

영상 기반 모델링 및 렌더링 Image-based Modeling and Rendering

한 정 현

성균관대학교 전기전자 및 컴퓨터공학부

요약 영상기반 모델링 및 렌더링은 1990년대 초반 이후 집중적으로 연구되기 시작한 분야로, 영상 자체를 입력으로 하여 출력 영상을 생성하여 자연스럽게 photorealism을 달성할 수 있고, scene의 복잡도에 무관한 렌더링을 가능케 한다. 본 논문은 파노라마 렌더링, light field 렌더링, LDI 렌더링을 중심으로 지난 10년간에 걸친 영상기반 모델링 및 렌더링의 연구 성과를 개괄한다.

1. 소 개

3차원 컴퓨터 그래픽스의 궁극적인 목표 중 하나는 사진과 같은 사실적인 영상을 생성하는 데 있다. 이러한 photorealism을 달성하기 위해서는 복잡한 3차원 기하 모델링과 긴 시간을 소모하는 렌더링 과정이 필수적이었다. 이러한 전통적인 기하 기반 기법의 대안으로 영상 기반 모델링과 렌더링 기법(image-based modeling and rendering: 이하 IBMR)이 1990년대 초반 이후 집중적으로 연구되기 시작하였다. IBMR은 영상을 그 자체를 입력받아 새로운 시점에서의 출력 영상을 생성하므로, 자연스럽게 photorealism을 달성할 수 있고, scene의 복잡도에 무관한 렌더링을 가능케 한다.

그림 1은 영상 처리, 컴퓨터 그래픽스, 컴퓨터 비전의 상호 관계를 도시한다. 그림 1의 도식에 따르면, 영상을 입력으로 받아들여 출력 영상을 생성하는 IBMR은 영상 처리의 한 분야로 간주될 수

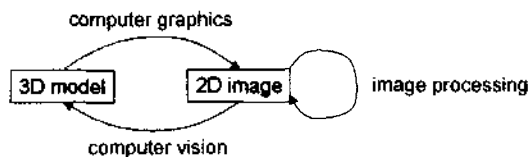


그림 1. Image processing, computer graphics and computer vision (from^[1])

있다. 하지만 일반적인 영상 처리와는 달리 IBMR 기법은 대개 입력 영상으로부터 일단 3차원 모델을 생성하고 이를 이용해 출력 영상을 만들어낸다¹. 따라서, IBMR은 컴퓨터 비전 기법에 기반한 그래픽스 렌더링 과정으로 묘사할 수도 있다. 이와 같이 IBMR에서는 영상 처리, 컴퓨터 그래픽스, 컴퓨터 비전 세 분야가 유기적으로 결합되어 있고, IBMR 연구를 위해서는 위 세 분야의 지식이 필수적이다.

IBMR은 이름 그대로 영상기반 모델링과 영상기반 렌더링의 두 단계로 구분된다. 영상기반 모델링의 결과물은 그림 2에 도시된 바와 같이 다양한 방식으로 표현될 수 있고, 이에 따라 영상기반 렌더링 과정 역시 달라지게 된다. 이들 중 1995년

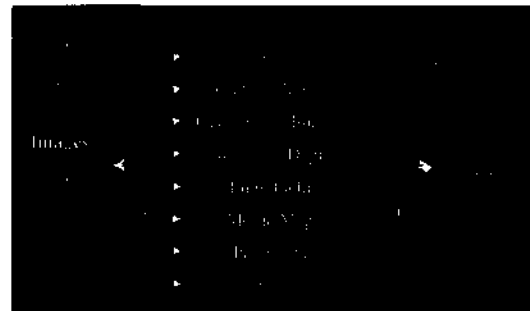


그림 2. The spectrum of IBMR (from^[2])

¹대개의 경우, 이러한 3차원 모델은 전통적인 3차원 기하 모델과는 다른 방식으로 표현된다.

