
컬러와 깊이 카메라를 이용한 시점 일치 영상 생성 기법

현지호* · 한재영* · 원종필** · 유지상***

Generation of an eye-contacted view using color and depth cameras

Jeeho Hyun* · Jaeyoung Han* · Jongpil Won** · Jisang Yoo***

이 논문은 2011년도 (주)LG전자 연구비를 지원받았음

요 약

일반적인 화상회의 시스템에서는 카메라의 위치가 모니터 중앙에 위치하여 있지 않기 때문에 화자간의 시점 불일치 현상이 발생하게 되며 따라서 대화 시 현장감이 떨어지는 문제가 발생한다. 본 논문에서는 컬러 카메라와 깊이 카메라(depth camera)를 이용하여 획득한 컬러 영상과 깊이 정보를 가지고 DIBR(Depth image based rendering) 기반의 시점 일치(eye-contacted) 영상을 생성하는 기법을 제안한다. 제안하는 기법에서는 생성된 시점 일치 영상에서 발생하는 홀(hole)을 인접 화소 값의 산술평균을 이용하여 효율적으로 채울 수 있는 홀 채움 기법과 객체 경계에서 발생하는 잡음을 효율적으로 제거하기 위해 깊이 정보에서 획득된 객체의 경계를 확장하여 보정하는 기법이 포함되어 있다. 제안하는 기법으로 생성된 시점 일치 영상은 매우 좋은 화질을 갖는다는 것을 실험을 통하여 보였다.

ABSTRACT

Generally, a camera isn't located at the center of display in a tele-presence system and it causes an incorrect eye contact between speakers which reduce the realistic feeling during the conversation. To solve this incorrect eye contact problem, we newly propose an intermediate view reconstruction algorithm using both a color camera and a depth camera and applying for the depth image based rendering (DIBR) algorithm. In the proposed algorithm, an efficient hole filling method using the arithmetic mean value of neighbor pixels and an efficient boundary noise removal method by expanding the edge region of depth image are included. We show that the generated eye-contacted image has good quality through experiments.

키워드

시점 일치, 깊이 영상 기반 렌더링, 깊이 카메라, 깊이, 화상회의, 홀 채움

Key word

Eye-contacted, DIBR, depth camera, depth, tele-presence, hole filling

* 준회원 : 광운대학교 전자공학과 석사과정
** 정회원 : LG전자기술원 연구원
*** 정회원 : 광운대학교 전자공학과 정교수 (교신저자)

접수일자 : 2012. 03. 29
심사완료일자 : 2012. 04. 20

I. 서 론

최근 디지털 카메라와 같은 영상 센서가 휴대폰, 노트북 등에도 기본 탑재되고 있고, 스마트폰이나 태블릿 PC 등의 단말기 성능이 좋아져 동영상을 손쉽게 다룰 수 있는 환경이 제공됨에 따라 일반인들의 영상 콘텐츠 제작 참여도 늘어나고 있다. 이러한 기술 발전으로 의사소통의 방법도 과거 문자나 소리에 의존하던 방법에서 서로의 얼굴을 마주보는 방향으로 바뀌고 있다.

최근 기존의 영상 센서와 함께 저가이면서 비교적 성능이 좋은 깊이 센서(depth sensor)가 개발되면서 영상 처리 분야는 또 다른 변화를 맞이하고 있다. 대표적인 깊이 센서로는 MESA-imaging의 SR4000과 PMD의 CamCube3.0 제품과 비교적 저가의 Kinect 카메라 등이 있다. 깊이 정보를 이용하면 음성 또는 동작 등의 사용자 환경(user interface, UI)에서 사용자 경험(user experience, UX)을 이용한 기능 등을 쉽게 구현할 수 있다[1,2]. 또한, 깊이 정보는 화상회의 시스템에서 현장감과 사실감을 높이기 위해서 사용될 수 있다.

일반적으로 카메라와 화자의 시선이 5도 이상 차이가 나는 경우 화상회의 시 화자의 시점이 일치되지 않는 현상이 발생한다[3]. 스마트폰과 같은 비교적 작은 디스플레이의 경우 화자 간 시점 불일치 현상을 느낄 수 없으나 컴퓨터 모니터 또는 TV를 이용하는 화상회의 시스템의 경우 카메라와 화자의 시선이 15도 이상의 차이가 생기게 된다. 이러한 경우 시점 불일치 현상이 발생하여 화상회의 시 현장감이 떨어지는 문제가 발생한다. 이 경우 현장감을 높이기 위해서는 화자 간의 시선의 방향이 일치된 시점 일치(eye-contacted) 영상을 생성하는 기술이 반드시 필요하다.

시점 일치 영상을 생성하는 기술에는 여러 방법이 있다. 초기에는 반거울(half-mirror)을 이용하여 디스플레이의 중앙에서 직접 영상을 획득하는 방식이 제안되었다. 이 방법은 시점 일치 영상을 직접 획득하기 때문에 매우 좋은 화질을 가지는 장점이 있으나 비용이 많이 들고 제작이 어려우며 화자가 바라보는 시선에 따라 카메라의 촬영 위치를 조절해야 하는 단점이 있다. 인간의 인지 감각을 이용하여 시점 일치 영상을 생성하는 방법도 제안되었다[4].

