2005년도 대한전자공학회 컴퓨터소사이어티 하계학술대회 논문집 제28권 제1호

깊이정보를 이용한 Light Field 예측 부호화 기법

이 은 경, 윤 승 욱, 호 요 성
광주과학기술원 정보통신공학과
전화 : 062-970-2263 / 핸드폰 : 016-610-0643

Predictive Coding of Light Field using Depth Information

Eun-Kyung Lee, Seung-Uk Yoon, and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
E-mail : {eklee78, suyoon, hoyo}@gist.ac.kr

Abstract

Recently, variance image-based rendering (IBR) techniques are used actively in many applications. A light field rendering is one of them that can render three-dimensional scenes using two dimensional images. Light field rendering (LFR) provides a simple and efficient way to generate arbitrary new views of the scene using the high correlation among adjacent pixels. This technique characterizes the flow of light through the unobstructed space in a static scene. Since an enormous amount of data required in a light field poses a key challenge in rendering, we need to compress light field data. In this paper, we propose a predictive coding scheme to compress light field data.

I. 서론

최근 멀티미디어와 가상현실에 대한 관심이 증가하면서 현실에 존재하는 장면을 좀더 사실감있게 표현하고자 하는 노력이 많은 분야에서 진행되고 있다. 지금까지의 가상공간을 표현하는 기술은 대부분 물체나 장면을 그래픽 툴을 이용해 기하학적으로 정교하게 모델링하고 이를 렌더링하였다. 그러나 장면을 생성할 때 많은 시간과 노력이 필요할 뿐만 아니라 생성된 결과도 인공적인 느낌이 강하여 가상현실과 같이 몰입감을 요구하는 분야에 적용할 때 어려움이 많았다.

영상기반 렌더링은 필요한 정보를 입력된 그림이나 사진을 통해 얻어 장면을 구성하면 입력 영상과 같은 수준의 사실적인 장면을 연출해 낼 수 있다. 뿐만 아니라, 출력 장면의 생성 속도가 장면 내의 기하학적 복잡도에 상관없이 일정하다. 따라서 객체들을 포함한 영상을 장면 관찰자가 바라보는 시점에 따라 실시간으로 생성하고 렌더링할 수 있다.

영상을 기반으로 현실감있는 장면을 연출하는 기술 중에 대표적인 것이 Light Field 렌더링이다. Light Field 렌더링은 다른 위치와 각도에서 여러 영상들을 획득하고 이를 이용하여 현재 시점에 맞는 영상을 렌더링한다. 렌더링 화질은 사용할 수 있는 2차원 영상의 수에 따라 결정된다. 현실감있는 렌더링 결과를 얻기 위해서는 수천 장의 영상이 필요하기 때문에 실시간 렌더링을 위해서는 압축 부호화 과정이 필요하다.

본 논문에서는 Light Field의 부호화 효율을 향상시키기 위해 기존의 비디오 부호화 방식인 벡터 양자화와 엔트로피 부호화 방법을 Light Field에 적합하게 재구성하였다. 또한 카메라 정보와 깊이정보를 이용해 3차원 워핑(warping)을 수행하고, 카메라 위치에 적합한 보간방법을 사용하여 좀더 효과적으로 주변 영상을 예측하는 방법을 제안한다.

II. 영상기반 렌더링 (Image-Based Rendering)

영상기반 렌더링은 지금까지의 기하기반 렌더링 기법의 단점을 보완하면서 좀더 현실적인 렌더링 결과를 얻을 수 있다는 장점 때문에 많은 분야에서 활발히 연구되고 있다. 영상기반 렌더링 기술은 기하학 정보의 사용 유무에 따라 크게 기하정보가 없는 렌더링, 암시적 기하정보가 있는 렌더링, 그리고 명확한 기하정보가 있는 렌더링으로 분류할 수 있다.

기하정보가 없는 렌더링 방법은 빛의 흐름을 매개변수화한 플렌옵틱 함수(plenoptic function)의 특성을 이용하여 새로운 시점의 장면을 재현하는 것으로, 빛 정보와 텍스처 정보만을 이용해 장면을 렌더링한다[1,2]. 기하정보가 없는 렌더링 방법의 경우에는 렌더링에 필요한 영상 데이터가 많이 필요한 것이 특징이다. 루미그래프(lumigraph)와 Light Field 렌더링, 영상 모자이크 등이 이 범주에 속한다.

암시적 기하정보가 있는 렌더링 방법은 새로운 시점의 장면을 재현하기 위해 입력 영상들의 위치적 관련성을 이용하는 렌더링 방법으로 기하정보를 직접적으로 렌더링에 이용하지 않는다. 3차원 위치정보를 투영행렬을 계산해서, 이 정보를 새로운 장면 재현을 위해 사용하여 원하는 시점의 장면을 렌더링한다. 시점 모핑 (view morphing) 기법이 그 예이다.

