DBpia

고해상도 스테레오 카메라와 저해상도 깊이 카메라를 이용한 다시점 영상 생성

정회원 이천*, 송혁**, 최병호**, 종신회원 호요성*°

Multi-view Generation using High Resolution Stereoscopic Cameras and a Low Resolution Time-of-Flight Camera

Cheon Lee*, Hyok Song**, Byeongho Choi** *Regular Members*, Yo-Sung Ho*° *Lifelong Member*

요 약

최근 자연스러운 3차원 영상의 재현을 위하여 깊이영상을 이용한 영상합성 방법이 널리 이용되고 있다. 깊이영상은 시청자의 눈에 보이지는 않지만 합성영상의 화질을 결정하는 중요한 정보이므로 정확한 깊이영상을 획득하는 것이 중요하다. 특히 적외선 센서를 이용한 깊이 카메라(time-of-flight camera)는 보다 정확한 깊이영상을 획득하는데 이용되고 있다. 깊이 카메라는 스테레오 정합(stereo matching)에 비해 정확하고 실시간으로 깊이정보를 추출할 수 있지만, 제공되는 해상도가 너무 낮다는 단점이 있다. 본 논문에서는 단시점의 깊이영상을 두 시점의 깊이영상으로 확장하고, 이를 이용하여 여러 시점의 중간영상을 생성하는 시스템을 제안한다. 특히 복잡도를 낮춰 빠른 속도로 다시점 영상을 생성하는 시스템을 제안한다. 고해상도의 컬러 영상을 획득하기 위하여 두 대의 컬러 카메라를 설치하고 중간에 깊이 카메라를 획득한다. 그리고 깊이 카메라에서 획득한 깊이영상을 3차원 워핑을 이용하여 양쪽의 컬러 카메라의 위치로 시점 이동한다. 깊이영상과 컬러영상간의 객체 불일치 문제는 깊이값의 신뢰도를 기반으로 한 조인트 양방향 필터(joint bilateral filter)를 이용하여 보정한다. 이러한 과정을 통해 얻은 깊이영상은 다시점 영상 합성 방법을 이용하여 다시점 영상을 획득한다. 이와 같은 과정은 다중 스레드를 이용하여 빠르게 처리할 수 있도록 구현했다. 실험을 통해 두 시점의 컬러영상과 두 시점의 깊이영상이 실시간으로 획득했고, 약 7 fps의 프레임율로 10시점의 중간시점을 동시에 생성했다.

Key Words : 병합형 스테레오 카메라, 3차원 카메라, 3차원 비디오 시스템, 깊이영상 생성, 3차원 영상 보정, 다시점 영상 생성

ABSTRACT

Recently, the virtual view generation method using depth data is employed to support the advanced stereoscopic and auto-stereoscopic displays. Although depth data is invisible to user at 3D video rendering, its accuracy is very important since it determines the quality of generated virtual view image. Many works are related to such depth enhancement exploiting a time-of-flight (TOF) camera. In this paper, we propose a fast 3D scene capturing system using one TOF camera at center and two high-resolution cameras at both sides. Since we need two depth data for both color cameras, we obtain two views' depth data from the center using the 3D warping technique. Holes in warped depth maps are filled by referring to the surrounded background depth

* 광주과학기술원 정보통신공학과 ({leecheon, hoyo}@gist.ac.kr) (°: 교신저자)
** 전자부품연구원 멀티미디어IP ({hsong, bhchoi}@keti.re.kr)
논문번호 : KICS2012-02-055, 접수일자 : 2012년 2월 9일, 최종논문접수일자 : 2012년 4월 23일

www.dbpia.co.kr

values. In order to reduce mismatches of object boundaries between the depth and color images, we used the joint bilateral filter on the warped depth data. Finally, using two color images and depth maps, we generated 10 additional intermediate images. To realize fast capturing system, we implemented the proposed system using multi-threading technique. Experimental results show that the proposed capturing system captured two viewpoints' color and depth videos in real-time and generated 10 additional views at 7 fps.