QuickTime VR^[3]에서 채택된 'panorama' 기법은 모델링 과정이 쉽고 저가의 컴퓨터에서도 쉽게 렌더링되는 커다란 장점을 가진 IBMR 기법이다. 한편 1996년에 Levoy와 Hanrahan^[4], Gorler 등^[5]에 의해 'light field' 렌더링 기법이 소개되었고, 뒤이어 1998년 Shade 등^[6]에 의해 'images+depth' 모델을 이용한 LDI(Layered Depth Images) 기법이 소개되었는데, 이 두 기법에 대해서는 최근까지 활발히 연구 성과가 보고되고 있다. 본 논문은 파노라마 렌더링, light field 렌더링, LDI에 집중해서 IBMR의 기술 동향을 개괄한다.

2. 파노라마 렌더링

2.1. QuickTime VR

QuickTime VR에서는 원통형의 파노라마 영상을 이용하여 2차원 panning 및 zooming을 지원한다. 파노라마 영상은 그림 3과 같이 중첩된 일반 영상을 이어 붙여 생성할 수도 있고², 파노라마 카메라를 이용하여 한번에 얻어낼 수도 있다³. 이와 같은 방법으로 얻어진 파노라마 영상에 왜곡을 적용하여 2차원 panning을 수행하며, 매끄러운 zooming을 위해서는 다중 해상도 영상을 이용하기도 한다.

이러한 파노라마 렌더링 기법은 IBMR의 가장 간단한 예인데, 특별한 하드웨어 사양의 필요 없이



그림 3. Image stitching (from^[3])

²일반 영상들로부터 파노라마 영상을 생성하는 기법은 이미 다수 상품화되어 있는데, 일례로 다음 site를 방문하면 trial version을 다운받을 수 있다.

▷ http://www.mgisoft.com/webtools/photovista/pv_main.asp

³파노라마 카메라의 예는 <http://www.panoramtech.com> 등에서 찾아볼 수 있다.

⁴이의 어원 및 이론적 기초에 대해서는 참고문헌^[7]을 참조하라.

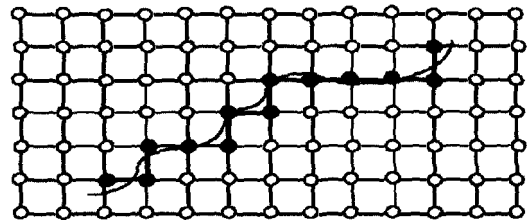


그림 4. Navigation with Panorama Grid (from^[3])

실시간 렌더링을 보장하는 아주 견고한 기법으로, 현재 인터넷에서 3차원 scene을 표현하는 사실상 표준의 하나로 사용되고 있다. 하지만 QuickTime VR의 파노라마 렌더링에서는 기본적으로 시점의 위치는 고정되고 뷰잉 방향의 변화만이 가능하다. 따라서, 시점의 이동 즉 내비게이션 구현을 위해 QuickTime VR에서는 원칙적으로 그림 4와 같이 2차원 격자를 구성하고 각 격자점에 파노라마 영상을 배치한다. 내비게이션 시에는 현재의 시점과 가장 가까운 격자점의 파노라마 영상을 이용하여 출력 영상을 생성한다. 따라서, 그림 4에 도시된 바와 같이 시점 이동 경로에 가장 가까운 파노라마 영상들을 차례차례 사용해야 하므로 출력 영상에서의 jumping/popping 현상을 피할 수 없다.

2.2. Beyond QuickTime VR

Plenoptic⁴ modeling 기법^[8]은 QuickTime VR과 마찬가지로 원통형 파노라마 영상을 사용하지만, 2차원 격자 대신 임의의 위치에서 얻어진 보다 적은 수의 파노라마 영상을 사용하여 시점 이동을 지원한다. 스테레오 영상으로부터 픽셀의 깊이 정보를 계산하기 위하여 컴퓨터 비전 분야에서 널리 사용되는 기법으로 epipolar geometry가 있는데, plenoptic modeling에서는 이를 파노라마 영상에 적용하여 두 영상간 disparity, 즉 광류(optical flow)를 계산하고, 이를 이용하여 새로운 시점에서의 영상을 생성한다. 하지만, 이러한 광류 기반 기법의 커다란 문제점은, 입력 영상에서는 보이지 않던 부분이 시

점 이동에 따라 출력 영상에서는 나타나야 할 때 이를 해결할 근본적인 방안이 존재하지 않는다는 데 있다. 이런 경우 결국 출력 영상에 홀(hole)이 발생하게 된다.