이 방법은 시점 불일치 영상 전체를 회전하여 시점

을 보상하면서 시점이 일치되었다고 느끼는 영상을 선택하는 방법이다.

이 방법은 완벽한 시점 일치 영상을 생성하는 것이 아니며 휴먼팩터(human factor)에 의존하기 때문에 정확하게 시점이 일치된 영상을 생성할 수 없다. 최근 들어 컴퓨터의 연산속도가 증가함에 따라 중간시점 영상 생성(intermediate view reconstruction, IVR) 방법을 응용하여 시점 일치 영상을 생성하는 방법도 제안되었다[5,6].

독일의 HHI(Heinrich Hertz Institute)는 2011년 CeBIT에서 가상 시점 일치 엔진(virtual eye contact engine)을 시연하였는데, 디스플레이 상단에 위치한 3대의 카메라로부터 획득된 영상을 가지고 깊이 정보를 획득하고 화자의 눈의 위치를 추적하여 알맞은 위치에서 가상의 시점 일치 영상을 생성하는 방법을 사용하였다. 컬러 카메라 3대는 눈의 위치 추적 및 깊이 정보를 획득하기 위한 스테레오 정합을 위해 사용된다. 이 방법의 경우 눈의 위치를 잘못 추적하거나 부정확한 깊이 정보를 획득하게 되면 잘못된 시점 일치 영상을 생성할 수 있다는 단점이 있다.

일반적으로 중간시점 영상 생성 기법으로 시점 일치 영상을 생성하는 경우에는 정확한 깊이 정보가 필요하다. 깊이 정보를 획득하는 방법은 능동형 깊이 센서(active depth sensor)나 수동형 깊이 센서(passive depth sensor) 또는 복합적 깊이 센서(hybrid depth sensor)를 이용하는 방법으로 구분된다. 수동형 깊이 센서 방식은 주어진 스테레오 영상을 이용한 스테레오 정합(stereo matching) 기법에 기반을 둔 방법이다[7-10]. 이 방법은 컬러 영상과 동일한 해상도를 가지는 깊이 정보를 획득할 수 있으나, 연산 복잡도가 매우 높으며 영상의 정렬화가 부정확할 경우 정확한 깊이 정보를 획득하는 것이 매우 어렵다는 단점이 있다. 이에 비해 능동형 깊이 센서 방식은 깊이 카메라(depth camera)와 같은 물리적 장치를 이용하여 깊이 정보를 직접 획득하는 방법이다[11]. 이 방식은 깊이 카메라 성능의 한계로 저해상도의 깊이 정보를 획득하는 단점이 있으나, 앞서 언급한 수동형 깊이 센서 방식에 비해 적은 연산으로 보다 정확한 깊이 정보를 획득할 수 있다는 장점이 있다. 따라서 본 논문에서는 깊이 카메라를 이용하여 직접 깊이 정보를 획득하는 방법을 이용한다.

먼저 획득된 깊이 영상을 캐니 연산자(Canny operator)를 이용하여 깊이 영상의 경계를 추출한다 [12,13]. 추출된 경계를 확장하여 보정할 영역을 생성한 후, 이 영역을 배경 깊이 값과 같은 값으로 채워 깊이 영상의 경계를 보정한다. 3차원 워핑(3D warping)을 이용하여 컬러 영상과 매핑되는 깊이 영상을 획득하고 가상의 시점 일치 영상을 생성하게 된다.

이 때 발생하는 홀(hole)은 홀 영역의 주변 화소 중에 홀이 아닌 화소들의 산술 평균값을 구하여 실시간으로 채우게 된다. 마지막으로 화상회회의 현상감 및 사실감을 향상시키기 위해 배경영역에 임의의 배경을 삽입함으로써 시점 최종적인 일치 영상을 생성하게 된다.

본 논문의 구성은 다음과 같다. 2장에서는 pin-hole 카메라 모델과 일반적인 카메라 보정 작업 및 3차원 워핑의 원리를 소개하고, 3장에서는 본 논문에서 제안하는 시점 일치 영상 생성 기법을 소개한다. 4장에서는 실험을 통하여 최종적으로 생성된 시점 일치된 영상의 결과를 보이고 5장에서 결론을 맺는다.

II. 본 론

2.1. 카메라 보정 작업

2.1.1. Pin-hole 카메라 모델

Pin-hole 카메라 모델은 내면을 검게 칠한 통의 한쪽 면에 작은 구멍을 내고 반대쪽 면에 필름을 위치하여 영상을 획득하는 장치이다. 그림 1은 pin-hole 카메라의 기하학적 구조를 보여준다.

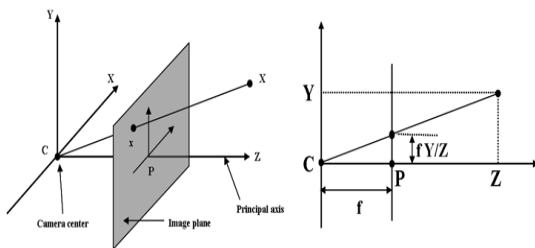


그림 1. Pin-hole 카메라의 기하학적 구조
Fig. 1 Geometrical model of pin-hole camera

실제 3차원 좌표(3D world coordinates)는 카메라의 구멍을 통하여 영상면(image plane)에 투영되고 비례 법에 의하여 카메라 중심과 영상면에 투영된 좌표까지의 거리가 산출된다. 이 거리를 이용하여 카메라의 정보에 의해 실제 3차원 좌표와 카메라의 영상에 투영되는 2차원 좌표와의 관계를 정의할 수 있다.

