

RGB 텍스처와 깊이 데이터를 이용한 가상 다시점 영상의 생성 및 그래픽스 합성

황원영 권준섭 김만배 최창열
강원대학교 컴퓨터정보통신공학과
e-mail:godz165@kangwon.ac.kr

Virtual Multiview Image Composition based on RGB Texture and Depth Data

Won Young Hwang, Jun Sup Kwon, Man Bae Kim, Chang Yeol Choi
Dept. of Computer Science and Engineering
Kangwon National University

요 약

2D 및 입체 영상 콘텐츠의 공급이 많아지면서 실감 콘텐츠에 대한 관심이 더욱 높아지고 있다. 실감 콘텐츠는 다시점 카메라로부터 획득한 다시점 비디오, 깊이 카메라에서 얻은 RGB 영상과 컴퓨터 그래픽스와 같은 synthetic data를 합성하여 보다 실감나게 제작된다. 다시점 카메라를 이용하면 다시점 비디오를 쉽게 획득할 수 있으나 제작비용이 많이 들고, 깊이 카메라를 사용하면 시스템 구성은 상대적으로 용이하나 시점 영상이 하나라는 단점이 있다. 본 논문에서는 깊이 카메라에서 얻은 RGB 텍스처 데이터와 깊이 데이터로부터 가상 다시점 영상을 생성하고, 생성된 영상에 컴퓨터 그래픽스를 합성하는 방법을 제안한다. 제안한 방법은 다시점 카메라 시스템을 사용하지 않고도 시점에 따른 가상 시점 화상을 용이하게 제작하여 그래픽 객체를 합성한다. 합성된 영상은 다시점 3D 모니터나 입체 모니터를 이용하여 3차원으로 실감나게 시청할 수 있다.

1. 서론

그동안 2-D 및 스테레오 영상에 대한 많은 연구와 기술 개발이 있어왔으며, 최근 들어서는 실감 미디어에 대한 소비자들의 요구가 점점 증가되고 있는 추세이다. 나아가 실감 방송은 다양한 시점(viewpoint)으로 표현되기 때문에 관찰자의 위치에 따라 서로 다른 3차원 영상을 보여줄 수 있어야 한다. 이에 다시점 3차원 영상의 획득이 필요하며, 그 중의 하나가 다시점 비디오(multiview video)이다 [1-2]. 다시점 비디오는 8~12개의 카메라를 평형식 또는 폭주식으로 배열하여 획득한 영상으로서, 넓은 시야를 제공하기 때문에 실감형 또는 몰입적 미디어를 제공할 수 있다. 그러나 다시점 비디오 방법은 여러 대의 카메라와 각 카메라에 대한 제어가 필요하여 시스템의 복잡성이 높아진다는 단점이 있다.

한편, 시스템이 복잡해지는 문제를 해결하고 비교적 간단하게 다시점 3D 모니터에 적합한 영상을 효율적으로 얻기 위해서는 깊이 정보를 이용하여 다시점 영상을 생성한다[3]. 즉, 깊이 영상으로부터

다시점 모니터의 시점 개수만큼의 시점 영상들을 생성하여 공급함으로써, 다시점 카메라 시스템의 복잡성을 극복할 수 있다.

본 논문에서는 이와 같이 다시점 비디오를 생성할 때 수반되는 이슈들에 대한 해결 방안으로서 깊이 정보를 이용한 다시점 영상의 생성 및 시점별 그래픽스 합성 기법을 제안한다. 제안된 기법은 일반적으로 깊이 카메라(depth camera)로부터 획득한 RGB 텍스처 데이터와 8비트의 깊이 정보를 이용하여 생성한 여러 개의 가상 시점(virtual view) 영상들, 그리고 3D 그래픽스 영상을 합성하여 더욱 실감나는 그래픽 영상을 제작할 수 있게 한다. 즉, 한대의 깊이 카메라를 이용하여 여러 개의 가상 시점 영상과 그래픽스 객체를 합성하여 보다 자연스러운 영상을 생성할 수 있다.

본 논문의 구성은 다음과 같다. 서론에 이어, 2절에서는 깊이 카메라를 이용한 다시점 가상 영상의 생성 방법을 살펴본다. 3절은 그래픽스 객체의 합성

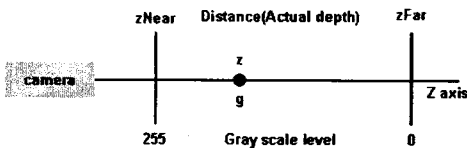
기법을 기술하고, 4절에서는 실험 결과를 보인다. 마지막 5절에서 결론을 내린다.