한편, 1999년에 발표된 동심 모자이크(concentric mosaics) 기법^[9]은 일반적인 파노라마와는 사뭇 다른 모델링 기법을 사용하지만, 파노라마 모델링만큼의 편이성을 제공하며 동시에 렌더링 과정에서 시점이동에 따른 변위 효과(parallax)를 얻을 수 있다는 측면에서 주목을 받았다. 하지만, 이 기법은 렌더링 과정에서 이른바 '수직 왜곡' 현상 등 치명적인 결함을 안고 있다.

3. Light field rendering

1996년에 Levoy와 Hanrahan^[4], 그리고 Gortler 등^[5]에 의해 동시에 발표된 light field 렌더링 기법은 하나의 오브젝트를 렌더링하는 이른바 object viewer에 적합한 기법으로 평가된다(Gortler 등^[5]은 이 기법을 Lumigraph라 부른다). 하나의 오브젝트를 둘러싸는 직육면체를 가정하면, 그림 5의 st -평면과 uv -평면은 이 직육면체를 구성하는 두 평면이 된다. st -평면상의 한 점 (s, t) 를 눈의 위치라 하고 uv -평면을 영상면으로 간주해 ray tracing을 수행해 보자. 눈에 해당하는 (s, t) 점으로부터 uv -평면의 (u, v) 점을 연결하는 벡터는 하나의 ray가 되는데, $L(s, t, u, v)$ 는 (s, t) 와 (u, v) 를 잇는 ray가 처음으로 오브젝트와 만나는 지점에서의 컬러 값을 나타낸다. 그림 6은 (s, t) -평면을 5×5 격자로 나누고 ray tracing 한 결과를 모은 예를 보여

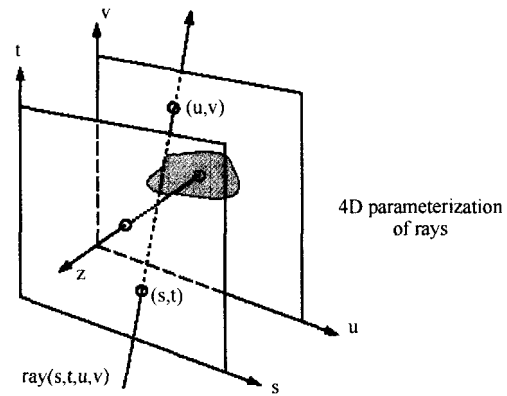


그림 5. Parameterization of light field (from^[5])

준다.

이렇게 생성된 light field를 사용하여 새로운 시점에서 렌더링을 수행하는 과정이 그림 7에 도시되어 있다. 새로운 시점으로부터 영상면의 한 픽셀을 통과하는 ray가 st - 및 uv -평면과 교차하는 점을 계산하여 $L(s, t, u, v)$ 를 가져오면 이 값이 바로 영상면의 해당 픽셀 값이 되는 것이다. 물론, 새로운 시점에서의 ray가 st - 및 uv -평면의 격자점과 정확히 일치하는 경우는 거의 없으므로, 보간 과정이 필요하고 이를 위해 Lumigraph 기법에서는 대상 오브젝트의 대략적인 기하 정보를 사용하였다.

