

## 가상시점 영상 생성을 위한 효율적인 홀 채움 방법

문지훈, 호요성

광주과학기술원

jhm@gist.ac.kr, hoyo@gist.ac.kr

### Efficient Hole Filling Method for Producing Virtual View Images

Ji-Hun Mun, Yo-Sung Ho

Gwangju Institute of Science and Technology

#### 요 약

본 논문에서는 가상시점 영상을 생성할 때 발생하는 홀 영역을 효율적으로 채우는 방법을 제안한다. 가상시점 영상을 생성하려면 우선 깊이 영상에 대해 3차원 워핑을 수행한 뒤, 이때 발생하는 작은 홀을 미디언 필터를 이용하여 적절히 채워야 한다. 홀이 채워진 깊이 영상을 기반으로 하여 입력받은 참조 컬러 영상에 대해 3차원 워핑을 수행하여 가상 위치에 새로운 시점 영상을 생성하게 된다. 이때 또한 마찬가지로 3차원 워핑을 수행하기 때문에 홀 영역이 발생하게 된다. 텍스처 영상을 워핑하여 새로운 좌표계로 옮긴 영상은 주변 컬러 화소들과의 관계들을 가지고 있다. 텍스처 영상을 워핑한 결과 영상에서 발생하는 홀 영역을 채우기 위해 방향성을 고려한 홀 채움 방법을 사용한다. 홀 주변 화소 영역의 값들을 홀을 채우게 될 후보 화소 값으로 설정한 뒤, 각각의 화소값에 대해 비용값을 계산한다. 이때 가장 적은 비용값을 갖게 하는 주변 화소 값을 해당 영역의 홀 채움 값으로 사용하게 된다. 좌영상과 우영상을 워핑할 때 발생하는 홀 영역의 위치가 각각 다르게 나타난다. 홀 영역은 배경화소 값을 이용해 채울 경우 자연스러운 결과를 보인다. 배경화소 값을 이용하기 위해 좌영상과 우영상에 따른 새로운 홀 스캔 방향 또한 제안한다. 능동적으로 홀 스캔 방향을 선택하여 홀 주변 화소값들을 스캔해가며 워핑 결과 발생하는 홀 영역을 효율적으로 채우게 된다. 결과적으로 제안한 방법을 통하여 생성된 가상시점 영상의 화질이 좋아지는 결과를 확인할 수 있었다.

## 1. 서론

기존 양안식 영상 시점보다 나은 몰입감과 입체감을 제공하는 다시점 영상에 대한 연구가 활발하게 진행되고 있다[1]. 다시점 영상을 넘어 홀로그래픽 영상을 제작하는 것을 최종 구현 단계로 연구를 진행하고 있으며, 최근에는 슈퍼 멀티뷰 카메라 시스템을 사용한 자유시점 TV 구현에 대한 연구를 중점적으로 하고 있다.

이와 같은 3차원 디스플레이의 발전으로 인해 3차원 TV는 사용자에게 보다 실감나는 콘텐츠를 제공할 수 있게 되었다. 방송 시스템이 아날로그에서 디지털로 변하게 되면서 영상 산업의 동향에 큰 변화를 가져왔듯이, 머지 않아 3차원 디지털 미디어가 영상 산업의 또 다른 개혁을 가져올 것으로 기대를 모으고 있다.

이러한 시스템을 바탕으로 점차 3차원 비디오 시스템에 대한 수요가 증가하고 있는 상황이다. 3차원 산업은 영화, 게임, 방송 등 다양한 멀티미디어 적용 분야에서 새로운 미디어 서비스로 자리잡아가고 있다. 3차원 영상 산업을 위한 시스템은 이미 많은 연구를 통해 기술의 미래성과 창출 효과에 대한 입증 이 완료되었다. 하지만 최근에는 3차원 영상을 시청하기 위해 보조로 사용하는 보조 안경의 불편함과 3차원 영상의 화질 저

하로 인한 눈의 피로감, 또한 3차원 비디오 영상 콘텐츠의 부재로 인해 많은 연구자들이 예상보다 3차원 영상의 기술 개발의 속도가 느리다는 평을 내리고 있다.