Ⅰ. 서 론

3차원 영상 기술은 최근 들어 차세대 멀티미디어 서비스를 이끌 핵심 기술로 주목을 받고 있다. 하지만 3차원 영상에 대한 연구는 1838년에 영국의 과학자인 찰스 휫스톤(Charles Wheatstone)이 입체시(stereopsis)에 대해 처음으로 설명한 것으로 거슬러 올라간다[1]. 이후 입체경(stereoscope)과 스테레오그램(stereogram) 등의 발명으로 꾸준히 연구되고 있다. 하지만 입체영상을 볼 때 느낄 수 있는 어지러움이나 구토 증세와 같은 안전상의 문제와, 카메라를 두 대 이상으로 이용하는 문제 때문에 발생할 수 있는 기술적/제정적인 문제로 인해 대중에 친숙하도록 보급되진 않았다. 하지만 최근 정보통신 기술의 비약적인 발전으로 인해 이와 같은 문제가 많이 해소되고 있고, 보다 실감 있는 컨텐츠에 대한 수요가 증가함에 따라 3차원 영상 기술이 다시 주목 받고 있다. 특히 두 시점만을 이용하여 입체영상을 제공하는 기존의 스테레오 영상의 한계에서 벗어나, 다시점 영상을 이용하여 보다 안전하고 생동감 있는 깊이영상이 활발하게 연구되고 있다[2].

깊이영상을 기반으로 한 다시점 입체영상 기술은 다시점영상과 다시점 깊이영상을 이용한 영상 형식(multi-view video-plus-depth, MVD)을 이용한다[3]. 깊이영상은 영상의 각 화소별 3차원 거리를 나타내는 정보이므로, 각 화소를 원래의 3차원 위치로 재사영할 수 있다. 그리고 이를 이용하여 임의의 카메라 시점으로 재사영하면 임의의 시점으로 영상 합성이 가능하다. 이러한 방법을 깊이기반 영상렌더링(depth-image-based rendering, DIBR)이라 한다[4]. 이 DIBR 기술을 이용하면 다시점 영상의 시점 수를 줄일 수 있는 장점이 있다. 자연스런 입체영상 재현을 위해서는 많은 시점의 영상이 필요로 하지만 시점 수에 따라 데이터의 양도 비례하여 증가하는 단점이 있다. 하지만, 깊이영상을 이용한 시점 합성 방법을 이용하면 일부 시점과 깊이영상만으로 다시점의 영상 생성이 가능하므로 전송 효율 면에서 큰 이득을 얻을 수 있다.

이러한 문제를 해결하기 위하여 2005년부터 국제 동영상 압축 표준화 그룹인 MPEG(Moving Picture Experts Group)과 ITU-T가 Joint-Video-Team (JVT)를 구성하여 다시점 비디오 부호화(multi-view video coding, MVC) 기술을 개발했다[5]. 이 기술은 다시점 컬러 영상을 시점간 상관도를 이용하여 부호화하는 기술이다. 이후 2007년부터는 깊이영상의 부호화까지 수행하는 3차원 비디오 부호화 (3D video coding)에 대한 연구가 진행되고 있다[6].

이와 같이, 미래의 3차원 비디오 기술은 깊이영상을 중요한 보조 데이터로 이용한다. 하지만, 정확한 깊이영상을 예측하거나 획득하는 것이 쉽지 않다. 두 시점 이상의 컬러 영상간의 시차를 이용하여 깊이값을 추정하는 스테레오 정합(stereo matching)은 하드웨어가 필요로 하지 않고 어떠한 조건에서 촬영한 영상이라도 추정이 가능하다는 장점이 있지만, 깊이값의 정확도가 항상 보장되지 않고 복잡도가 매우 높다는 단점이 있다[7]. 이에 반해 하드웨어를 이용한 깊이영상 획득 방법은 비용과 부피가 크지만[8-9], 비교적 정확한 깊이영상을 빠른 속도로 획득할 수 있다는 단점이 있다.

최근 널리 이용되고 있는 time-of-flight(TOF) 방식의 깊이 카메라는 실시간으로 비교적 정확한 깊이영상을 획득할 수 있는 장점이 있다. 하지만, 지원 해상도가 매우 작아(QCIF) 고화질의 깊이영상을 획득하는데 어려움이 있다. 이러한 단점을 보완하고 고해상도의 깊이영상을 획득하기 위한 다양한 연구들이 진행되고 있다. Lindner[10]와 Huhle[11]은 한 대의 컬러 카메라와 한 대의 깊이 카메라를 이용하여 3차원 모델을 획득했다. 그리고 Kim[12]와 Hahne[13]는 고해상도의 스테레오 카메라와 한 대의 깊이 카메라를 이용하여 불안정한 깊이영상을 개선했다. Zhu[14]와 Lee[15]의 방법은 정확한 깊이값 측정을 위하여 깊이값을 보정하는 방법을 제안했다.