식 (1)은 투영행렬로서 pin-hole 카메라의 기하학적 구조를 이용하여 3차원 좌표와 2차원 좌표 사이의 관계를 수학적으로 표현한 것이다. 식 (1)을 이용하여 영상면 내에 투영된 2차원 좌표와 카메라의 내, 외부 파라미터를 이용하면 실제 3차원 좌표 X, Y, Z를 획득할 수 있다[14].

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

여기서 x, y는 영상면 내에 투영되는 2차원 좌표이고, K는 3x3 행렬로서 카메라 보정에 대한 내부 파라미터(intrinsic parameter), R은 3x3 행렬로서 카메라의 회전(rotation)에 대한 외부 파라미터(extrinsic parameter), T는 3x1 행렬로서 카메라의 병진(translation)에 대한 외부 파라미터이다. 그리고 X, Y, Z는 실제 3차원 좌표로서 각각 가로, 세로, 깊이에 관한 정보를 나타낸다.

2.1.2. 일반적인 카메라 보정 과정

카메라 보정(camera calibration) 과정은 크게 보정판을 사용하는 방법과 보정판을 사용하지 않는 방법으로 나눌 수 있다. 보정판을 사용하지 않는 방법은 카메라를 이동시킴으로써 얻은 정적인 영상들을 바탕으로 두 영상의 상관도를 이용하는 방법이다.

따라서 많은 유연성을 제공하지만 여러 파라미터들이 예측되어야 하기 때문에 신뢰 높은 결과를 얻을 수 없다. 이에 비해 보정판을 사용하는 방법은 일정한 패턴을 여러 장 촬영하여 사용하기 때문에 신뢰성이 매우 높다는 장점이 있다. 따라서 본 논문에서는 2차원 보정판을 사용하는 보정 방법을 이용한다. 그림 2는 카메라 보정에 사용되는 2차원/3차원 보정판이다 [15-20].

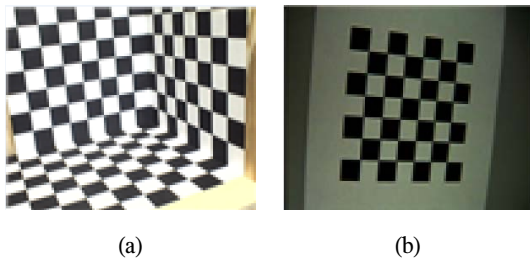


그림 2. 카메라 보정판
 (a) 3차원 보정판 (b) 2차원 보정판
 Fig. 2 The camera calibration panel
 (a) three-dimensional and (b) two-dimensional panel

2.1.3. 컬러 카메라와 깊이 카메라 간 보정 과정

일반적으로 카메라 간 보정 작업은 각 카메라에서 획득된 영상의 해상도가 같아야만 가능하다. 따라서 컬러 카메라에서 획득한 고해상도의 컬러 정보를 저해상도의 깊이 정보와 동일한 해상도로 변환하거나 저해상도의 깊이 영상을 고해상도의 컬러 영상과 동일한 해상도로 변환한 후 보정 작업을 진행하여야 한다. 본 논문에서 사용한 깊이 카메라는 MesaImaging 사의 SR4000 모델로 QCIF(176x144) 해상도의 깊이 영상을 초당 30 프레임 이상 획득할 수 있고 컬러 카메라는 Pointgrey 사의 Grasshopper 모델로 최대 2448x2048 해상도의 영상을 초당 15 프레임까지 획득 가능하다.

컬러 영상과 깊이 영상의 해상도를 동일하게 만들기 위하여 깊이 영상을 QCIF 해상도의 64배인 1408x1152의 해상도로 조정하고 컬러 영상은 1408x1152의 해상도로 획득한다. 이때 깊이 영상의 보간을 위해서 고등차수 보간법(bi-cubic interpolation)을 이용한다[12]. 깊이 정보의 해상도 조절 이후 두 카메라의 보정 작업을 통해 획득된 파라미터를 이용하여 3차원 워핑을 수행하여 획득된 깊이 정보를 컬러 영상으로 매핑하고, 최종적으로 컬러 영상과 매핑된 깊이 정보를 이용하여 시점이 일치된 영상을 생성하게 된다.

2.2. 3차원 워핑

3차원 워핑(3D warping) 과정은 크게 전 방향 워핑(forward warping)과 후 방향 워핑(backward warping)으로 구성된다. 일반적으로 스테레오 영상에서의 3차원 워핑 과정은 그림 3과 같이 진행된다. 전 방향 워핑의 경우 좌

시점 깊이를 영상을 우 시점 깊이 영상으로 변환하고, 생성된 우 시점 깊이 영상을 이용하여 후 방향 워핑을 통해 우 시점의 컬러 영상을 생성한다. 또한 이 과정에서 생성된 우 시점 영상에는 좌 시점 영상에 존재하지 않는 영역이 발생하며 화소 값이 존재하지 않아 홀(hole) 영역으로 남게 된다.