마지막으로, 명확한 기하정보가 있는 렌더링 방법은 직접적으로 깊이정보와 같은 3차원 정보를 사용하여 렌더링한다. 이는 깊이정보 기반의 3차원 워핑 (warping)을 수행하여 새로운 시점에서의 영상을 생성하는 방법으로 전통적인 텍스처 매핑 방법과 계층적 깊이영상 (layed depth image, LDI) 방법이 이 범주에 속한다[3, 4].

III. Light Field 렌더링

모든 빛의 흐름은 3차원 공간 좌표 (x, y, z)와 2차원 방향정보 (θ, ϕ)를 이용해 5차원 플렌옵틱함수 (plenoptic function)로 표현할 수 있다. 이 플렌옵틱함수는 폐색 (occlusion) 영역이 존재하지 않는 조건하에서 4차원 함수로 표현될 수 있다. 그림 1에서 나타낸 것처럼, 카메라 평면 uv와 영상 평면 st가 평행하게 존재하고, 그 두 평면 사이를 광선이 통과한다. Light Field는 이 두 평면과 광선간의 교차점을 이용하여 Light Field 함수 L(u,v,s,t)로 매개변수화 할 수 있다. 단, 모든 빛의 흐름은 두 평면에 의해 구성되는 평행사변형 안에 존재해야 한다[3].

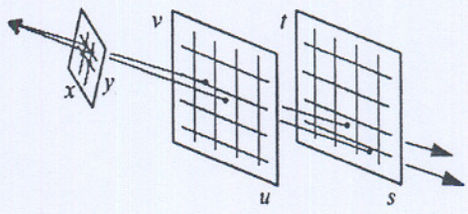


그림 1. Light Field의 4차원 매개변수화

그림 2는 영상평면 st에서의 시점 영상과 그 시점에 해당하는 재표본화 (resampling) 영상를 보여주고 있다. 이는 시점의 변화에 따라 각기 다른 영상을 생성한다. Light Field는 이렇게 카메라 평면과 영상 평면간의 인덱싱과정과 리샘플링 과정을 통해 현재 관찰자가 원하는 시점의 장면을 생성하고 렌더링할 수 있다.

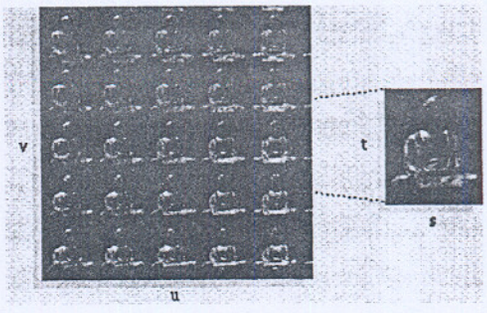


그림 2. 영상평면 st과 카메라 평면 uv의 인덱싱 구조

Light Field를 부호화하기 위해 두 단계의 부호화 알고리즘을 사용한다. 이는 벡터 양자화와 엔트로피 부호화 방법을 4차원 구조인 Light Field에 적합하게 재구성한 방법이다. 그림 3에 보인 것처럼, 먼저 Light Field의 벡터정보를 손실 부호화기법인 벡터 양자화 방법을 통해 일정 비트 내의 새로운 벡터로 생성한다. 생성한 벡터를 코드워드(codeword)라 하고, 코드워드의 집합을 코드북(codebook)이라 한다. 영상 안에 존재하는 모든 Light Field 벡터는 코드북 안에 존재하는 코드워드로 부호화한다. 그리고 생성한 코드워드를 엔트로피 부호화한다[5, 6]. 이는 시점 변화에 따라 장면을 빠르게 렌더링하기 위한 방법으로, 모든 Light Field 부호화의 기본 구조이다.

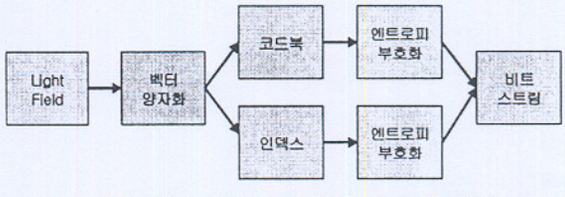


그림 3. Light Field 렌더링 부호화 방법

Ⅳ. 기존의 Light Field 부호화 방법

현실감있는 장면 연출을 위한 Light Field는 방대한 양의 데이터를 포함한다. 적게는 수백에서 많게는 수천장에 해당하는 영상을 효율적으로 부호화하기 위해 크게 두 가지 부호화 알고리즘을 사용한다.