Light field 기법에 대해서는 최근까지 활발히 연구 성과가 보고되고 있는데, Wood 등^[10]은 오브젝트 표면의 모든 점에서 방출되는 ray의 모든 가능한 방향에 컬러 값을 할당한다는 개념의 surface light field 기법을, Isaksen 등^[11]은 Lumigraph와는 달리 오브젝트의 대략적인 기하 정보를 필요로

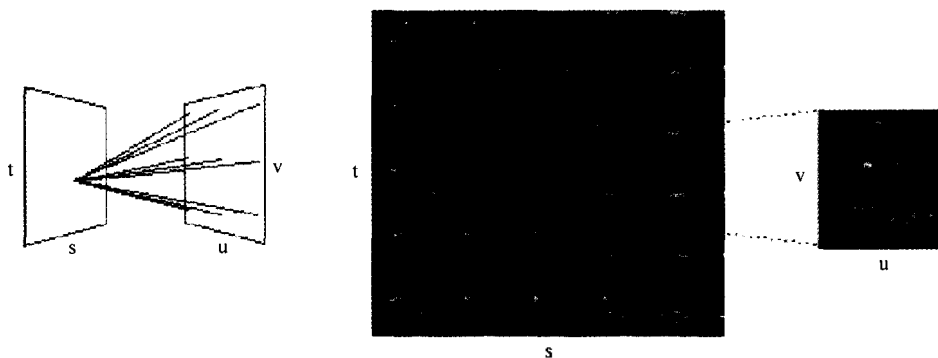


그림 6. Visualization of light field (from^[4])

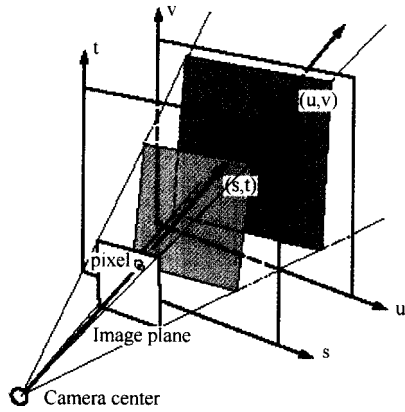


그림 7. Rendering using light field (from⁵⁾)

하지 않는 기법을 제안했다.

Light field 기법의 가장 커다란 문제점은 데이터 크기가 방대하다는 것으로, 출력 화질 개선을 위해 고밀도 샘플링을 수행하면 쉽게 기가 바이트 단위에 도달하게 된다. 따라서 데이터 압축이 필수적으로 요구되는데, 고성능의 압축을 수행한다해도 인터넷 전송 등에는 부적절한 크기가 된다. 예를 들어, 영상 기반 렌더링의 주요 응용 분야 중 하나인 Web commerce에 light field 모델을 사용하는 것은 현재의 압축 기술로는 매우 회의적이다.

4. Layered depth images

앞서 2.2절에서 언급한 바와 같이, 픽셀의 깊이 정보를 이용한 광류(optical flow) 기반 기법은 홀발생의 문제를 가지고 있다. Shade 등⁶⁾은 깊이 정보를 가진 영상을 복수 개 이용하여 이 문제점을 해결하고자 Layered depth images(LDI)를 제안했다.

LDI 기법의 입력을 구성하는 각각의 영상은 그 영상을 얻은 카메라 좌표계 정보와 모든 픽셀에 대한 깊이 정보를 필요로 한다. 이렇게 모든 픽셀이 깊이 정보를 가지고 있으므로 각각의 영상을 '깊이 영상(depth image)'이라 한다. 이러한 '깊이 영상'을 3차원 외평¹¹⁾을 통해 하나의 카메라 좌표계로 재구성하면, 이 카메라로부터의 각각의 ray에는 여러 개의 픽셀이 놓이게 된다. 예를 들면, 그림 8에 도시된 바와 같이, 카메라 C_2 에서 얻은 픽셀 b와 c는 카메라 C_1 좌표계로 3차원 외평되면서 동일한

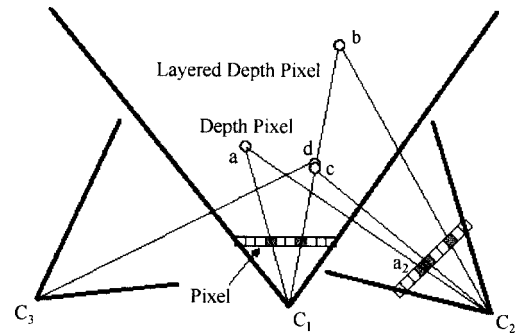


그림 8. Layered depth images (from⁶⁾)

ray 상에 놓이게 된다. 한편 카메라 C_3 에서 얻은 픽셀 d는 c와 동일한 위치로 옮겨지는데 이 경우는 c와 d의 컬러 값을 평균해서 새로운 픽셀을 정의한다. 이러한 방식으로 3차원 외평된 픽셀들은 그 깊이 정보에 따라 계층적으로 저장되게 된다. 이렇게 구성된 영상 모델을 '계층적 깊이 영상(layered depth image: LDI)'이라 부른다.