2. 깊이 정보를 이용한 다시점 영상의 생성

다시점 영상을 생성하는 데는 깊이 카메라를 이용한 깊이 정보의 획득과 표현 방법이 필요하다. 본 절에서는 다시점 영상의 생성 방법을 간단하게 요약한다.

2.1 깊이 정보의 획득 및 표현

깊이 정보는 적외선이나 레이저를 사용하여 카메라로부터 사물까지의 거리를 측정, 표현한다. 사물과 카메라 사이의 거리 정보인 8비트의 깊이 정보와 RGB 텍스처 데이터를 획득하면 현실감 있는 3차원 영상을 생성할 수 있다. 깊이 카메라에 관한 연구가 많이 있어 왔다[3-4].



(그림 1) Gray scale 값과 깊이 좌표의 관계

깊이가 표현된 영상을 만들기 위해서는 깊이 정보를 카메라와 물체 사이의 거리인 깊이 좌표(z 좌표)로 변환해야 한다. 이러한 깊이 정보와 깊이 좌표와의 관계는 (그림 1)과 같다. (그림 1)에서 g 는 깊이 영상에서 각 픽셀의 깊이를 gray scale 값으로 표현한 것이고, z 는 깊이 카메라로부터 물체가 위치한 곳까지의 거리로서 관측 공간의 z -축 좌표를 의미한다. $zNear$ 와 $zFar$ 는 각각 깊이 카메라가 깊이를 측정할 수 있는 최소, 최대범위이다. 카메라의 관측 공간에서 볼 때 z 와 g 는 같은 위치를 의미하지만 서로 다른 형식으로 표현된다.

RGB 텍스처 영상에 그래픽 객체를 합성하기 위해서는 영상에 깊이를 나타내야 한다. 즉, 각 픽셀의 깊이 좌표 z 를 표현해야 한다. 각 픽셀의 깊이 좌표는 식 (1)을 이용하여 얻는다.

$$z = \frac{1.0}{\left(\left(\frac{g}{255}\right) \times \left(\frac{1.0}{zNear} - \frac{1.0}{zFar}\right) + \left(-\frac{1.0}{zFar}\right)\right)} \quad (1)$$

깊이가 표현된 장면을 만드는데 필요한 각 픽셀

의 깊이 값은 깊이 버퍼(depth buffer)에 저장되며, 깊이 버퍼의 깊이 값은 0과 1 사이의 실수로 표현된다. 깊이 좌표와 깊이 버퍼에 저장된 값 사이의 관계는 식 (2)와 같고, d 는 깊이 버퍼의 깊이 값이다.

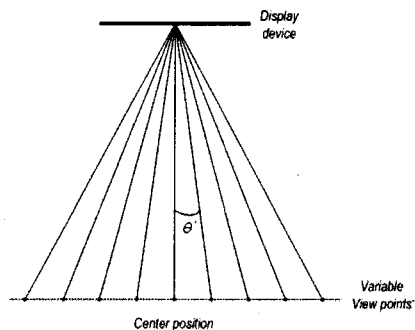
$$d = \left(\frac{1.0}{zNear} - \frac{1.0}{z}\right) / \left(\frac{1.0}{zNear} - \frac{1.0}{zFar}\right) \quad (2)$$

이렇게 변환된 깊이 값 d 를 깊이 버퍼에 채워 넣어 장면의 깊이 값을 표현한 후, 그래픽스 합성을 수행한다.

2.2 다시점 영상의 생성

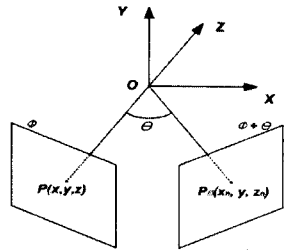
가상 다시점 영상은 한 대의 깊이 카메라로 획득한 RGB 텍스처 데이터와 깊이 정보를 사용하여 생성한다. 이때 깊이 정보가 시차(disparity) 정보로 변환되어 각 가상 시점의 영상을 생성할 수 있다[4].

관찰자의 시점에 따라 필요한 시점의 영상을 생성하기 위해서는 관찰자의 움직임을 고려해야 한다. 관찰자의 시점 변화를 각 θ 로 정의하면 관찰자의 시점에 해당하는 가상의 카메라에 투영되는 피사체의 상은 중앙 위치의 카메라에 투영되는 상에 대해 θ 만큼 회전한다[4-5].