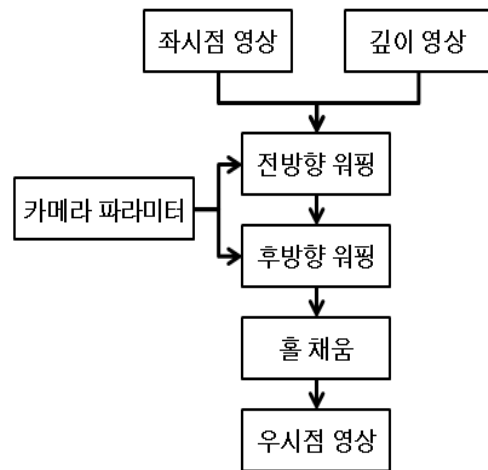


그림 3. DIBR 기법의 순서도
 Fig. 3 Block diagram of DIBR algorithm

III. 제안하는 기법

본 논문에서 제안하는 기법의 순서도를 그림 4에 보였다. 먼저 깊이 카메라의 해상도를 고등차수 보간법을 이용하여 컬러 카메라의 해상도와 같게 하고[12], 객체의 경계 부분을 보정한 후 전 방향과 후 방향 워핑을 적용하여 시점이 일치된 영상을 생성하며 이 때 발생하는 홀을 채우게 된다. 마지막으로 임의의 배경을 객체 이외 부분에 삽입하여 최종 결과를 얻게 된다[21]. 본 논문에서 깊이 카메라에서 획득한 깊이 영상의 해상도 확장 과정 전 방향 워핑과 후 방향 워핑 과정은 2장에서 설명한 기법을 이용한다.

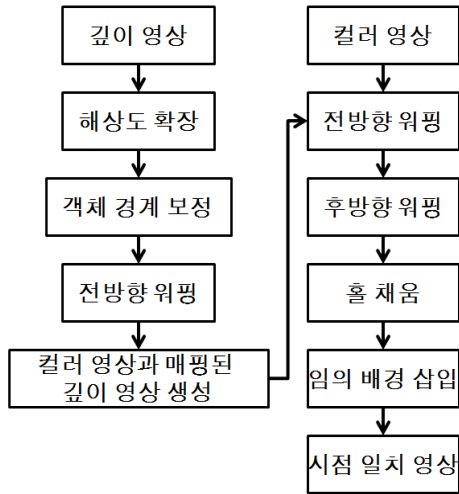


그림 4. 제안하는 기법의 순서도
Fig. 4 Block diagram of proposed method

3.1. 깊이 영상 경계 보정

영상 보간법을 이용하여 해상도가 확장된 깊이 영상은 그림 5와 같이 객체의 경계 부분에 흐려짐 (blurring) 현상이 발생하게 된다. 이러한 경계에서의 흐려짐 현상은 깊이 값의 오류를 발생시켜 3차원 워핑 시 경계 잡음을 발생시킨다[21]. 본 논문에서 제안하는 경계 잡음을 제거하기 위한 기법의 블록도를 그림 6에 나타내었다.

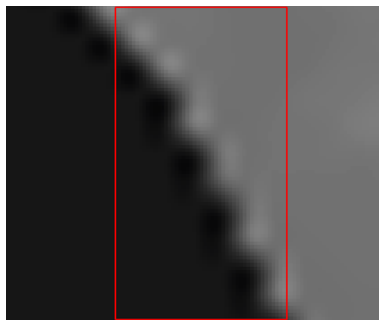


그림 5. 깊이 영상의 경계 흐려짐 현상
Fig. 5 Blur effect in boundary of depth image

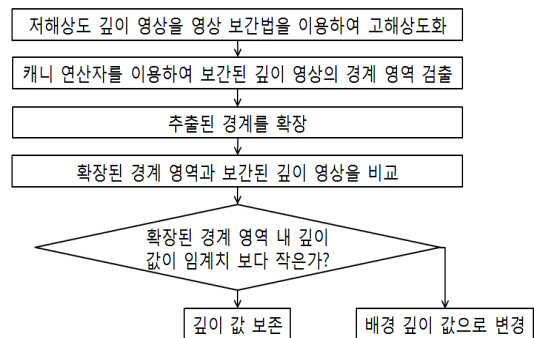


그림 6. 깊이 영상의 경계 잡음 제거 순서도
Fig. 6 Block diagram of boundary noise elimination for depth images

우선, 깊이 카메라에서 일정 거리 내에 존재하는 객체의 깊이 정보만 획득하고, 배경 영역의 깊이 값은 일정한 값으로 할당으로 한 후 깊이 영상을 컬러 영상과 동일한 해상도로 확장한다. 해상도가 높아진 깊이 영상의 경계를 Canny 연산자를 통해 추출한다. 그림 7(a)와 (b)에 보간된 깊이 영상과 추출된 깊이 영상의 경계를 각각 보았다. 경계 부분의 잡음을 제거하기 위해 경계 화소 중심으로 그림 8(a)와 같이 8-이웃(8-neighbor) 화소를 비교하여 배경인 부분을 연속해서 채워나가게 된다. 이 과정을 전체 경계 화소에 대하여 반복적으로 수행한다. 이렇게 하여 확장된 경계 영역과 깊이 영상을 다시 비교하여 해당되는 영역 내에 위치한 객체의 깊이 값을 임계 값과 비교한 후 임계 값보다 작을 경우는 배경의 깊이 값으로 대체함으로써 깊이 정보의 경계 부분에서 발생하는 흐려짐 현상도 보정한다.