3차원 영상 산업이 보다 빠르게 발전하기 위해서는 현재 많이 사용되고 있는 2차원 영상의 콘텐츠 처럼 3차원 콘텐츠를 제작하기 쉬워야 하며, 시청자들에게 접근하기 용이하도록 해야 한다. 또한 보조 안경을 착용함으로써 발생하는 눈의 피로감, 불편함을 개선해야 하며, 무엇보다 고화질의 3차원 영상을 제공하여 어느 디스플레이 장치에서 시청해도 불편함이 없도록 해야 한다[2].

세계 여러 나라에서는 3차원 방송을 위한 3차원 영상 콘텐츠를 제작하고 있으며 국내에서도 여러 방송 사업기관들을 중심으로 3차원 시험방송을 시행하고 있다. 2002년부터 현재까지 유럽에서는 3차원 TV 에 관한 연구를 지속적으로 수행하고 있다. ATTEST (advanced three-dimensional television system technologies) 는 3차원 콘텐츠의 생성과 부호화 전송 그리고 렌더링과 재현 총 4단계의 구조를 가지고 있는 3차원 방송 시스템에 대해 소개하기도 했다 [3].

지금까지 연구되었던 3차원 방송 시스템이 스테레오 영상 스트림에 대한 내용인 반면, ATTEST 프로젝트는 단일 시점의 색상 및 깊이 영상 스트림을 사용하는 깊이 영상 기반 렌더링 (DIBR, depth image-based rendering) 기술을 사용했다는 차

이점이 있다.

그림 1은 가장 빈번하게 사용되는 3차원 비디오 시스템의 개념도를 크게 부호화기와 복호화기를 중심으로 하여 간략히 나타낸 것이다

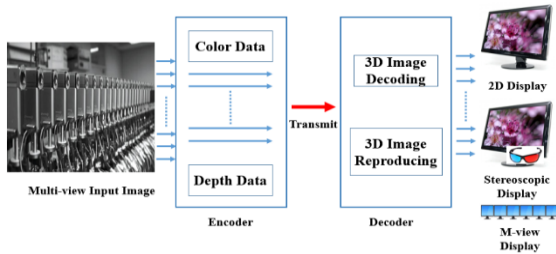


그림 1. 3차원 비디오 시스템

3차원 비디오 시스템은 각 시점의 깊이 영상을 포함하는 3차원 비디오의 모든 처리 과정을 포함하고 있다. 기본적으로 N 시점의 영상과 함께 N 시점의 깊이 영상을 동시에 처리 가능하도록 설계되어 있다.

3차원 비디오는 스테레오 카메라, 깊이 카메라, 다시점 카메라 등 다양한 제작 도구를 이용하여 제작할 수 있는데, 깊이 영상을 직접적으로 획득할 수 있는 깊이 카메라를 제외한 나머지 제작 도구들을 이용할 경우 장면의 3차원 정보를 계산적으로 탐색하는 방법을 이용하여 깊이 영상을 획득한다. 입력 받은 컬러 영상과 깊이 영상은 부호화 되어 전송되고 복호기에서 이러한 정보를 복원한다. 그리고 이때 사용자가 사용하는 디스플레이 기기에 따라 다르게 영상을 보여지게 된다.

다시점 3차원 영상 정보를 전송할 경우 보내야 할 정보의 양이 많아져 정보를 효율적으로 보낼 수 없다는 문제가 발생하게 된다. 이러한 단점을 보완하기 위해 가상시점 영상을 합성(virtual view synthesis)하는 방법을 사용하여 3차원 영상 전송을 하도록 하는 방법이 사용되는데, 이때 발생하는 홀 영역을 효율적으로 채우는 방법을 본 논문에서 다루도록 한다.