본 논문에서는 이러한 연구의 일환으로 두 대의 고해상도 컬러 카메라와 한 대의 깊이 카메라를 이용하여 빠르고 효율적으로 두 컬러 카메라 시점의 깊이영

www.dbpia.co.kr

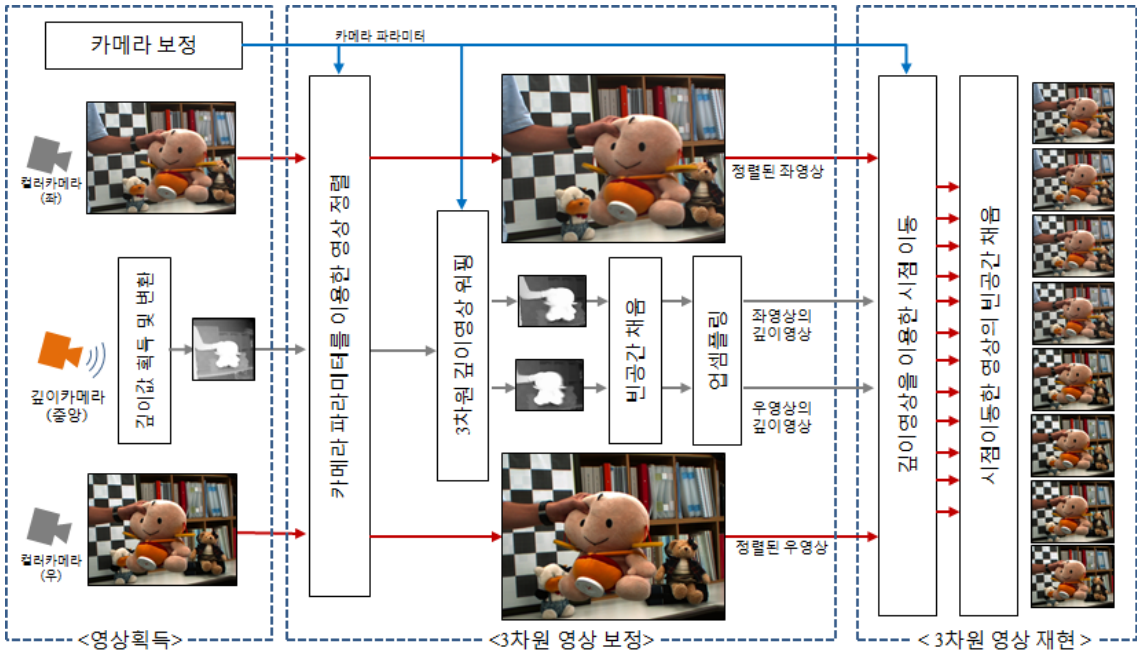


그림 1. 두 대의 컬러 카메라와 한 대의 깊이 카메라를 이용한 다시점 비디오 획득방법
Fig. 1. Overall structure of multi-view generation using one TOF camera and two color cameras

상을 획득하고, 다시점 비디오까지 생성하는 방법을 제안한다. 특히, 시점 이동한 깊이영상의 깊이값 오류를 줄이기 위하여 컬러와 깊이값을 동시에 고려한 깊이영상 오류 수정방법을 이용했다.

본 논문의 구성은 다음과 같다. 이어지는 장에서는 본 연구를 통해 구축한 카메라 시스템을 소개하고, 이어서 제 3장에서는 두 시점의 깊이영상을 획득하기 위한 깊이영상 처리 방법을 제안한다. 제 4장에서는 제안한 방법을 빠른 속도로 처리하기 위한 구현 방법과 실험 결과를 보인다. 그리고 제 5장에서는 결론을 맺는다.