이렇게 구성된 LDI를 렌더링하기 위해서는 의료 영상 가시화 등의 분야에서 많이 사용되는 스플래팅 기법¹¹⁾을 사용한다. 그런데 3차원 복셀을 영상면으로 스플래팅할 때 스플래팅 커널의 크기를 결

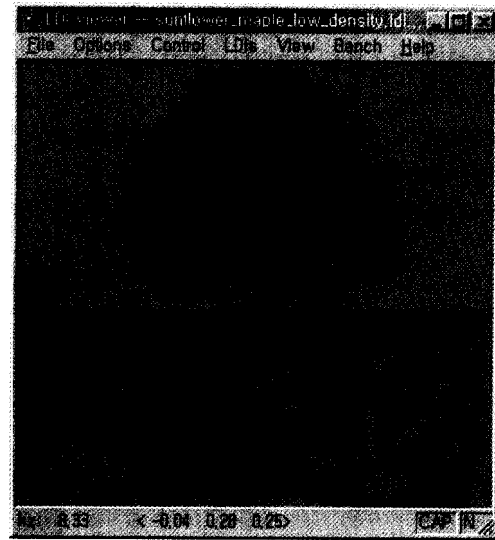


그림 9. Hole problems in LDI rendering⁵

⁵<http://www.cs.washington.edu/research/projects/grail2/www/projects/ldi/>

정하는 것은 까다로운 문제로, 커널의 크기가 필요 이상 작을 경우 출력 영상에 홀이 발생할 수 있으며, 반대의 경우는 출력 영상이 흐려지는 현상이 발생한다. LDI 렌더링 역시 이런 문제를 드러낸다. 그림 9는 LDI를 이용하여 입력 영상과는 다른 새로운 시점에서 렌더링을 한 결과를 보여주는데, 심각한 홀 발생 문제를 보여준다.

한편, LDI와 같이 '깊이 픽셀'을 이용한 IBMR 기법은 그 뿌리를 Chen의 뷰 보간 기법^[13]에 두고 있는데, 이 계열의 IBMR 기법으로는 2.2절의 plenoptic modeling 기법^[8] 이외에도, 슬릿(slit) 카메라를 이용한 모델링을 제안한 MCOP 기법^[15], LDI 데이터를 계층적으로 저장하는 LDI 트리 기법^[16] 등이 있고, 최근에는 렌더링의 새로운 기법으로 이른 바 '포인트 샘플 렌더링'을 지향하는 연구 결과^[16-18]와 접목되는 양상을 보이고 있다.

5. 결 론

본 논문은 1990년대 컴퓨터 그래픽스 분야에서 가장 활발히 연구되어 왔다고 말할 수 있는 영상 기반 모델링 및 렌더링(IBMR)을 파노라마 렌더링, light field 렌더링, LDI에 집중해서 간단히 고찰하였다. 짧은 지면을 이용한 개괄이었지만, 영상 처리, 컴퓨터 그래픽스, 컴퓨터 비전 세 분야의 유기적인 결합체인 IBMR의 현 수준을 파악할 수 있는 기회가 되었으면 한다.

