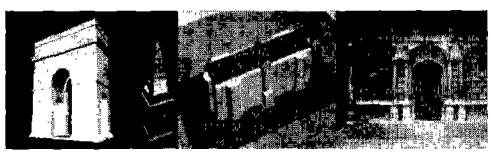


그림 11 Model-base Stereo

그림 12의 Campanile Movie라는 애니메이션이 SIGGRAPH97에서 상영되었는데 이것은 연(kite)을 이용한 항공 촬영을 통하여 얻은 영상으로 탑을 구성하고 주변의 건물들을 모델링 하여 가상환경을 구축하였다. 또한, 시점에 관계없이 텍스처를 맵핑하는 기술[9]을 제안하여 생동감 있고 실감 있는 영상기반모델링 및 렌더링 기술의 진수를 보여주었다.

그림 12 Campanile Movie

3.4 영상기반라이팅 기술

◆ Reflection Mapping

그래픽스 연구의 많은 부분은 70, 80년대 후반의 환경매핑[4]이나 반사매핑을 근간으로 하고 있으며, 실세계 영상을 반사맵으로서 사용하기 시작한 것은 80년대 초반, Gene Miller와 Michael Chou에 의해서였다. 그림 13은 Miller에 의해 실세계 배경 영상에 합성 물체를 추가하고 환경의 조명 조건을 측정하여 합성 물체에 반사매핑을 수행한 영상이다.

그림 13 Reflection-mapped blobby dog

◆ Rendering Synthetic Objects into Real Scene

Debevec은 환경의 래디언스를 측정하고 전역 조명 알고리즘을 사용하여 실세계 영상에 합성 물체를 자연스럽게 추가하기 위한 방법을 소개하였다. Light Probe를 사용하여 환경의 래디언스를 측정하고 실세계 영상에 추가되는 합성물체의 조명 효과를 계산하기 위해서 세 가지 조합의 환경을 고려하였다. 첫째, 물체에 의해서 영향을 받지 않는 것으로 가정하는 원거리영상과 둘째, 측정된 반사

모델에 영향을 받아 그림자를 표현하거나 합성될 새로운 물체로부터의 반사광에 의해 영향을 받는 근거리 영상, 마지막으로 합성물체 이다. 그림 14는 각각 원거리 영상(a), 합성물체와 근거리 영상의 추가(b), 최종 렌더링 결과(c)를 보여준다.

(a) (b) (c)

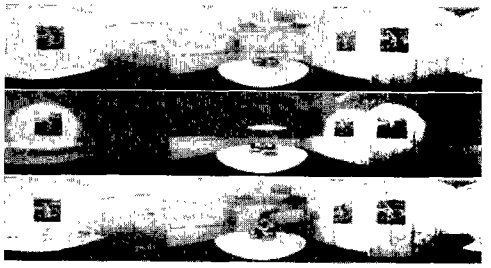
그림 14 실 영상에 합성물체의 추가

◆ Reflectance Modeling

정적인 조명환경에서 물체를 합성하는 것과 달리 환경의 조명 자체를 변화시키는 것은 보다 복잡한 계산이 요구된다. Sato는 움직이는 광원을 가지는 환경에서 촬영된 다수의 영상을 이용하여 Lambertian 모델과 Torrance-Sparrow 모델의 조합으로 설명되는 반사 모델을 추출하고, 이를 임의의 조명 조건에서 합성 물체를 렌더링하는데 사용하였다.

◆ Recovering Reflectance

Yizhou는 사진 영상으로부터 실세계 환경의 반사 모델을 복구하는 기법을 소개하고, Inverse Global Illumination이라 하였다. 환경에 대한 기하모델과 High Dynamic Range 영상을 입력으로 사용하여 기하모델을 폴리곤 메쉬로 분할한 뒤, 영상기반렌더링 기법을 사용하여 각 패치의 래디언스를 계산하였다. 그림 15는 소스 영상, (a)를 이용하여 다른 조명 조건에서의 영상, (b)와 합성 물체를 추가한 경우의 영상, (c)를 파노라마 영상으로 구성한 것이다.



(a)

(b)

(c)

그림 15 Inverse Global Illumination

◆ Acquiring Reflectance of a Human Face

Debevec은 Light Stage를 사용하여 여러 시점에서 촬영한 얼굴 영상을 가지고 임의의 조명 조건과 시점에서 인간의 얼굴을 렌더링하기 위한 기법을 소개하였다. 입력 영상을 이용하여 생성된 반사함수를 사용하면 소스 영상과 같은 시점에서 다른 조명 조건의 영상을 렌더링할 수 있다. 또한, 변경된 시점에서의 렌더링을 위하여 Skin Reflectance 모델을 사용하였다. 그림 16은 다른 조명 조건에서의 얼굴을 렌더링한 영상이다.