Ⅱ. 깊이 카메라를 이용한 3D 카메라 시스템

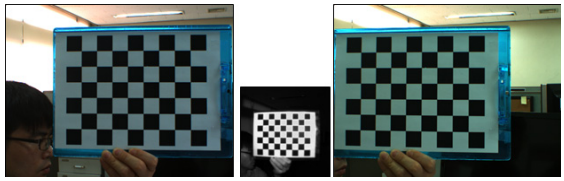
2.1 3차원 영상 획득을 위한 카메라 시스템

그림 1은 제안하는 3차원 영상 획득 시스템을 보인 것이다. 일반적인 스테레오 카메라의 카메라간 거리 인간의 양안의 거리가 약 60-70mm인 것을 고려하여 평균적으로 약 65mm의 간격을 유지한다. 하지만, 제안하는 카메라 시스템은 중간에 깊이 카메라를 설치해야 하므로 두 컬러 카메라의 거리가 약 10.5cm로 고정되어 있다. 제안하는 카메라 시스템을 이용하여 다시점 영상을 획득하는 과정을 총 세 단계로 구분할 수 있다. 먼저 두 대의 컬러 카메라로 고해상도의 스테레오 영상을 획득하고, 깊이 카메라로 저해상도의 깊이영상을 획득하는 '영상 획득단'과 각 영상을 다시점 영상 합성에 이용할 수 있도록 컬러 및 깊이 데이터를 변형/처리 하는 '3차원 영상 보정단', 최종 획득한 컬러영상과 깊이영상을 이용하여 다시점 영상을 생성하는 '3차원 영상 재현단'으로 나눈다.

2.2 카메라 변수 추정

두 대 이상의 카메라를 사용하여 다시점 영상을 생성하기 위해서는 카메라간의 위치적 상관관계를 정확하게 정의해야 한다. 특히, 깊이정보를 이용하여 각 시점별로 화소간 상관관계를 정의할 때는 각 카메라의 정확한 변수를 이용한다. 다시점 카메라의 카메라 변수를 추정하는 방법을 카메라 보정(camera calibration)이라고 한다[16]. 카메라 변수는 카메라 내부 구조를 나타내는 내부 변수(intrinsic parameter)와 카메라의 방향과 위치를 나타내는 외부 변수(extrinsic parameter)로 구성된다. 카메라 내부 변수는 초점거리, 광축의 위치 등의 변수로 구성된 3x3 행렬이다. 카메라 외부 변수는 카메라가 가리키는 방향을 나타내는 3x3 방향 행렬(rotation matrix)과 카메라의 3차원 위치를 가리키는 3x1 천이 벡터(translation vector)로 이루어진다. 단시점 카메라의 카메라 보정과는 달리 다시점 카메라의 카메라 보정은 카메라간 위치관계가 명확해야 하므로 보다 세밀한 추정 과정이 필요하다. 그림 2는 다시점 카메라의 카메라 변수 추정을 위한 패턴 영상을 촬영한 것이다.

www.dbpia.co.kr



(a) 왼쪽 카메라 (b) ToF 카메라 (c) 오른쪽 카메라

그림 2. 카메라 변수 추정을 위한 패턴 영상 획득
Fig. 2. Pattern images for camera calibration

이렇게 추정한 카메라 변수는 3차원 공간상의 점과 영상에 찍힌 화소의 관계를 설명한다. 예를 들어, 3차원 공간상의 한 점 **Mw** = $[X, Y, Z]^T$이 한 카메라로 투영되어 영상의 한 화소 **m**=$[u\ v]^T$로 사영되었을 경우, 이 두 점의 상관관계를 다음의 식 (1)로 표현할 수 있다.

$$\tilde{\mathbf{m}} = \mathbf{A} \cdot \mathbf{R} \cdot \mathbf{M}_w + \mathbf{A} \cdot \mathbf{t} \qquad (1)$$

이러한 기본 개념을 두 컬러 카메라와 중앙의 깊이 카메라에 적용하면, 3차원 공간상의 한 점 **Mw**은 각 카메라의 영상에 다음과 같이 사영된다.

$$\begin{cases} \tilde{\mathbf{m}}_L = \mathbf{A}_L \cdot \mathbf{R}_L \cdot \mathbf{M}_w + \mathbf{A}_L \cdot \mathbf{t}_L \\ \tilde{\mathbf{m}}_R = \mathbf{A}_R \cdot \mathbf{R}_R \cdot \mathbf{M}_w + \mathbf{A}_R \cdot \mathbf{t}_R \\ \tilde{\mathbf{m}}_D = \mathbf{A}_D \cdot \mathbf{R}_D \cdot \mathbf{M}_w + \mathbf{A}_D \cdot \mathbf{t}_D \end{cases} \qquad (2)$$

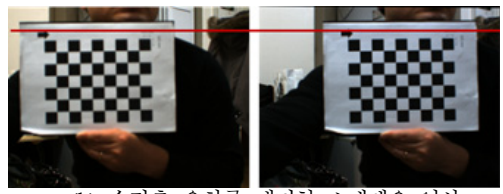
여기서 *L*, *R*, *D*는 왼쪽, 오른쪽 컬러 카메라와 중앙의 깊이 카메라를 표현한 것이다.