(그림 2) 관찰자의 위치에 따른 시점

(그림 2)에서 관찰자의 수평 이동 정도를 각 θ 로 정의하여 회전(rotation)과 이동(translation) 등의 기하학적 변환을 적용하면 해당 시점에 대한 영상을 생성할 수 있다. 깊이 정보를 이용하여 여러 시점의 영상을 만들어내기 위해서는 (그림 3)과 같은 폭주식(convergence) 카메라 모델을 적용한다. (그림 3)에서 Φ 로부터 θ 만큼 회전한 영상 $P_\theta(x_n, y, z_n)$ 는 영상 $P(x, y, z)$ 의 Y 축에 대한 회전으로 표현된다[5].



(그림 3) 폭주식 카메라 모델

θ만큼 회전한 화소의 새 좌표를, 식 (3)에 따라 생성하면 입력 영상으로부터 θ만큼 회전한 영상을 생성할 수 있다. 즉, 식 (3)에 의해 (그림 3)의 왼쪽 영상 내 화소 $P(x, y, z)$ 은 오른쪽 영상 내 화소인 $P_{\theta}(x_n, y, z_n)$ 로 이동된다. 유사하게 $2\theta, -\theta$ 등의 여러 각도로 회전된 영상을 생성할 수 있다. 식 (3)에서 $I = (x, y, z)$ 이고 $I_{\theta} = (x_n, y, z_n)$ 이다.

$$\begin{bmatrix} x_n \\ y \\ z_n \end{bmatrix} = R_{\theta} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, I_{\theta}^T = R_{\theta} I^T \quad (3)$$

3. 영상과 그래픽스 객체의 합성

RGB 텍스처 영상과 그래픽스 객체를 합성하는 방법은 그래픽스 객체를 텍스처 영상을 구성하는 공간 상의 한 위치에 두는 것이다. 하지만 그래픽스 객체가 텍스처 영상의 한 물체보다 더 멀리 위치하면 그래픽스 물체의 일부가 텍스처 영상의 물체에 가려지는 경우가 발생한다. 이러한 문제는 텍스처 영상에 깊이 정보를 표현하여 해결할 수 있다. 텍스처 영상을 구성하는 각 픽셀들의 깊이가 표현되어 있으면 그래픽스 객체가 어디에 위치하던지 서로의 깊이를 비교할 수 있으므로 정확한 합성이 가능하다.

3.1 그래픽스 객체의 합성

그래픽스 합성을 위한 영상의 처리 과정에서 깊이 카메라로부터 얻은 RGB 텍스처 데이터와 깊이 데이터는 각각 컬러 버퍼와 깊이 버퍼에 저장된다. 그 후 그래픽스 객체의 RGB 텍스처와 깊이 데이터를 생성하여 RGB 텍스처 데이터와 깊이 데이터가 표현된 영상에 합성한다.

합성된 RGB 텍스처 영상은 읽어들인 컬러 버퍼의 값을 사용할 파일 포맷에 맞게 저장하고, 깊이

영상은 읽어들인 깊이 버퍼의 값을 식 (4)와 식 (5)에 차례로 적용시켜 깊이 좌표로 변환하여 gray scale 값으로 표현한다.

$$z = \frac{zNear \times zFar}{zFar - d(zFar - zNear)} \quad (4)$$

$$g = \frac{255.0 \times \left(\frac{1.0}{z} - \frac{1.0}{zFar} \right)}{\frac{1.0}{zNear} - \frac{1.0}{zFar}} \quad (5)$$

(그림 4)는 깊이가 표현된 RGB 텍스처 영상에 그래픽스 객체를 합성하여 얻은 RGB 텍스처 영상 (a)와 깊이 영상 (b)이다.



(a) RGB 텍스처 데이터 (b) Gray scale의 깊이 데이터

(그림 4) 그래픽스 합성된 데이터

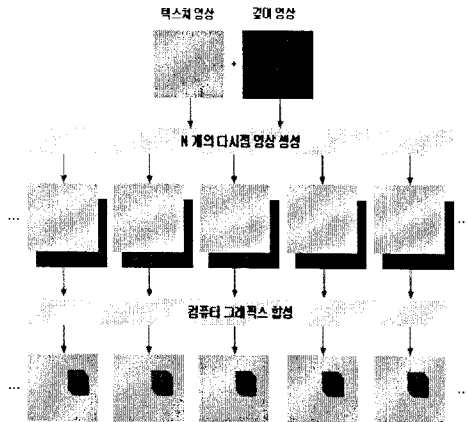
3.2 합성 방법

RGB 텍스처 영상과 깊이 영상을 입력으로 사용하여 다시점 영상을 생성한 후, 각 다시점 가상 영상에 그래픽스 객체를 합성한다. 각 다시점 영상은 카메라의 시점이 인간의 양안 시차만큼 이동된 결과 영상이므로 이동 및 회전 등의 변환을 통해 얻어진다. 그러므로 각 가상 시점 영상에 현실감 있는 그래픽스 객체의 합성을 위해서는 합성될 그래픽스 객체의 이동 및 회전 변환이 필요하다. 그래픽스 객체의 변환은 다시점 영상의 변환과 동일하게 이루어지는 것으로, 그래픽스 객체를 바라보는 가상 카메라의 위치를 이동시켜 얻을 수 있다.