첫 번째 접근 방법은 기존의 비디오 부호화 방식을 Light Field에 적합하게 재구성한 비디오 부호화기반 Light Field 부호화와 주변 영상간의 화소 차이값을 이용한 변이기반 Light Field 부호화를 포함한다. 먼저 비디오 부호화기반 Light Field 부호화 방법은 기존의 비디오 부호화 방식을 Light Field 구조에 맞게 확장한 방법으로, 그림 4에 보인 것처럼, 영상을 크게 I-영상과 P-영상으로 부호화한다. 여기서 I-영상과 P-영상의 개념은 MPEG 비디오 부호화에서와 같은 개념으로 I-영상은 전체 영상을 인트라 부호화하고, P-영상은 I-영상을 이용해 예측 부호화한다.

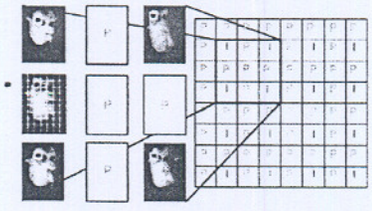


그림 4. 비디오 부호화기반 Light Field 부호화

Light Field 부호화의 또 다른 접근 방법은 변이기반 부호화 방법으로, 그림 5에 보인 것처럼, 주변 영상과의 변이차를 계산하고, 그 값을 하나의 영상으로 생성한다. 변이값은 현재 영상의 화소값이 주변영상의 어느 위치인지를 찾은 후, 두 화소의 위치를 뺀 값이다. 그렇기 때문에 변이값을 이용해 주변영상의 화소 위치를 쉽게 예측할 수 있다. 그림 6은 변이영상을 이용해 주변 영상을 예측하고 원 영상을 다시 복원하는 과정을 보여준다[6].

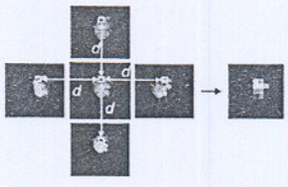


그림 5. 주변 영상을 이용한 변이영상 생성

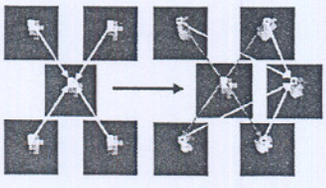


그림 6. 변이 영상을 이용한 원본영상 생성

Ⅴ. Light Field 예측 부호화 시스템

본 논문에서는 2차원 배열의 카메라로부터 영상을 획득하고, 이를 예측 부호화하는 방법을 제안한다. 기존의 Light Field 부호화 방법을 기반으로, 주변 영상의 효율적인 예측을 위해 각 시점의 카메라 정보와 깊이정보를 이용한다.

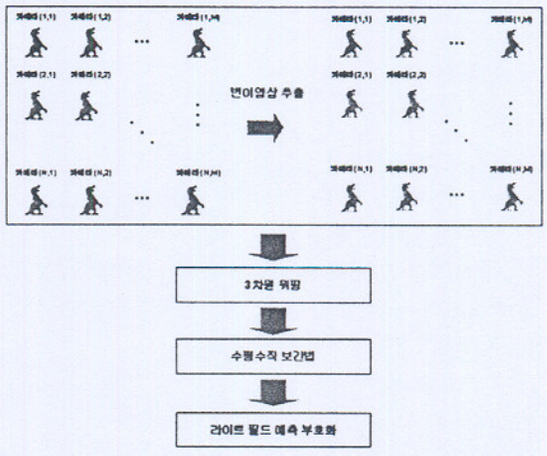


그림 7. Light Field 예측 부호화 시스템

먼저 획득한 모든 영상의 변이 (disparity) 영상을 생성하는데, 이는 그 영상의 깊이 값을 표현한다. 생성한 깊이 정보와 주어진 카메라 정보를 이용해 3차원 워핑 (warping) 과정을 수행한다. 3차원 워핑은 현재 영상과 깊이영상, 그리고 카메라 정보를 이용해 다른 카메라 위치에서 획득한 영상을 생성하고자 할 때 사용하는 기법이다. 따라서 이를 통해 주변 영상을 비교적 정확하게 예측할 수 있다.