## 2. 3차원 워핑을 이용한 가상시점 합성

### 1. 깊이 영상 기반 렌더링

임의의 가상시점 영상을 생성하기 위해서는 각각의 좌영상과 우영상의 참조 영상들을 생성하고자 하는 가상시점의 위치로 좌표계 변환 과정을 수행해야 한다. 즉 참조 시점의 영상에 있는 객체를 가상시점 공간으로 옮겨야 하는데 이 방법은 깊이 영상 기반 렌더링(depth image based rendering) 방법 중 하나인 3차원 워핑 기술을 이용하면 쉽게 결과를 얻을 수 있다.

그림 2는 깊이 영상 기반 렌더링 기술의 간략한 흐름도를 나타낸다. DIBR에서는 색상 영상과 깊이 영상을 이용하여 가상시점을 합성한다. 가상시점을 합성할 때 3차원 워핑을 이용하는데 이 과정에서 빈 영역(hole)이 발생하게 된다. 이러한 빈 영역은 다양한 보간법(interpolation) 과정을 이용해 채워나가게 된다. 가상시점 합성을 수행하는 과정에서 비폐색 영역이 발생하는 경우 비폐색 영역에 대한 정보가 없기 때문에 이 부분을 얼마나 자연스럽게 채우느냐가 합성 영상의 화질을 좌우한다.

비폐색 영역을 채우는 방법을 설명하기에 앞서 우선 3차원 워핑의 원리에 대해 설명하도록 한다.

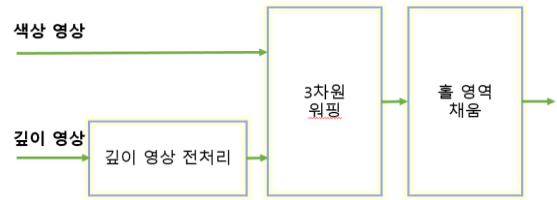


그림 2. 깊이 영상 기반 렌더링 구조

### 2. 3차원 워핑

3차원 워핑의 도입은 기존 3차원 공간에 있는 물체의 좌표를 카메라로 촬영하게 되면 멧히게 되는 2차원 공간상의 위치로 옮기기 위해 좌표계 변환의 개념으로부터 발전되어왔다. 그림 3은 핀홀 카메라 모델의 영상 투영 구조를 2D 형태로 표현한 것이다.

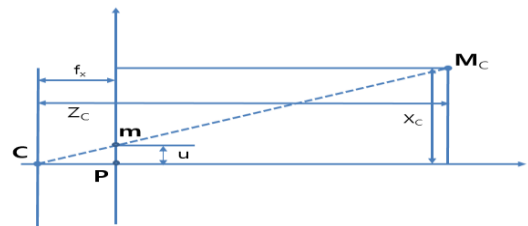


그림 3. 핀홀 카메라 모델의 투영 구조

핀홀 카메라로 촬영한 영상은 주로 x 축의  $f_x$  에 역으로 상이 멧히게 된다. 카메라를 통해 촬영한 영상은 2D 구조로 저장되지만 실제로 촬영된 사물들은 3D 구조상에 위치해 있으므로 정확한 위치를 정의하기 어렵다. 이 때문에 영상이 멧히는 위치를 2D 구조의  $f_x$  의 위치로 옮겨서 해석하게 된다.  $f_x$  의 위치를 카메라의 초점거리라고 하는데 이는 실제 물체의 3D 좌표가 영상 평면에 투영되어 상이 멧히는 거리를 나타낸다. 이 거리는 삼각 비례법을 사용하여 알아낼 수 있다. 아래 식 (1)은 그림 3에 나타나 있는 변수들의 관계를 삼각 비례법을 사용하여 정리한 결과이다.

$$\frac{u}{X_c} = \frac{f_x}{Z_c} \quad (1)$$

식 (1)에서 변수  $u$ 에 대해 정리하면 다음 식(2)와 같이 정리할 수 있고, 이는 실제 3차원 좌표계의 영상이 멧히는 영상 평면의 위치를 나타낸다.