2.3 다시점 영상 정렬

카메라 변수는 장면의 각 점이 카메라로 투영되어 영상으로 기록될 때 각 점과 영상의 각 화소간의 위치 관계를 설명한다. 이러한 정보를 이용하면 다시점 카메라로 획득할 때 발생할 수 있는 기하학적 오류를 줄일 수 있다. 특히, 3차원 영상은 인간이 사물을 통해 입체감을 인지하는 조건과 최대한 비슷하게 처리를 한 후 양안으로 주사해야 하므로, 입체영상간의 불필요한 요소는 미리 제거해야 한다. 그 중 가장 큰 기하학적 오류는 시점간 높이차이와 방향 차이이다. 시점간 방향차이가 있는 3차원 영상의 경우는 서로 상응하는 화소가 서로 다른 높이에 존재하기 때문에 입체영상을 관람할 때 인식의 혼동이 유발되어 어지러움을 느낄 수 있다. 또한 두 대의 카메라가 가리키는 방향이 서로 다를 경우도 사용자에게 어지러움을 유발할 수 있다. 이러한 기하학적 오류를 줄이고, 마치 1차원 평행형 카메라로 촬영한 것처럼 영상을 보정하는 기술을 영상 정렬(image rectification)이라고 한다[17]. 일반적으로 영상 정렬은 스테레오스코픽 영상에 적용하여 시점간 기하학적 오류를 제거하지만, 본 시스템에서 이용하는 카메라는 세 대이고, 영상의 해상도가 서로 다르므로 적응적인 영상 정렬을 수행해야 한다. 영상 정렬을 수행하면 공통된 방향벡터와 거리가 일정한 천이벡터를 얻을 수 있고, 통일된 컬러 카메라의 내부 변수를 얻을 수 있다. 영상 정렬된 세 카메라 변수의 관계를 수식으로 표현하면 다음과 같이 나타낼 수 있다.

$$\begin{cases} \mathbf{A}'_L = \mathbf{A}'_R \neq \mathbf{A}_D \\ \mathbf{R}'_L = \mathbf{R}'_R = \mathbf{R}'_D \\ \mathbf{t}'_L \neq \mathbf{t}'_R \neq \mathbf{t}'_D \end{cases} \qquad (3)$$

(a) 수직축으로 차이가 있는 원본 스테레오 영상



(b) 수직축 오차를 제거한 스테레오 영상

그림 3. 영상 정렬 결과
Fig. 3. Result of image rectification

방향 행렬은 세 카메라가 통일된다. 이로써 세 카메라가 바라보는 방향이 같아진다. 그리고 좌/우의 컬러 카메라들은 내부 변수가 통일되고, 왼쪽의 컬러 카메라와 중앙의 깊이 카메라의 거리는 중앙의 깊이 카메라와 오른쪽의 컬러 카메라와 거리가 같아진다. 식에서 B는 두 컬러 카메라간의 거리를 나타내는 상수이다. 이때 중요한 것은 두 컬러 카메라는 동일한 초점거리를 갖지만, 깊이 카메라는 컬러 카메라와 다른 초점 거리를 갖는다.