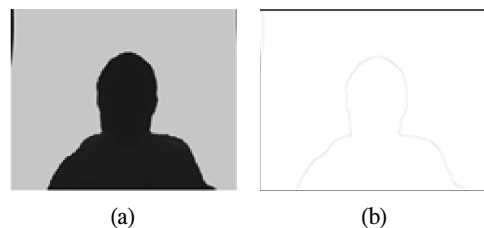


그림 7. (a) 고해상도로 변환된 깊이 정보
(b) 캐니 연산자를 이용하여 추출된 깊이 영상의 경계
Fig. 7 (a) Interpolated depth image with higher resolution
(b) edge map of interpolated depth image



그림 8. (a) 경계 영역 확장 과정
 (b) 확장된 깊이 정보의 경계 영역
 Fig. 8 (a) Procedure of edge expansion
 (b) expanded edge of depth image

그림 9는 제한한 깊이 영상의 경계 보정 기법의 성능을 실험하기 위해 획득된 깊이 영상을 3차원 워핑을 통해 시점 일치 영상으로 이동시킨 결과이다. 그림 9(a)는 경계 보정을 적용하지 않은 영상이며, 그림 9(b)는 경계 보정을 적용한 후의 깊이 영상이다. 그림 9(b)에서 보듯 경계 보정 기법을 적용하여 객체 외부에 발생하는 잡음을 효과적으로 제거할 수 있다는 것을 알 수 있다.

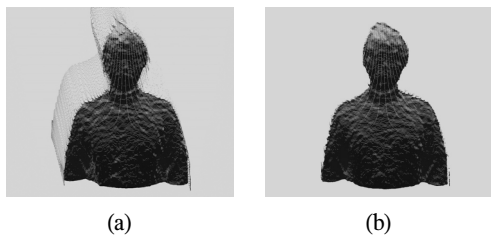


그림 9. 시점이 일치된 3차원 워핑의 결과
 (a) 경계 보정 전 (b) 후의 깊이 영상
 Fig. 9 Result of 3D warping (a) before and (b) after eliminating boundary noise of depth image

3.2. 실시간 처리가 가능한 홀 채움 기법

시점 일치 영상에서 홀은 깊이 영상의 잡음을 제거하고 3차원 워핑을 적용하는 과정에서 대응하는 원본 컬러 영상의 정보를 찾을 수 없어 발생한다. 그림 10은 전 방향 워핑과 후 방향 워핑을 이용해 생성된 시점 일치 영상에서 발생한 홀을 보여준다. 그림 10(a)는 홀이 발생한 시점 일치 영상이며, 그림 10(b)와 (c)는 그림 10(a)에서 발생한 홀 영역을 확대하여 나타낸 것이다.

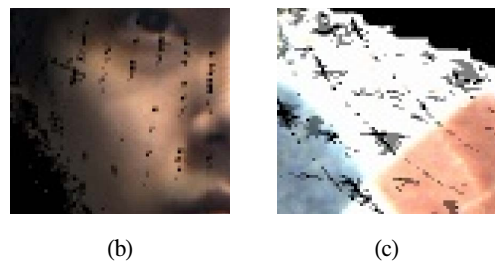
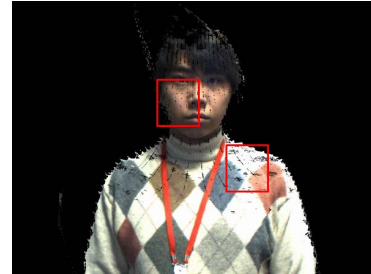


그림 10. (a) 홀이 발생한 시점 일치된 영상 (b) 얼굴 영역 홀 확대 영상 (c) 몸 영역 홀 확대 영상
 Fig. 10 (a) Generated eye-contacted view with holes (b) zoomed images of (a) around face region and (c) around body region

본 논문에서 제안하는 홀 채움 기법에서는 화소 단위로 홀 영역 여부를 판정한다. 먼저 영상 내에서 화소 값이 존재하지 않는 화소를 탐색하게 된다. 화소 값을 갖지 않는 화소의 경우, 이웃 화소들을 순차적으로 탐색하여 홀 영역 여부를 판정한다. 그림 11에는 홀 탐색 순서의 예시와 과정을 보였다.

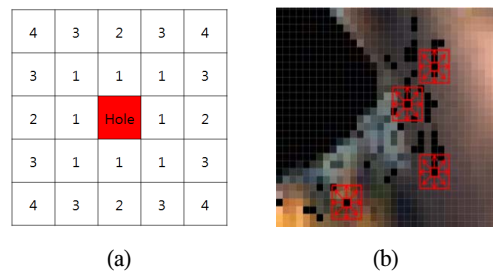


그림 11. (a) 홀 탐색 순서 (b) 홀 탐색 과정
 Fig. 11 (a) Order of hole searching (b) procedure of hole searching

그림 11(a)와 같이 8-이웃 화소를 탐색하고 모두 화소 값을 갖지 않는 경우 해당 화소를 배경으로 간주하고 탐색을 종료한다. 값이 존재하는 화소가 있는 경우, 유클리디언 거리(Euclidean distance)를 기반으로 가까운 순서로 탐색을 하여 식 (2)와 같이 각 단계마다 화소들의 R, G, B 색상 별 화소 값의 산술 평균을 구하고 다음 탐색 단계를 수행한다.