그림 16 다른 조명조건에서의 얼굴 렌더링

4. 응용 및 사례

영상기반 기술의 진보는 건축물과 애니메이션, 영화 및 광고와 같은 다양한 분야의 가시화 문제를 해결하는 새로운 방법의 요구와 함께 많은 연구자와 개발자들로 하여금 2차원 영상이 가지고 있는 정보를 이용하여 3차원 모델과 새로운 2차원 영상을 만드는데 기여하고 있다.

4.1 영상기반렌더링, 라이팅 기술의 응용

◆ Virtual Camera

Virtual Camera는 시간의 시작과 정지, 종료 등의 효과를 생성하고 정지 모션과 일반 모션의 합성에도 사용된다. 특히, 그림 17과 같이 영화 매트릭스를 통하여 잘 알려져 있다.

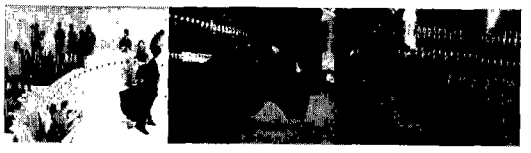


그림 17 Virtual Camera의 활용(The Matrix)

◆ Reflection mapping in Films

SIGGRAPH85 Film Show에 반사 매핑을 이용

한 애니메이션인 Interface [그림 18(a)]가 처음 선을 보인 이후로 많은 영화에서 영상기반라이팅 기술이 사용되었다. 그림 18(b)는 반사매핑을 사용한 최초의 영화인 Flight of the Navigator이고 (c)는 TerminatorⅡ의 한 장면이다.

(a) (b) (c)

그림 18 영화에서의 영상기반라이팅 기술

4.2 영상기반모델링 기술의 응용

영상기반모델링 기술을 적용한 대부분의 상용제품은 카메라의 위치와 시점을 정하기 위하여 사진측량 원리에 근거하고 있다. 일단 위치가 정해지면 적어도 2개 이상의 다른 시점에서 촬영된 사진간의 삼각측량에 의해서 3차원 모델을 만들어진다. 또한, 이러한 기술은 시각적으로 정확히 구별되는 점, 선, 평면 등을 이용하여 모델을 구축하게 된다. 현재 출시 되어있는 상용제품들은 대부분 윈도우 기반으로 구동되고 범용 그래픽 파일로의 저장기능이 제공되고 있다.

◆ PhotoModeler(EOS Systems)[32]

한 장 이상의 입력영상에서 특징이 되는 포인트를 사용자가 입력하고 이를 연결하는 선을 구성하거나 면을 생성한 뒤 특징점들 간의 정합관계를 계산하여 3차원 모델을 생성한다. 타제품과 달리 카메라 캘리브레이션 모듈을 포함하고 있어 사용자의 필요에 따라 더욱 정확한 3차원 모델을 생성할 수 있다.

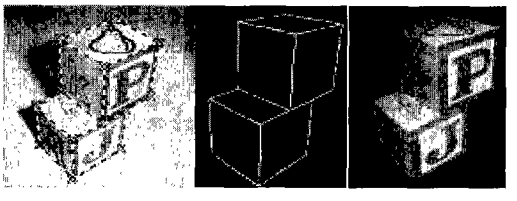


그림 19 PhotoModeler 활용

◆ Image Modeler(Realviz)[33]

다른 제품과 달리 Image Modeler는 한 장 이상의 입력영상에서 사용자가 몇 점을 입력하면 정합 알고리즘에 의해서 전체 사진 영상의 특징점의 정합점을 자동으로 찾아내는 모듈을 포함하고 있다. 또한, 건축물과 같이 대칭적이거나 직선을 포함하는 모델이외에도 인물과 같은 비정형체에 대한 3차원 모델 생성도 가능하다. 최근에는 구성된 3차원 모델에 애니메이션을 추가하여 사용자와 상호작용이 가능하도록 하고 있다.

그림 20 Image Modeler의 활용

◆ Canoma(MetaCreations)[34]

Canoma는 다른 상용제품이 사용자가 특징점을 직접 찍어 주어 그 상관관계를 추출하는 방법 대신에 Debevec이 제안한 방법과 같이 미리 정해진 3차원 프리미티브-면, 박스, 웨지 등-를 사용하여 사진 위에 적절히 배치함으로써 3차원 모델을 생성하도록 하였다.