2.4 깊이카메라를 이용한 깊이영상 변환

깊이 카메라에서 획득한 깊이값은 실수로 획득되는데, 이를 양자화하여 회색톤의 영상으로 저장한다. 이때 저장된 깊이 영상은 다시점 영상 합성에 이용되므로 컬러 카메라의 위치관계를 이용하여 각 깊이값을 양자화 해야 한다. 먼저 두 시점간의 시차로 인해 발

www.dbpia.co.kr

생하는 변이값(disparity)는 다음과 같이 정의된다.

$$d = \frac{F \cdot B}{Z} \tag{4}$$

여기서 Z는 깊이 카메라에서 획득한 실수단위의 깊이값을 나타내고, F는 영상정렬된 컬러 카메라의 초점거리(focal length)를 나타내고, B는 앞에서 설명한 컬러 카메라의 거리를 나타낸 것이다. 식 (4)를 이용하여 깊이값에서 변이값으로 변환한 값은 다음의 식 (5)를 이용하여 깊이영상으로 저장한다.

$$v = (2^N - 1) \cdot \left(\frac{d - d_{\min}}{d_{\max} - d_{\min}} \right) \tag{5}$$

여기서 v는 깊이영상의 화소값을 나타낸다. d_{min}은 가장 먼 깊이값을 나타내는 Z_{max}를 식 (4)를 이용하여 변이값으로 변환한 값이고, d_{max}는 가장 가까운 깊이값을 나타내는 Z_{min}을 식 (4)를 이용하여 변이값으로 변환한 값이다. 그림 4는 깊이영상 획득의 한 예로, 1m에서 3m 사이의 장면을 깊이영상으로 변환한 것이다. 추가적으로 깊이 카메라를 이용하여 깊이영상을 획득하면 동일한 객체의 내부에 작은 오류(noise)가 존재하게 된다. 이러한 오류를 줄이기 위하여 간단한 중간값 필터(median filter)를 적용하였다.

그림 4. 1m~3m 거리에서 촬영한 깊이영상 결과
Fig. 4. Captured depth image from 1~3m

Ⅲ. 스테레오 영상 시점의 깊이영상 획득

중앙 시점의 깊이영상을 이용하여 양쪽의 컬러 카메라에서 획득한 것과 같은 깊이영상을 생성하기 위해서는 3차원 워핑(3D warping)을 이용해야 한다. 3차원 워핑은 깊이값을 이용하여 다른 시점으로 원본 시점을 임의의 시점으로 이동하는 영상 합성 방법으로 3차원 영상 처리에 널리 사용되고 있다.

3.1 3차원 영상 워핑을 이용한 상응화소 결정

3차원 워핑은 참조 영상의 각 화소가 임의의 시점의 영상의 어디에 정합이 되는지를 결정해야 한다. 그리고 결정된 임의 시점의 화소에 참조영상의 깊이값을 할당하여 임의 시점의 깊이영상을 생성할 수 있다. 하지만, 원본 깊이영상의 해상도가 매우 낮고, 컬러 카메라와의 거리가 차이가 나므로 단순 3차원 워핑을 이용하면 정확한 깊이영상을 획득할 수 없다. 그러므로 해상도 차이의 영향을 줄이고 효율적으로 두 시점의 깊이영상을 획득하는 방법을 설명한다.

3차원 워핑을 수행하기 위한 시점간 상응 화소 결정 방법은 그림 5에 소개한 것과 같이 깊이영상을 역사영(backward projection)하여 세계 좌표계(world coordinates)로 이동하고, 이 영상을 다시 좌/우 시점으로 각각 전사영(forward projection)하여 영상을 시점 이동한 깊이영상을 생성 할 수 있다.

먼저 중앙시점의 깊이영상의 각 화소와 세계 좌표계의 점의 상응관계를 찾는 방법은 식 (6)과 같다.

$$\mathbf{M}_w = \mathbf{R}_D^{-1}\mathbf{A}_D^{-1}[\mathbf{m}_D \mid 1]^T Z(\mathbf{m}_D) - \mathbf{R}_D^{-1}\mathbf{t}_D \tag{6}$$

여기서 $\mathbf{A}_D$는 깊이 카메라의 내부 변수이고, $\mathbf{R}_D$와 $\mathbf{t}_D$는 각각 깊이 카메라의 방향 행렬과 천이벡터를 나타낸다. 그리고 $\mathbf{m}_D$는 깊이영상의 한 화소를 가리킨다. 이 때 Z는 실제 깊이값을 나타내는데, 깊이영상의 값을 양자화 된 값이므로, 식 (7)과 같이 실제 깊이값을 구한다.

$$Z(\mathbf{m}) = \frac{1}{\frac{D(\mathbf{m})}{255} \cdot \left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{1}{Z_{far}}} \tag{6}$$

주목할 점은, 참조시점의 영상과 임의시점의 영상간의 대응화소를 결정하기 위해서는 정확한 카메라 변수가 있어야 한다. 이때, 영상의 해상도는 카메라 내부 변수와 직접 관련이 있으므로, 두 시점간의 해상도에 따라 정확한 내부 변수를 결정해야 한다.