실시간 처리를 위해 탐색 단계의 화소가 모두 화소 값을 가지지 않거나 지정된 탐색 범위를 벗어나면 탐색을 종료하고 식 (3)과 같이 각 단계별 평균 화소 값과 거리의 가중치를 이용하여 해당 홀 화소를 채우게 된다.

$$R_k = \frac{1}{N} \sum_{n=0}^{N-1} r(n) \quad (2)$$

$$G_k = \frac{1}{N} \sum_{n=0}^{N-1} g(n)$$

$$B_k = \frac{1}{N} \sum_{n=0}^{N-1} b(n)$$

R_k, G_k, B_k, N is constant

$$R_{final} = \sum_{k=0}^{range-1} \left(\frac{1}{d_k} R_k \right) \quad (3)$$

$$G_{final} = \sum_{k=0}^{range-1} \left(\frac{1}{d_k} G_k \right)$$

$$B_{final} = \sum_{k=0}^{range-1} \left(\frac{1}{d_k} B_k \right)$$

여기서 N 은 각 탐색 단계별 홀이 아닌 화소의 개수를 나타내며, $r(n), g(n), b(n)$ 은 각 탐색 단계의 화소 값들의 배열, d 는 홀 화소와 이웃 화소 간의 유클리디언 거리, k 는 각 탐색 단계의 순서를 나타낸다.

3.3. 사실감을 높이기 위한 배경 삽입

배경 부분이 없으면 화상회의 시스템에서 사용자의 사실감이나 현장감이 떨어지게 된다. 본 논문에서는 이러한 문제를 해결하기 위하여 객체 이외의 영역에 임의의 배경 영상을 삽입한다. 그림 12(b)는 생성된 시점 일치 영상에 임의의 배경을 삽입한 결과이다.



그림 12. (a) 임의배경 삽입 전 영상 (b) 임의배경 삽입 후 영상

Fig. 12 (a) With and (b) without a background image

IV. 실험 조건 및 결과

4.1. 시점 일치된 영상의 생성

그림 13은 실험에 사용된 카메라 환경이다. 깊이 카메라와 컬러 카메라를 53인치 모니터 하단에 일렬로 배치하여 시점 일치 영상을 생성한다. 먼저 제안하는 기법으로 경계 잡음을 제거한 후 중간값 필터(median filter), 인페인팅(in-painting) 기법과 제안하는 홀 채움 기법을 각각 적용하여 그 결과를 그림 14(b), (c)와 (d)에 보였다 [22]. 그림 14(b)에서 보면 중간값 필터를 사용하는 경우 전체적인 시점 일치 영상에 흐려짐 효과가 심하게 나타나고 채워지지 않는 홀이 남아있는 것을 확인할 수 있으며, 그림 14(c)의 인페인팅 기법의 경우는 홀 채움 효과에 비해 연산 복잡도가 높아 실시간 처리가 어렵다는 것을 알 수 있었다. 이에 비해 제안하는 기법의 경우는 그림 14(d)와 같이 흐려짐 현상도 많이 제거되고, 대부분의 홀 영역도 잘 채워지는 것을 확인할 수 있다.

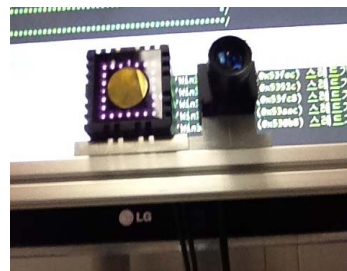


그림 13. 실험 환경
Fig. 13 Test environment

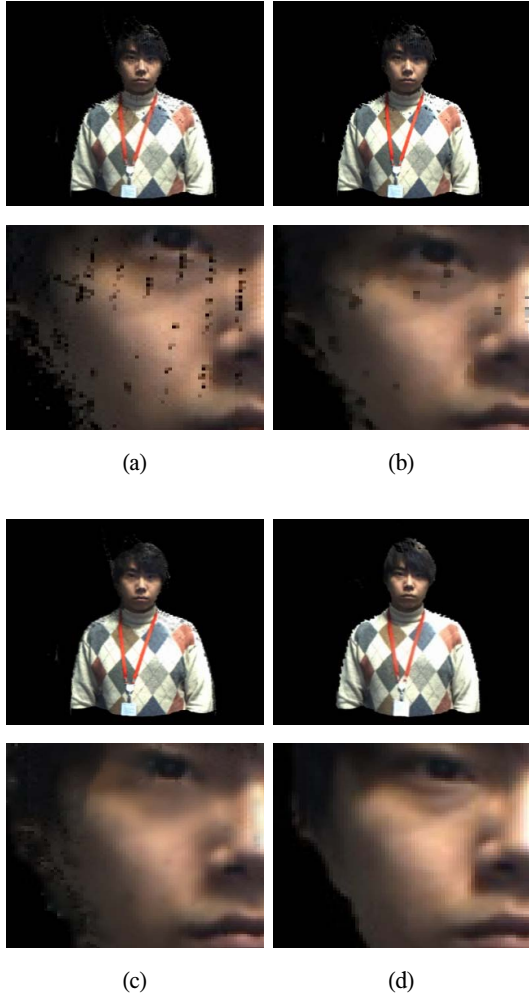


그림 14. 홀 채움 기법 결과 (a) 원본 영상 (b) 3x3 중간값 필터를 이용한 기법 (c) 인페인팅 기법 (d) 제안한 기법 (상 : 전체 영상 / 하 : 얼굴 영역 확대 영상)
 Fig. 14 Result of hole filling (a) original image (b) algorithm using 3x3 median filter (c) an in-painting algorithm (d) the proposed algorithm

그림 15(a)는 컬러 카메라에서 획득된 원본 컬러 영상이며 그림 15(b)는 깊이 카메라에서 획득된 원본 깊이 영상, 그림 15(c)는 본 논문에서 생성된 시점 일치 영상에 배경을 삽입하여 최종적으로 생성된 시점 일치 영상을 보여 준다.