(그림 5)는 합성 방법을 도식화 한 것으로 N 개의 다시점 영상을 생성한 이후에, 입력 영상과 그래픽스 객체의 합성이 이루어짐을 볼 수 있다. 다시점 영상의 깊이 영상들은 각 시점의 텍스처 영상의 생성 방법과 같은 방법을 사용하여 얻는다.

제안된 합성 방법은 다시점 영상을 생성한 후 그래픽스 객체를 합성하므로 각 시점마다 원하는 그래픽

스 객체를 정확하게 합성할 수 있다. 또한 입체 영상에 합성되는 그래픽스 객체의 실감 정도를 쉽게 조절할 수 있어, 그래픽스 객체가 화면 밖으로 돌출되어 보이는 정도를 자유롭게 조절할 수 있다.



(그림 5) 합성 방법

4. 실험 및 분석

본 논문에서 제안한 텍스처 영상과 깊이 영상을 이용한 그래픽스 합성을 위해 OpenGL을 사용한다. 그래픽스 객체의 합성 및 다시점 영상 생성 결과를 확인하기 위해 Microsoft Research의 break-dancing 중 100 프레임의 RGB 텍스처 영상과 깊이 영상을 입력으로 채택하였다.



(a) 좌 영상 (b) 우 영상

(그림 6) 합성 결과

(그림 6)은 제안된 합성 방법을 사용하여 MSR의 breakdancing frame과 적색 큐브를 합성한 결과이다. (a)와 (b)는 각각 가상 카메라의 두 시점인 좌, 우의 영상을 보여준다. 좌, 우 영상을 생성하기 위해서는 깊이 카메라로부터 얻은 영상을 입력으로 하여 카메라를 좌, 우로 이동시켜 생성되는 다시점 영상에 그래픽스 객체를 합성한다.

일반적으로 카메라로부터 멀어지면 카메라의 이동에 따른 그래픽스 객체의 변화가 적고, 카메라가

가운데 위치하면 카메라의 이동에 따른 변화가 크다. 결과 영상 (a), (b)에서 적색 큐브는 중앙의 댄서와 비슷하게 이동되었고, 뒤쪽의 회색 옷의 인물보다 많이 이동되어 자연스러운 가상 다시점 영상이 합성되었음을 알 수 있다.

5. 결론

본 논문에서는 기존의 다시점 카메라를 이용하는 다시점 비디오 생성 방법과는 다르게, 한 개의 RGB 텍스처와 깊이 정보로부터 다시점 가상 영상을 생성하는 방법에 기반을 둔 그래픽스 합성 방법을 제안하였다. 제안된 방법은 시점의 개수에 비례하여 그래픽스 합성 과정이 반복되지만, 그래픽스 객체를 텍스처 영상에 원하는 모습으로 합성할 수 있으므로 비실시간 다시점 그래픽스 영상 합성 과정에 적용하기가 아주 적합하다. 비실시간 다시점 그래픽스 영상 합성은 시간, 자원이 충분한 작업이므로 보다 정확한 합성 결과를 얻는 것이 더욱 중요하기 때문이다.

감사의 글

본 연구는 정보통신부 및 정보통신연구진흥원의 SmartTV 개발 및 대학 IT연구센터 지원사업의 연구 결과로 수행되었음.(IITA-2006-C1090-0603-0017)

참고문헌

- [1] Veksler, O. "Semi-dense stereo correspondence with dense features", IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001), pp. 149~157, Dec. 2001.
- [2] A. Redert, et al., "ATTEST : Advanced Three-Dimensional Television System Technologies", First International Symposium on 3D Data Processing Visualization and Transmission, pp. 313-319, Jun. 2002.
- [3] C. Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV", Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems XI, pp. 93-104, San Jose, CA, USA, January 2004.
- [4] Se hoon Jang, Chung shin Han, Jin woo Bea, Ji sang Yoo, "Real-time Multiple stereo Image Synthesis using Depth Information", The Journal of Korean Institute of Communication Sciences Vol.3 No.4C, April 2005.
- [5] Yao Wang, Jorn Ostermann, Ya-Qin Zhang, "Video processing and communications", Prentice Hall, pp.374-383, 2002.