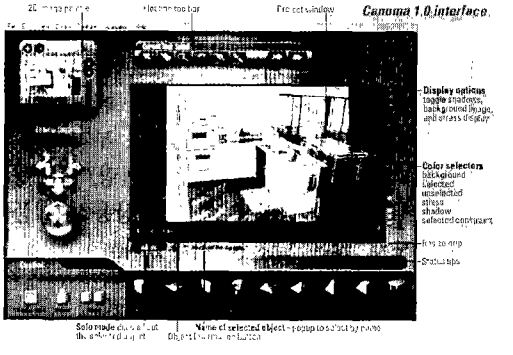


그림 21 Canoma 활용

◆ Gen-Trix(R) Studio(Gentech)[35]

Gen-Trix(R) Studio는 정확한 3차원 모델을 구성하기 위하여 자체 제작된 보정 시스템에 대상 물체 사이의 보정을 수행한 뒤 다른 영상간의 정합을 수행하여 3차원 모델을 추출한다. VR 환경의 3차원 데

이터 제작을 위한 전체 시스템이 설계되어 있다.

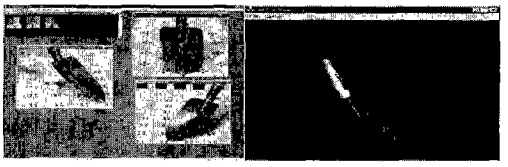


그림 22 Gen-Trix(R) Studio 활용

◆ ArcView 3D Producer(Evans & Sutherland)[36]

ArcView 3D Producer는 CAD 데이터로부터 건물을 구성하는 Building Modeler와 사진으로부터 3차원 모델을 구성하는 Picture Modeler, 자연경관이나 나무 등을 3차원으로 구성하는 CutOut Modeler의 세 가지 제품군으로 구성된다. 이들 제품은 자연물이나 인조물로 특화된 제품을 제작, 판매 서비스하는 것이 특징이다.

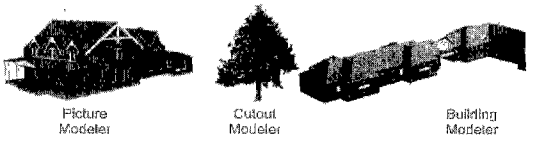


그림 23 ArcView 3D Producer

5. 맺음말

지금까지 영상기반 기술과 그 응용에 대하여 살펴보았다. 전통적인 그래픽스 기법의 렌더링과 모델링의 한계를 극복하기 위한 요구로 부각되기 시작한 영상기반 기술들은 렌더링, 모델링, 라이팅의 각 분야에서 다양한 응용뿐 아니라 상용화 제품을 선보일 정도로 많은 발전을 이루었다.

참고문헌

[1] D. G. Aliaga, Visualization of complex models using dynamic texture based simplification, IEEE Visualization '96, 1996.

[2] D. G. Aliaga, Anselmo A. Lastra, Architectural Walkthroughs Using Portal Textures, IEEE Visualization '97, pp. 355-362, 1997.

[3] Beardsley, Paul, Phil Torr, Andrew Zisserman, 3D Model Acquisition from Extended Image Sequences, European Conference on Computer Vision, pp. 683-695, 1996.

[4] J. F. Blinn, M. E. Newell, Texture and reflection in computer generated images, Communications of the ACM 9(10), pp. 542-547, 1976.

[5] S. E. Chen, L. Williams, View Interpolation for Image Synthesis, SIGGRAPH93, pp. 279-288, 1993.

[6] S. E. Chen, QuickTime VR-An Image-Based Approach to Virtual Environment Navigation, SIGGRAPH95, pp. 29-38, 1995.

[7] P. Debevec, Modeling and Rendering Architecture from Photographs, Ph.D. Thesis, UC Berkeley, December, 1996.

[8] P. Debevec, Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image based Graphics with Global Illumination and High Dynamic Range Photography, SIGGRAPH98, pp. 189-198, 1998.

[9] P. Debevec, Yizhou Yu, G. D. Borshukov, Efficient View-Dependent Image Based Rendering with Projective Texture Mapping, In 9th Eurographics Workshop on Rendering, pp. 105-116, June 1998.

[10] P. Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, Mark Sagar, Acquiring the Reflectance Field of a Human Face, SIGGRAPH2000, 2000.

[11] X. Decoret, G. Schaufler, F. X. Sillion, J. Dorsey, Multi-Layered Impostors for Accelerated Rendering, Eurographics '99, pp. C61-C72, 1999.