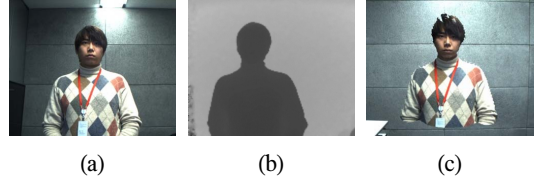


그림 15. (a) 원본 컬러 영상 (b) 원본 깊이 정보 (c) 최종적으로 생성된 시점 일치된 영상
 Fig. 15 (a) Original color image (b) interpolated depth information (c) final eye-contacted view

4.2. 제안하는 기법의 성능 평가

제안된 기법이 화상회의 시스템에 적용하려면 생성된 시점 일치 영상의 품질과 더불어 실시간 처리를 위한 연산 속도도 매우 중요한 요소이다. 표 1에서는 실험에 사용된 PC 사양을 보여준다. 표 2는 제안한 기법을 적용하여 시점 일치 영상을 생성하는데 소요되는 시간을 단계별로 나타내었다.

표 1. 실험에 사용된 PC 사양
 Table. 1 Test environment

CPU	Intel i7-2600 CPU @ 3.40GHz
RAM	4.00GB
OS	Windows 7 (x86)
GPU	GeForce GTX 570

표 2. 제안한 기법의 단계별 처리 시간
 Table. 2 Processing time of each process

과 정	수행시간 (msec)	비 율 (%)
컬러 카메라와 깊이 카메라의 영상 획득	4	3.65
깊이 영상의 해상도 확장	20	9.13
깊이 정보 경계 보정	71	32.42
전 방향 워핑을 통한 컬러 영상에 깊이 영상 매핑	18	8.22
3차원 워핑을 통해 시점 일치 영상 생성	21	9.60
홀 채움	55	25.11
입의 배경 삽입	16	7.30
결과 영상 출력	10	4.57
총 수행 시간	219	100

본 논문에서 깊이 영상의 해상도 확장은 OpenCV 라이브러리를 이용하여 GPU로 병렬 처리하였으며, 전 방향 워핑으로 컬러 영상을 깊이 영상에 매핑하는 과정과 시점 일치 영상으로 3차원 워핑을 하는 과정 또한 GPU를 이용한 병렬 연산을 통해 속도를 향상시켰다. 향후 제안한 깊이 영상의 경계 보정 기법과 홀 채움 기법을 GPU를 이용한 병렬 처리로 구현한다면 초당 15프레임 이상의 실시간 처리가 가능할 것으로 판단된다.

화질 평가를 위해 주관적인 화질평가도 수행하였다. 평가 방법으로는 ITU-R 권고로서 규정되어 있는 이중자극 연속품질척도법인 DSCQS(Double Stimulus Continuous Quality Scale) 방법을 이용하였다[23]. 제안한 기법의 성능을 비교하기 위해 중간값 필터를 이용한 기법과 인페인팅 기법을 이용하여 홀 채움을 수행한 결과를 비교하였다. 총 15명이 평가를 했으며 100프레임에 각각의 기법을 적용한 결과를 이용하였다. 각각의 기법을 적용한 결과를 무작위로 순서를 설정하고 순서를 바꿔가며 총 3회의 평가를 실시하였다. 표 3의 평가 결과와 같이 제안한 기법이 다른 기법에 비해 성능이 우수하다는 것을 알 수 있다.

표 3. 주관적 화질 평가표
Table. 3 DSCQS table

	중간값 필터를 이용한 기법	인페인팅 기법	제안한 기법
점수	Poor(39)	Fair(64)	Fair(68)

V. 결 론

본 논문에서는 화상회의 시스템에 적용하여 화자의 사실감 및 현장감을 높이고자 고화질의 시점 일치된 영상을 생성하기 위한 기법을 제안하였다. 제안된 기법에서는 먼저 깊이 카메라에서 획득된 깊이 영상을 고등차수 보간법을 통해 컬러 영상과 동일한 해상도로 보간한다. 이렇게 보간된 깊이 영상에 캐니 연산자를 적용하여 객체의 경계를 구하고, 구한 경계를 확장하여 경계 부분에 발생하는 흐려짐 현상을 보정하였다. 또한 이웃 화소를 단계별로 탐색하여 홀을 판정하고 단계별 산술

평균값과 홀 화소와 이웃 화소의 유클리디언 거리를 이용하여 홀 채움을 수행하였다. 화상회의 시스템의 사실감과 현장감을 개선하기 위하여 생성된 시점 일치 영상에 배경을 삽입하였다. 실험 결과를 통해 새로이 제안하는 기법의 성능이 기존 기법들 보다 우수하다는 것을 보였다. 주관적 화질평가도 실시하였는데 제안된 기법으로 생성된 시점 일치 영상의 화질이 우수함을 확인할 수 있었다.