Photorealism을 실시간에 구현한다는 목표로 진행되어온 IBMR 연구는 QuickTime VR의 성공 이후 현재까지 제법 긴 모색기를 거치는 중인 듯하다. 즉, 다수의 연구자가 인정하는 탁월한 IBMR 기법은 아직 제시되지 않았다. 그 커다란 이유 중 하나는, IBMR을 구성하는 한 축인 컴퓨터 비전 분야에서 스테레오 영상 기술 등이 아직 견고한 해법을 찾아내지 못하고 있는데 있다. 한편, 컴퓨터 비전 등에서의 요소 기술이 뒷받침된다면, 현재의 렌더링 파이프라인에 필적하는 범용적인 IBMR 기법이 출현할 수 있을까? 이에 대해서는 1997년 SIGGRAPH에서의 IBR 패널, 1998년 스탠포드 대학에서의 IBMR 워크숍⁶에서 각각 Malik과 Szeliski

가 내린 진단이 현재에도 여전히 우리에게 시사점을 던져주고 있는 듯 하다. "IBMR 범용 기법은 적어도 단기간에는 나올 수 없다. 현 시점에서 가장 실용적인 해법은 IBMR이 필요한 각 분야에서 그 영역에 고유한 기법을 발견하는 것이다."

참고문헌

1. Pavlidis, T. *Algorithms for Graphics and Image Processing*, Computer Science Press, 1982.
2. Debevec, P. E. "Image-based Modeling, Rendering, and Lighting," *SIGGRAPH2000 Course Notes*.
3. Chen, S. E.. "QuickTime VR - An Image-based Approach to Virtual Environment Navigation," *SIGGRAPH95*, 1995.
4. Levoy, M. and Hanrahan, P. "Light Field Rendering," *SIGGRAPH96*, pp. 31-42, 1996.
5. Gortler, S. J., Grzeszczuk, R., Szeliski, R. and Cohen, M. F., "The Lumigraph," *SIGGRAPH96*, pp. 43-54, 1996.
6. Shade, J. W., Gortler, S. J., He, L., and Szeliski, R., "Layered Depth Image," *SIGGRAPH98*, pp. 231-242, 1998.
7. Adelson, E. H. and Bergen, J. R., "The Plenoptic Function and the Elements of Early Vision," *Computational Models of Visual Processing*, Chapter 1, Ed. by Landy, M. and Movshon, J. A., The MIT press, Cambridge, Mass. 1991.
8. McMillan, L. and Bishop, G., "Plenoptic Modeling: An Image-based Rendering System," *SIGGRAPH95*, pp. 39-46, 1995.
9. Shum, H. and He, L., "Rendering with Concentric Mosaics," *SIGGRAPH99*, pp. 299-306, 1999.
10. Wood, D., Azuma, D., Aldinger, K., Curless, B., Duchamp, T., Salesin, D., and Stuetzle, W., "Surface Light Fields for 3D Photography," *SIGGRAPH2000*, pp. 287-296, 2000.
11. Isaksen, A., McMillan, L., and Gortler, S., "Dynamically Reparameterized Light Fields," *SIGGRAPH2000*, pp. 297-306, 2000.
12. Mark, W., McMillan, L., and Bishop, G., "Post-rendering 3D Warping," *Proc. of 1997 Symposium on Interactive 3D Graphics*, pp. 7-16, 1997.
13. Westover, L.. "Footprint Evaluation for Volume Rendering," *SIGGRAPH90*, pp. 367-376, 1990.
14. Chen, S. and Williams, L.. "View Interpolation

⁶<http://www-graphics.stanford.edu/workshops/ibr98>

- for Image Synthesis," *SIGGRAPH93*, pp. 279-288, 1993.
15. Rademacher, P. and Bishop, G., "Multiple-Center-of-Projection Images," *SIGGRAPH98*, pp. 199-206, 1998.
 16. Grossman, J. and Dally, W., "Point Sample Rendering," *Proc. of 1998 Symposium on Interactive 3D Graphics*, pp. 181-192, 1998.
 17. Lischinski, D. and Rappoport, A., "Image-based Rendering for Non-diffuse Synthetic Scenes," *Proc. of 1998 Symposium on Interactive 3D Graphics*, pp. 301-314, 1998.
 18. Pfister, H., Zwicker, M., Barr, J. and Gross, M., "Surfels: Surface Elements as Rendering Primitives," *SIGGRAPH2000*, pp. 335-342, 2000.