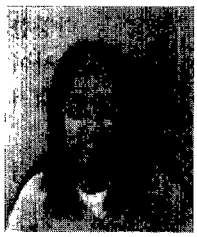
[12] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, Michael F. Cohen, The Lumigraph, SIGGRAPH96, pp. 43-54, 1996.

[13] Y. Horry, K. I. Anjyo, K. Arai, Tour into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image, SIGGRAPH97, pp. 225-232, 1997.

[14] The Immersion '94 project at Interval Research Corp.

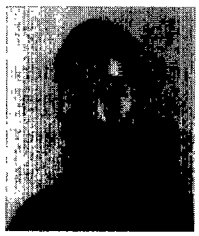
[15] T. Kanade et. al., Constructing Virtual

Worlds Using Dense Stereo, ICCV' 98, Bombay, India, pp. 3-10, January 1998.
[16] Jed Lengyel, John Snyder, Rendering with coherent layers, SIGGRAPH97, pp. 233-242, 1997.
[17] M. Levoy, P. Hanrahan, Light Field Rendering, SIGGRAPH96, pp. 31-42, 1996.
[18] M. Cohen, P. Shirley, Visual Navigation of Large Environments Using Textured Clusters, Symposium on Interactive 3D Graphics, pp. 95-102, 1995.
[19] L. McMillan, G. Bishop. Plenoptic Modeling: An Image-Based Rendering System, SIGGRAPH95, pp. 39-46, 1995.
[20] M. Regan, R. Pose, Priority rendering with a virtual reality address recalculation pipeline, SIGGRAPH94, pp. 155-162, 1994.
[21] Sato, Yoichi, Mark D. Wheeler, Katsushi Ikeuchi, Object Shape and Reflectance Modeling from Observation, SIGGRAPH97, pp. 379-387, 1997.
[22] G. Schaufler, Wolfgan Sturzlinger, A three-dimensional image cache for virtual reality, Computer Graphics Forum, 15(3), 1996.
[23] G. Schaufler, Per-Object Image Warping with Layered Imposters, In 9th Eurographics Workshop on Rendering 98, pp. 145-156, 1998.
[24] S. M. Seitz, C. R. Dyer, View Morphing, SIGGRAPH96, pp. 21-30, 1996.
[25] J. Shade, D. Lischinski, D. Salesin, T. DeRose, J. Snyder, Hierarchical image caching for accelerated walkthroughs of complex environments, SIGGRAPH 96, pp. 75-82, 1996.
[26] J. W. Shade, S. J. Gortler, L. He, R. Szeliski, Layered Depth Images, SIGGRAPH98, 1998.
[27] R. Szeliski, Rapid Octree Construction from Image Sequences, CVGIP: Image Understanding, 58(1), pp. 23-32, July 1993.
[28] R. Szeliski, Heung-Yeung Shum, Creating Full View Panoramic Image Mosaics and Environment Maps, SIGGRAPH97, pp. 251-258, 1997.
[29] J. Torborg, Jim Kajiya, Talisman: Commodity Real-time 3D graphics for the PC, SIGGRAPH96, pp. 353-364, 1996.
[30] Yizhou Yu, P. Debevec, Jitendra Malik, Tim Hawkins, Inverse Global Illumination: Recovering Reflectance Models of Real Scenes From Photographs, SIGGRAPH99, pp. 215-224, 1999.
[31] Xiong, Rebecca, CityScape-A Virtual Navigational System for Large Environments, MIT Master Thesis, 1996.
[32] http://www.photomodeler.com
[33] http://www.realviz.com
[34] http://www.metacreations.com
[35] http://www.gen.co.jp/
[36] http://www.es.com



김 성 예

1998 중앙대학교 컴퓨터공학과(학사)
2000 중앙대학교 컴퓨터공학과(석사)
2000~현재 ETRI VR 센터 영상기반렌더링팀
관심분야:컴퓨터그래픽스, 전역조명 모델링, 래디오시티, NPR, 영상기반렌더링, 가상현실
E-mail:inside@etri.re.kr



구 본 기

1995 경북대학교 전자공학과(학사)
1997 경북대학교 전자공학과(석사)
1996~현재 ETRI VR 센터 영상기반렌더링팀
관심분야:3차원 이미징, 영상기반모델링 및 렌더링, 컴퓨터비전, 영상처리, 가상현실
E-mail:bkkoo@etri.re.kr

최 병 태

1986 경북대학교 전자공학과(학사)
1991 KAIST 전산학과(석사)
1991~현재 ETRI VR센터 영상기반렌더링팀 팀장
관심분야:컴퓨터그래픽스, 컴퓨터비전, 네트웍 가상현실
E-mail:btchoi@etri.re.kr