$$u = \frac{f_x X_c}{Z_c} \quad (2)$$

카메라 좌표계 C의 시점에서 보면 초점거리 P의 위치가  $(0,0)^T$ 가 아니므로 원점의 위치를 찾아내기 위해 보정할 수 있도록 보정 변수를 사용하여 식 (3)과 같이 나타낼 수 있다.

$$u - p_x = \frac{f_x X_c}{Z_c} \quad (3)$$

$$\therefore u = \frac{f_x X_c}{Z_c} + p_x, v = \frac{f_y Y_c}{Z_c} + p_y$$

3차원 워핑은 지금까지 설명한 원리를 기반으로 하여 수행된다. 임의의 가상시점 영상을 생성하기 위해서는

좌영상과 우영상 각각의 참조 영상들을 3차원 워핑 방법을 이용하여 가상시점의 위치로 이동시켜야 한다.

MacMilan 이 제안한 3차원 워핑을 이용한 중간시점 영상의 생성 방법은 카메라 파라미터와 깊이 정보를 이용하여 참조시점의 각 화소가 가상시점영상의 어느 위치에 생성되는지 정의하여 가상시점 영상을 생성하게 된다 [5]. 3차원 좌표계에 있는 물체의 영상을 2차원 카메라 좌표계로 이동하기 위해서는 회전(rotation)과 이동(translation) 변환을 수행해야 한다.

카메라 파라미터는 세계 좌표계와 카메라 좌표계와의 관계를 정의하기 위해 필요한 수치들이다. 카메라 파라미터는 외부와 내부 파라미터 값으로 나뉘지는데, 회전과 이동 변환은 카메라 외부 파라미터로 정의된다. 내부 파라미터 값은 그림 3에서 구한 실제 촬영된 영상에 맺히는 카메라 평면의 위치  $f_x$ 를 포함하는 정보로 정의된다. 카메라 파라미터를 이용한 3차원 워핑 과정을 식 (4)와 같이 나타낼 수 있다.

$$Z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = A \cdot M_c \quad (4)$$

식 (4)에 나타나 있는 A와  $M_c$ 는 각각 카메라 내부, 외부 파라미터 값을 의미한다.  $M_c$ 는 회전변환 R과 이동변환 t 그리고 좌변에 있는  $Z_c(u,v,1)^T$ 를  $m'$ 으로 변환하여 식을 재 정리하면 식 (5)와 같이 표현 할 수 있다.

$$m' = A \cdot R \cdot M_w + A \cdot t \quad (5)$$

식 (5)는 세계좌표계의 값을 가상시점으로 이동시킨 결과를 나타내고 위 정보에 깊이 정보  $d_r$ 를 이용하여 세계 좌표계로 재투영한 결과를 얻어낼 수 있다. 깊이 정보를 이용한 재 투영 결과는 식 (6)과 같이 표현 가능하다.

$$M = R^{-1} \cdot A^{-1} \cdot m_s \cdot d_r - R^{-1} \cdot t \quad (6)$$

식 (6)은 세계 좌표계에 있는 영상의 정보를 가상시점의 위치에 옮긴뒤 다시 재 투영한 결과를 표현하는 식을 나타내는데, 본 논문에서 중점적으로 다루게 될 가상시점 영상 합성에서는 이전 단계까지만 주로 사용되게 된다. 3차원 워핑을 이용하여 가상시점 영상을 생성하기 위해 입력 영상으로 받아들이는 좌영상과 우영상 각각의 워핑 결과는 그림 5와 같다.



(a) 좌영상과 3차원 워핑 결과



(b) 우영상과 3차원 워핑 결과

그림 5. 좌영상과 우영상의 3차원 워핑 결과

그림 5에서 각각 위 아래는 좌영상과 우영상을 워핑한 결과이다. 우측의 그림이 워핑한 결과를 나타내는데, 워핑 결과 검정색인 부분이 많이 나타남을 확인할 수 있다. 이 영역을 홀 영역 또는 비폐색 영역이라고 하는데 이러한 영역은 3차원 워핑의 수행 결과 나타나게 된다.