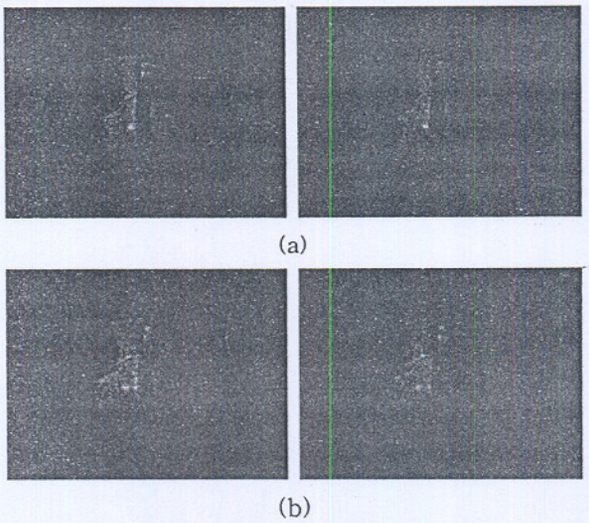
그림 8. 카메라 위치에 따른 홀

그러나, 그림 8에 보인 것처럼, 워핑을 수행하면 영상에 많은 홀(hole)이 생기게 된다. 이 홀을 효과적으로 채우기 위해 주로 수평수직 보간기법을 사용한다. 즉, 주변 값을 이용하여 수평 방향으로 홀을 채운 후, 수직 방향으로 나머지 홀을 채우는 방법을 사용한다. 그리고 현재 영상과 다음 영상의 카메라 위치 정보를 고려하여 폐색 영역을 채운다.

그림 8은 예측할 영상의 카메라가 현재 카메라 위치의 왼쪽에 위치한 경우로 폐색 영역은 객체의 오른쪽에 발생한다. 따라서 폐색 영역 영상의 오른쪽 화소값을 이용하여 폐색 영역을 채운다. 반대의 경우에는 왼쪽 화소값을 이용한다. 마지막으로 생성한 영상을 이용해 예측 부호화한다.

VI. 실험 결과 및 분석

본 논문에서는 8×8 크기를 가지는 카메라 평면 uv와 640×680 크기를 가지는 st영상에 대하여 실험을 수행했으며, PSNR을 사용하여 실험 결과를 비교하였다. 그림 9는 워핑할 원영상과 워핑한 결과 영상으로, 워핑을 수행한 후, 수평수직 보간을 수행한 결과 영상이 원영상과 거의 일치하는 것을 볼 수 있다. 그림 10은 기존의 부호화 방법과 제안한 예측 부호화 방법의 PSNR값을 표시한 것이다.



(a)

(b)

그림 9. 워핑 결과 영상

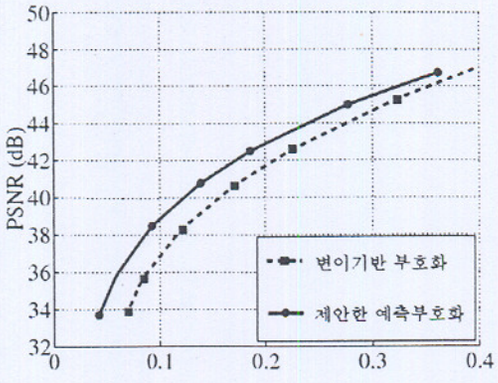


그림 10. 원본영상과 예측영상의 PSNR 비교

VII. 결론

본 논문에서는 깊이 정보와 카메라 정보를 이용한 3차원 워핑 기법을 통해 좀더 정확하게 주변 영상예측하는 Light Field 영상의 예측 부호화 알고리즘을 제안하였다. 3차워 워핑을 통해 생기는 홀을 채우기 위해 수평수직 보간기법을 사용하였고, 폐색영역의 효율적인 보간을 위해 카메라의 위치정보를 사용하였다. 제안된 예측 부호화 방법을 사용하여 부호화해야 하는 데이터 양을 현저히 줄임으로써, 방대한 양의 Light Field를 효과적으로 예측 부호화할 수 있었다. 향후에는 좀더 정확한 영상 예측을 위해 정확한 변이영상을 추출하는 방법에 대해 연구할 예정이다.

감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부 대학IT연구센터(ITRC), 교육인적자원부 두뇌한국21(BK21) 정보기술사업, 디지털 콘텐츠협동연구센터(DCRC)의 지원에 의한 것입니다.

참고문헌

[1] L. McMillan, "An image-based approach to three-dimensional Computer graphics," Ph.D. Dissertation. UNC Computer Science Technical Report TR97-013, April 1997.

[2] P. Debevec, C. Taylor, and J. Malik, "Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach," Proceedings of SIGGRAPH, pp. 11 - 20, August 1996

[3] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The lumigraph," Proceedings of SIGGRAPH, pp. 43 - 54., 1996

[4] J. Shade, S.J. Gortler, and R. Szeliski, "Layered depth image," Proceedings of SIGGRAPH, pp. 291-298, July 1998.

[5] M. Levoy, and P. Hanrahan, "Light field rendering," Proceedings of SIGGRAPH, pp. 31 - 42, 1996

[6] M. Magnor and B. Girod, "Data compression for light field rendering," IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, pp. 338 - 343, April 2000.