

## 이종 카메라를 이용한 Eye-contacted 영상 생성 기법

\*현지호, \*한재영, \*유지상

광운대학교

\*notoriousjay@kw.ac.kr, \*carism@kw.ac.kr, \*jsyoo@kw.ac.kr

### Eye Contacted View Generation by using Color and Depth Cameras

\*Jeeho Hyun, \*Jaeyoung Han, \*Jisang Yoo

Kwangwoon University

#### 요 약

중간시점 영상은 스테레오 정합 방식을 이용하여 구한 깊이 지도를 이용하여 생성하는 것이 일반적인 방법이다. 그러나 대부분의 스테레오 정합 방식들은 좌, 우 영상의 조명환경이나 기하학적으로 수평 또는 수직방향이 일치하지 않으면 잘못된 깊이 지도를 획득하는 단점이 있다. 이러한 단점들은 정합 과정을 통해 획득한 깊이 지도를 이용하여 다시점 영상 생성 시 더 많은 홀과 경계 잡음을 생성하게 된다. 이러한 문제점을 보완하기 위하여 본 논문에서는 RGB 컬러 카메라 1 대와 깊이 카메라를 이용하여 중간 영상을 생성하는 방법을 제안한다. 제안된 기법을 이용하여 화상회의 시 사실감 및 현실감을 증대할 수 있는 eye-contact 시점 영상을 생성하고 이때 발생하는 홀과 경계잡음 제거를 실시간으로 처리하기 위한 효율적인 기법을 제안한다.

#### 1. 서론

최근 기존의 대중화된 CCD, CMOS 센서 등과 함께 저렴하면서도 비교적 고성능으로 깊이 정보를 획득할 수 있는 센서의 출현으로 인해 영상 처리 분야는 또 다른 변화를 맞이하고 있다. 특히 깊이 정보를 정확히 측정할 수 있게 됨에 따라 Tele-presence 분야에서는 음성 또는 동작 인식과 융합하여 더 사실적인 사용자 경험(user experience)을 전달하려는 노력을 하고 있다[1,2].

하지만 현재 기술 수준에서 얻을 수 있는 깊이 정보의 해상도와 품질은 일반 컬러 영상과 비교하여 매우 부족한 수준이다. 이러한 이유로 깊이 영상의 해상도와 품질을 향상시키기 위한 적절한 신호처리 기술의 개발이 필요하다[3].

화상회의 시스템에서 화자간 시선을 마주보고 대화할 수 있는 eye-contacted 영상의 생성은 현재의 기술로 물리적인 방법으로는 불가능하다. 따라서 가상시점을 생성하는 방법 중 하나인 DIBR(depth image-based rendering) 기법을 이용하는 경우가 대부분이다. 이 방법은 텍스처 정보를 가진 컬러 영상과

영상의 각 화소에 대응하는 깊이 정보로 이루어진 깊이 지도(depth map)를 이용하여, 임의의 시점 영상을 생성하고 렌더링하는 것이다[4,5].

본 논문에서는 화상회의 시스템의 현장감 및 사실감 향상을 위한 eye-contacted 시점을 실시간으로 생성하는 기법을 제안한다. 제안한 기법에서는 객체와 배경 사이에 발생하는 잡음과 객체 내부에 생성되는 홀 영역을 효율적이고 빠르게 제거하는 방법도 포함되어 있다.

#### 2. 제안하는 기법

본 논문에서는 먼저 3 차원 워핑(3D warping)에 필요한 파라미터를 추출하기 위해 두 카메라에서 획득된 영상의 해상도를 동일하게 한다. 실험에 사용된 깊이 카메라는 MesaImaging 의 SR4000 으로 QCIF(176x144) 해상도의 깊이 영상을 초당 30 프레임 이상 획득할 수 있다. RGB 컬러 영상을 촬영하는데 사용된 카메라는 Pointgrey 사의 Grasshopper 모델로 최대 2448x2048 해상도의 영상을 초당 15 프레임 획득이 가능하다.

두 카메라에서 획득된 영상이 동일한 해상도를 갖도록 하기 위하여 QCIF 해상도의 64 배인 1408x1152 의 해상도로 조정한다. 이때 깊이 카메라에서 획득된 깊이 정보는 영상 보간에 일반적으로 사용되는 고등 차수 보간법(bi-cubic interpolation)을 이용한다.