가상시점 영상 합성을 위해 깊이 영상 기반 3차원 워핑을 수행하게 되는데, 이때 깊이 영상의 깊이 정보가 정확하지 않기 때문에 해당 영역의 정보 부재로 인해 홀 영역이 워핑 결과 발생하게 된다. 가상시점의 영상 합성 과정에서 홀 영역을 얼마나 효율적으로 채우는가 하는 것이 고화질의 가상시점 영상 합성 결과를 얻게 하는 중요한 요인이 된다.

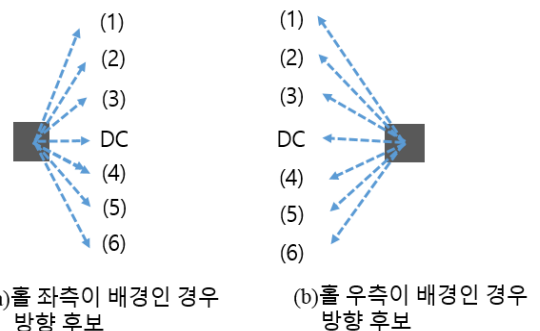
### 3. 기존의 홀 채움 방법

3차원 워핑 결과 발생하는 다양한 홀 영역을 채우기 위해 VSRS (view synthesis reference software)에는 깊이 영상 워핑 결과 나타나는 작은 홀 영역을 미디언 필터를 통해 제거한뒤 후처리된 깊이 영상 기반 텍스처 영상에 대해 3차원 워핑을 수행한다.

이러한 방법으로 홀을 채운 경우 홀 주변 텍스처 영역의 정보와 배경 영역의 정보를 동시에 사용하기 때문에 채워진 홀 영역에서 계단 현상과 같은 잡음이 발생하게 된다. 홀 채움을 효율적으로 채우기 위해 좌영상과 우영상 각각에 대한 다른 홀 스캔 방향을 사용하고 홀 주변의 방향성을 고려한 홀 채움 방법이 제안되었다[4].

본 논문에서는 잡음 현상을 제거하기 위해 홀 주변의 7가지 방향성을 고려한 홀 채움 방법을 제안한다. 3차원 워핑은 좌영상과 우영상에 대해 각각 워핑을 수행하기 때문에 발생하는 홀 영역의 위치 또한 제각각이다. 그래서 제안하는 홀 채움 방법에서는 홀이 발생하는 위치를 능동적으로 고려하여 지그재그 방식이 아닌 수직 홀 스캔 방법을 제안한다.

그림 6은 좌영상과 우영상에 대해 7가지 방향성을 고려하는 경우를 나타내고 있다.



(a) 홀 좌측이 배경인 경우 방향 후보 (b) 홀 우측이 배경인 경우 방향 후보

그림 6. 방향성을 고려한 홀 주변 영역 탐색

홀 좌측이 배경인 경우와 홀 우측이 배경인 경우에는 홀 영역을 스캔하는 방향성 또한 고려해줘야 한다. 홀 영역이 좌측과 우측에 있을 경우 각각 좌측과 우측에서부터 시작하는 수직 스캔 방법을 이용하여 홀 영역을 효

울적으로 탐색할 수 있도록 한다. 이 과정을 그림 7에서 보이고 있다.

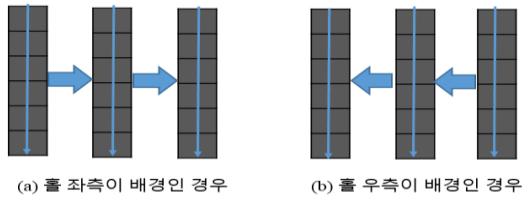


그림 7. 홀 영역에 따른 스캔 방향

방향성을 고려한 홀 채움 후보군의 적합한 화소값을 찾는 비용값은 식 (7)과 같이 나타낼 수 있는데, 이는 SAD (sum of absolute differentiation)를 사용하여 그 차이 값을 구해내는 과정이다.

$$Cost = \frac{\sum_{x,y \in Neighbor} |ZS(x,y) - DC(x,y)|}{NeighborNumber} \quad (7)$$