향후 3차원 구조의 카메라 배열을 통한 정확한 깊이 정보를 획득하는 방법 등의 연구를 통해 더욱 성능이 좋은 고화질의 시점 일치 영상을 구현할 수 있을 것으로 기대한다.

감사의 글

본 연구는 2011년도 (주)LG전자 산학연구 과제의 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

참고문헌

- [1] M. N. Do, Q. H. Nguyen, H. T. Nguyen, D. Kubacki, and S. J. Patel, "Immersive visual communication," *IEEE Signal Processing Magazine*, vol.28, pp.58-66, Jan. 2011.
- [2] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free-viewpoint TV," *IEEE Signal Processing Magazine*, vol.28, pp.67-76, Jan. 2011.
- [3] R. Stokes, "Human factors and appearance design considerations of the mod II PICTUREPHONE® station set," *IEEE Transactions on Communication Technology*, vol.17, pp.318-323, Apr. 1969.
- [4] F. Solina, and R. Ravnik, "Fixing missing eye-contact in video conferencing systems," *ITI 2011 33rd International Conference on*, pp.233-236, June 2011.

- [5] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV," *Signal Processing: Image Communication*, vol.24, pp.65-72, Jan. 2009.
- [6] Q. H. Nguyen, M. N. Do, and S. J. Patel, "Depth image-based rendering from multiple cameras with 3D propagation algorithm," *Proceedings of the 2nd International Conference on Immersive Telecommunications*, pp.1-6, May 2009.
- [7] 서용석, 유지상, "영상 분할을 이용한 다이내믹 프로그래밍 기반의 스테레오 정합," 한국통신학회 논문지, 제35권, 제8호, pp.680-688, 2010.
- [8] 김태준, 유지상, "영상의 컬러 정보를 이용한 계층적 스테레오 정합," 한국통신학회 논문지, 제34권, 제3호, pp.279-287, 2009.
- [9] 최미남, 윤정환, 유지상, "다시점 동영상의 중간 시점영상 생성을 위한 변이 예측 기법," 한국방송공학회 논문지, 제13권, 제6호, pp.915-929, 2008.
- [10] 김태준, 장은영, 허남호, 김진웅, 유지상, "다시점 카메라 모델의 기하학적 특성을 이용한 가상시점 영상 생성 기법," 한국통신학회 논문지, 제34권, 제12호, pp.1154-1166, 2009.
- [11] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics," *Eurographics State of the Art Reports*, vol.1, Mar. 2009.
- [12] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.29, pp.1153-1160, Dec. 1981.
- [13] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.6, pp.679-698, Nov. 1986.
- [14] O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, MIT Press, 1993.
- [15] H. R. Bakken, G. B. Eilertsen, U. G. Matus, and H. J. Nilsen, "Semi-automatic camera calibration using coplanar control points," *Proceedings of NIK*, 2009.
- [16] T. Svoboda, D. Martinec, and T. Pajdla, "A convenient multi camera self-calibration for virtual environments," *Presence: Teleoperators & Virtual Environments*, vol.14, pp.407-422, Aug. 2005.
- [17] Y. J. Bouguet, "Camera calibration toolbox for matlab," *Computational Vision at CALTECH*, 2010.
- [18] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, pp.1330-1334, Nov. 2000.
- [19] Z. Zhang, "Camera calibration with one-dimensional objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.26, pp.892-899, July 2004.
- [20] P. F. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.1, June 1999.
- [21] 현지호, 한재영, 유지상, "이종 카메라를 이용한 eye-contacted 영상 생성 기법," 한국방송공학회 추계학술대회, 2011.
- [22] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of Graphics Tools*, vol.9, pp.25-36, 2004.
- [23] ITU-R Recommendation BT.500-11, *Methodology for subjective assessment of the quality of television picture*, Mar. 2011.

저자소개

현지호(Jeeho Hyun)



2011년 2월 : 광운대학교
전자공학과 학사
2011년 3월 ~ 현재 : 광운대학교
전자공학과 석사과정

※ 관심분야 : 디지털 신호처리, 깊이 영상 획득



한재영(Jaeyoung Han)

2012년 2월 : 광운대학교
전자공학과 학사
2012년 3월 ~ 현재 : 광운대학교
전자공학과 석사과정

※ 관심분야 : 영상 신호처리, 스테레오 매칭



원종필(Jongpil Won)

2003년 2월 : 한국항공대학교
통신정보공학과 학사
2005년 2월 : 한국항공대학교
정보통신공학과 석사

2012년 3월 ~ 현재 : LG전자기술원 연구원

※ 관심분야 : Embedded System 개발, 영상 신호처리



유지상(Jisang Yoo)

1985년 : 서울대학교
전자공학과 학사
1987년 : 서울대학교
전자공학과 석사

1993년 : Purdue Univ. EE, Ph.D.

1997년 ~ 현재 : 광운대학교 전자공학과 교수

※ 관심분야 : 3D 입체영상처리/압축, 웨이블릿 기반
영상처리, 비선형 디지털 신호처리