그림 1 과 같이 두 카메라의 보정(stereo camera calibration) 작업을 통해 획득된 파라미터를 이용하여 3 차원 워핑을 통해 깊이 카메라에서 획득된 깊이 정보를 컬러 영상으로 매핑하고, 컬러 영상과 매핑된 깊이 정보를 이용하여 eye-contacted 시점의 영상을 생성하게 된다.

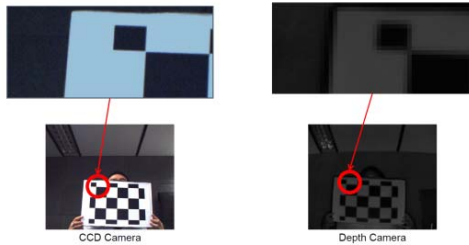


그림 1. 컬러 카메라와 깊이 카메라간의 보정 작업

### 2.1. 보간된 깊이 정보의 잡음 제거

고등 차수 보간법으로 업샘플링된 깊이 영상의 경우, 보간될 화소의 주변 16 개 화소의 값을 이용하는 특성에 따라 배경과 객체 사이에 그림 2 와 같은 잡음이 발생한다. 이러한 잡음은 깊이 값의 오류를 유발하여 3 차원 워핑 시 원하지 않는 또 다른 잡음을 발생하게 한다. 객체 내부의 깊이 정보의 경우 연속적인 값을 갖게 되는 특징을 이용하여 객체 외각 부분에 생기는 잡음을 워핑 전에 제거하여 화질을 향상시킨다.



그림 2. 깊이 정보의 경계 잡음

본 논문에서는 보간된 깊이 정보의 경계 부분에 생기는 오류를 제거하기 위하여 그림 3 과 같은 전체리 과정을 수행한다.

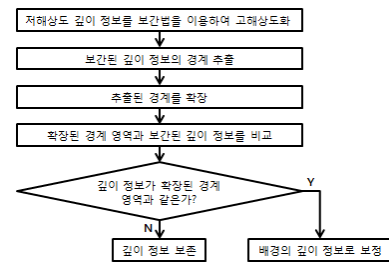


그림 3. 깊이 정보 경계 잡음 제거 순서도

### 2.2. 실시간 홀 채움

제안하는 기법에서는 최종 워핑하여 생성한 eye-contacted 시점의 컬러 영상에서 생기는 홀을 채우기 위하여 그림 4 와 같은 방법을 이용한다. 먼저 화소 단위로 홀 영역 여부를 판정하고 홀 영역일 경우, 그림 4(a)와 같이 8-이웃(8-neighbor) 화소의 값을 확인하여 이웃한 홀 영역을 제외한 나머지 화소값의 중간값을 계산하여 홀 영역을 채우게 된다. 그림 4(b)에 워핑 후 홀 영역을 채우는 방법을 보였다.

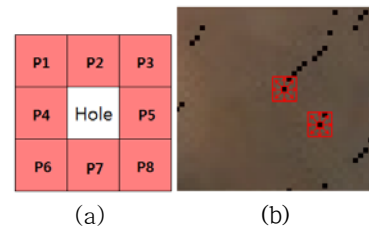


그림 4. (a) 8-이웃 화소 (b) 제안하는 홀 채움 방법

## 3. 실험 조건 및 결과

고등 차수 보간법을 이용하여 업샘플링된 깊이 영상의 경계는 Canny 연산자를 이용하여 얻는다. 그림 5(a)는 업샘플링된 깊이 영상이고 (b)는 추출된 경계 부분을 나타낸다.



그림 5. (a) 업샘플링된 깊이 정보 (b) Canny 연산자를 이용하여 추출된 깊이 영상의 경계

추출된 경계 정보를 이용하여 깊이 영상의 잡음을 제거하기 위해 그림 6 과 같이 경계 부분을 연속적으로 확장한다.