식 (7)에서 ZS는 홀 주변 후보군들의 화소값들을, DC는 홀을 채우게 될 화소값 그리고 분모는 홀 채움 후보군들의 총 화소 개수를 의미한다. 이 비용값의 최소를 갖게 하는 화소가 홀 영역을 채우게 될 화소값으로 선택 되게 된다. 이 과정은 식 (8)에 나타나 있다.

$$Direction = \arg \text{Min}\{Cost\} \quad (8)$$

### 3. 실험 결과 및 분석

2장에서 제안한 새로운 홀 채움 방법의 성능을 확인 하기 위해 MPEG 에서 제공하는 다시점 영상들을 가지고 실험을 수행했다. 실험 결과를 그림 8에서 나타내고 있다. 실험 결과 그림에서 좌측그림은 VSRS 에 구현되어 있는 원래의 홀 채움 방법을, 가운데는 홀 스캔 방향만을 바꿔서 수행한 홀 채움 결과를 그리고 오른쪽은 제안한 방향성과 홀 스캔 방향 두가지 동시에 고려한 경우의 실험 결과를 나타낸다.

실험 결과 방향성까지 동시에 고려한 홀 채움 결과가 기존의 방법과 홀 스캔 방향만을 바꿔서 수행한 결과보다 더 효율적으로 홀 영역이 채워짐을 확인 할 수 있다.

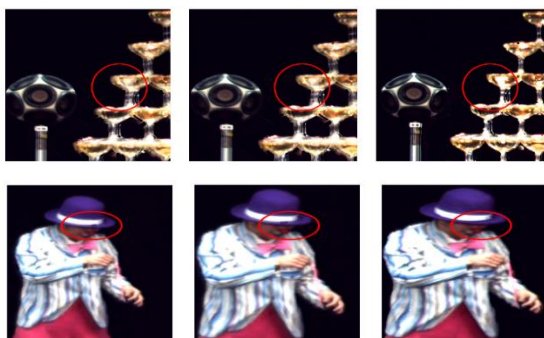


그림 8. 실험 결과 영상

### 4. 결론

본 논문에서는 가상시점 영상 합성시 워핑결과 발생 하는 홀 영역을 효율적으로 채우는 방법을 제안한다. 홀 주변 영역에 대해 홀 영역을 채우게 될 주변 화소값들을 후보군으로 선정후, 방향성을 고려하여 홀 영역을 채울 수 있도록 한다. 또한 홀 주변의 배경 영역을 고려하여 서로 다른 홀 스캔 방향 또한 제안한다. 제안한 홀 채움 기법을 사용하여 홀을 채운 결과 기존 방법에 비해 자연스럽게 홀 영역을 채울 수 있음을 확인할 수 있었다.

### 감사의 글

이 논문은 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2014-050807)

### 참고문헌

- [1] G.Riva, F.Davide, W.A.Ijsselsteijn, Being There Concept, "Effects and Measurement of User Presence in Synthetic Environments," Ios Press, Amsterdam, Netherlands, 2003.
- [2] G.M.Um, G.H.Cheong, W.S.Cheong, N.H.Hur, "Technical Development and Standardization Trends of Multi-view 3D and Free-view Point Video," The magazine of the IEEK, vol.38, no.2, pp. 18-23, Feb. 2011.
- [3] A.Redert, M.O.Beeck, C.Fehn, W.Ijsselsteijn, M. Pollefeys, L.Van Gool, E.Ofek, I.Sexton, P.Surman, "ATTEST: Advanced Three-dimensional Television System Techniques," Proc. of International Symposium on 3D Data Processing, pp.313-319, June 2002.
- [4] J.H.Mun, Y.S.Ho, "New Hole Scan and Considering Direction for Effectvie Hole Filling Method," The 2014 Korea Signal Processing Conference, pp. 9-12, Oct. 2014.
- [5] L.McMillan, "An Image-based Approach to Three Dimensional Computer Graphics Technical Report," Ph.D Dissertation, UNC Computer Science, 1999.