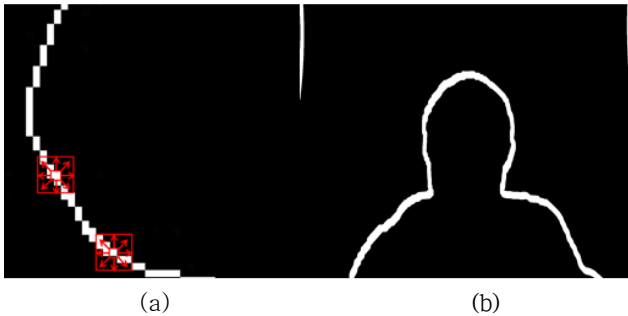


그림 6. (a) 경계 영역 확장 과정 (b) 확장된 깊이 정보의 경계 영역

객체와 배경의 경계를 나타내는 화소를 중심으로 그림 6(a)와 같이 8 개의 이웃화소를 비교하여 이 중에 경계 화소가 아닌 부분을 연속해서 채워나가는 방식을 사용한다. 이것을 연속적으로 15 회 정도 반복하여 (b)와 같이 경계를 확장한다. 이러한 방법으로 확장된 경계 정보와 업샘플링된 깊이 정보를 비교하여 그림 7 과 같이 깊이 정보를 보정한다.



그림 7. (a) 업샘플링된 깊이 영상 (b) 경계영역이 보정된 깊이 영상

그림 7(a)는 업샘플링된 깊이 영상이고 (b)는 업샘플링된 깊이 영상의 잡음이 제거된 결과이다. 각각의 깊이 정보를 이용하여 컬러 카메라의 위치로 위평한 결과는 그림 8 과 같다. 그림 8(a)는 경계 부분을 보정하지 않은 깊이 정보를 이용하여 컬러 영상과 매핑시킨 결과이며 (b)는 경계 부분을 보정한 깊이 정보를 컬러 영상과 매핑시킨 결과이다. (c)와 (d)는 (a)와 (b)를 확대한 영상으로 경계 보정을 통해 객체 외부에 나타나는 오류가 수정된 것을 명확하게 확인할 수 있다.

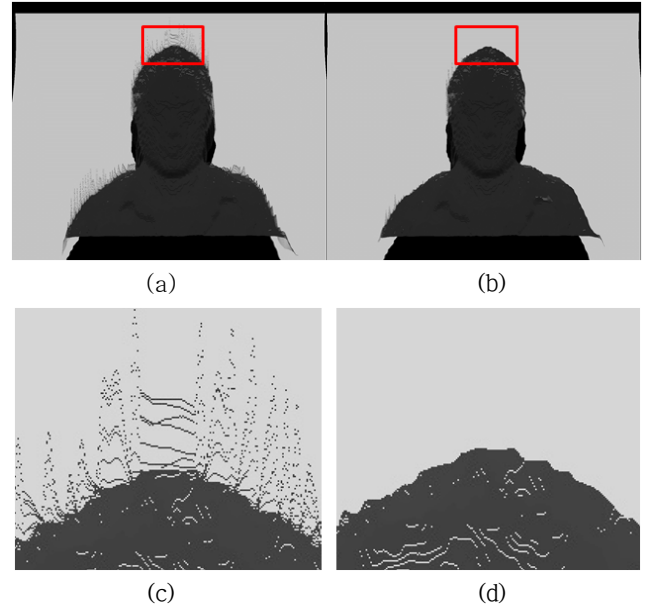


그림 8. 깊이 정보를 컬러 영상과 매핑한 결과 (a) 경계 보정 전, (b) 경계 보정 후 (c) (a)를 확대한 영상 (d) (b)를 확대한 영상

그림 8(a)와 (b)를 이용하여 각각의 위평된 깊이 정보를 이용하여 eye-contacted 시점으로 위평한 영상은 그림 9 와 같다. 그림 9(a)는 보정하지 않은 깊이 정보를 이용한 결과이며 (b)는 보정된 깊이 정보를 이용한 결과이다.

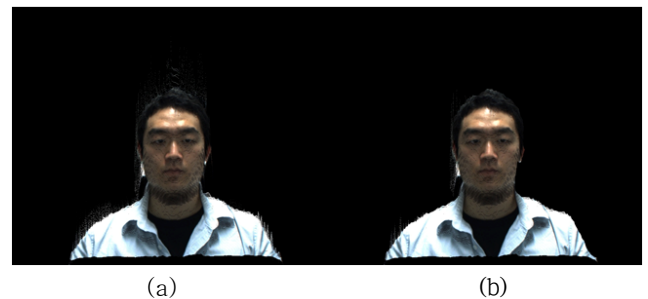


그림 9. Eye-contacted 시점 영상 (a)보정 전 (b)보정 후

깊이 영상의 경계 부분을 보정함으로써 최종 생성된 Eye-contacted 시점 영상에서 인물 외부에 생기는 잡음을 작은 연산 복잡도를 가진 방법을 통해 제거하는데 효과가 있음을 그림 9 에서 확인할 수 있다. 그러나 이러한 깊이 정보의 경계부분 보정을 통해서도 객체 내부에 생기는 홀 영역은 채울 수 없다.

높은 복잡도를 가진 홀 채움 기법은 고품질의 결과를 획득할 수 있으나 긴 연산 시간으로 인해 사용할

수 없다. 따라서 본 논문에서는 2.2 절에서 제안하는 홀 채움 기법을 통하여 객체 내부에 생긴 홀을 효율적으로 제거한다. 그림 10(a)는 홀 채움 기법을 적용하기 전의 영상이며, (b)는 홀 채움 기법을 적용한 최종 결과 영상이다.

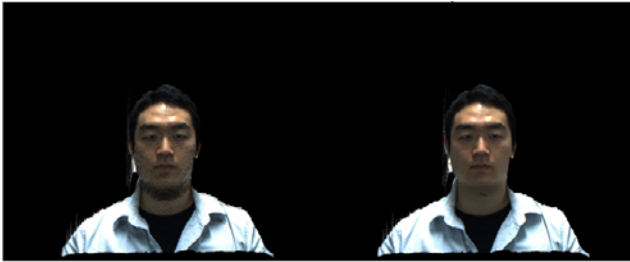


그림 10. (a) 홀 채움 전 영상 (b) 홀 채움 후 영상

실시간 처리를 위해 3 차원 워핑 과정은 GPU 기반 OpenCV 라이브러리를 이용하였으며 제안하는 기법의 처리 시간은 표 1 과 같으며 실험에 사용된 PC 의 사양은 표 2 와 같다.

표 1. 제안한 기법의 처리 시간

Progress	Processing Time(ms)
Post-processing of interpolated depth information	150
Warping depth information to texture information	14
Warping texture & depth information to eye-contact view	16
Hole filling	61
Total time	241

표 2. 실험에 사용된 PC 사양

Test environment	
CPU	Intel® Core™ i5-2500 CPU @ 3.30GHz
RAM	4.0GB
GPU	GeForce GTX 550Ti
OS	Windows 7 64-bit
IDE	Visual Studio 2008 SP1 (x86)

표 1 과 같이 eye-contacted 시점 한 프레임을 생성하는데 약 241msec 로 초당 약 4 프레임 정도 획득하는 것을 확인하였다.

#### 4. 결론

본 논문에서는 텍스처 정보를 촬영할 수 있는 컬러 카메라 1 대와 깊이 정보를 획득할 수 있는 깊이 카메라 1 대를 이용하여 화상 회의 시 현장감과 사실감을 극대화하기 위한 eye-contacted 시점 영상 생성 기법을 제안하였다. 제안한 기법에는 깊이 정보의 오류를 제거하고 홀 영역을 채우는 방법도 포함하고 있다. 화상 회의의 특성 상 실시간성이 충족되어야하기 때문에 연산 복잡도가 높지 않은 간단한 기법을 사용하여 높은 품질의 영상을 생성하였다. 향후 GPU 등의 프로세서를 활용한 연산속도 증대를 통해 최적화 과정 등의 연구가 필요하다고 생각된다.

#### <감사의 글>

이 논문은 2011 년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2011-0006791)

#### 참 고 문 헌

[1] Do M. N., Nguyen Q.H., Nguyen H. T., Kubacki D., and Patel S. J., "Immersive visual communication", *Signal Processing Magazine, IEEE*, vol. 28, pp. 58-66, Jan. 2011.

[2] Tanimoto M. Tehrani M. P., Fujii T., and Yendo T., "Free-viewpoint TV", *Signal Processing Magazine, IEEE*, vol. 28, pp. 67-76, Jan. 2011.

[3] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics", *Eurographics (State-of-the-Art Report)*, pp.119-134, Mar. 2009.

[4] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV", *Image commun.*, vol. 24, pp. 65-72, 2009.

[5] 김태준, 장은영, 허남호, 김진웅, 유지상, "다시점 카메라 모델의 기하학적 특성을 이용한 가상시점 영상 생성 기법", *한국통신학회논문지*, 제 34 권 12 호(통신이론 및 시스템), pp. 1154-1166, 12, 2009.