3.2 깊이카메라에서 획득한 깊이영상 처리

앞에서 설명한 것과 같이, 3차원 워핑은 대상 시점의 해상도와는 관계없이 카메라 변수와 거리 정보만 있으면 수행할 수 있다. 그러므로 제안한 카메라 시스템의 컬러 카메라와 깊이 카메라의 해상도가 크게 차이가 나지만 3차원 워핑을 수행하는 데엔 문제가 없다. 하지만, 깊이영상의 해상도가 작다는 것은 이용할 수 있는 정보가 적다는 것이므로, 정확한 고해상도의 깊이영상을 획득하기 위해서는 그림 6과 같이 단계적으로 깊이영상을 처리해야 한다.

www.dbpia.co.kr

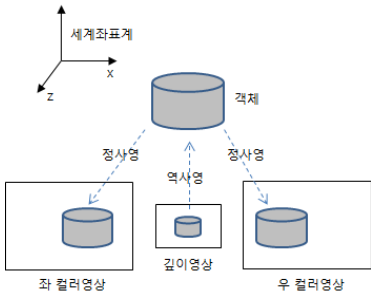


그림 5. 3차원 워핑 방법
Fig. 5. 3D image warping

먼저 저해상도의 깊이영상을 이용하여 양쪽 시점으로 시점 이동하고, 이 때문에 발생한 빈공간을 효율적으로 채워야 한다. 이어서 깊이영상과 컬러영상간의 경계가 불일치하는 문제를 해결한다. 그리고 마지막 단계로 목표로 하는 고해상도의 깊이영상으로 단계적으로 업샘플링하면 컬러영상과 동일한 해상도의 깊이영상을 획득할 수 있다.

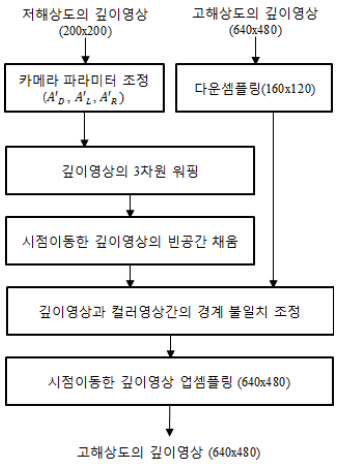


그림 6. 스테레오 카메라 시점의 깊이영상 생성
Fig. 6. Depth image acquisition for stereo cameras

3.2.1 3차원 워핑을 이용한 깊이영상 시점 이동

컬러영상과 같은 시점의 두 깊이영상을 획득하기 위해서는 3차원 워핑을 이용한 시점이동이 필요하다. 이때 컬러영상과 깊이영상간의 해상도 차이가 크므로, 컬러영상의 해상도를 1/4로 주는 줄인다. 이 때, 각 영상의 해상도가 조정되었으므로, 각 카메라의 내부 변수를 다음의 식 (7)과 같이 조정해야 한다.

$$\mathbf{A}'' \equiv s \circ \mathbf{A}' \equiv \begin{bmatrix} s \cdot f & 0 & s \cdot p_x \\ 0 & s \cdot f & s \cdot p_y \\ 0 & 0 & 1 \end{bmatrix} \qquad (7)$$

여기서 $\circ$는 원래의 내부 파라미터를 s로 스케일링 하는 연산자로 정의하고, 컬러영상을 다운샘플링할 때는 s=1/4로 식 (7)에 대입하여 카메라 내부 변수를 조정한다. 이어서 수행하는 3차원 워핑은 앞에서 설명한 상응화소 결정 방법을 이용하여 시점이 이동된 깊이영상을 획득 할 수 있다.

3.2.2 시점 이동한 깊이영상의 빈공간 채움

3차원 워핑을 통해 중앙 시점에서 양쪽으로 시점이동을 할 경우, 그림 7과 같이 빈공간이 발생한다. 중앙의 깊이 카메라에서 좌측 컬러 카메라로 워핑을 수행하면 객체의 왼쪽(O_L)에 빈공간이 생성되고, 반대의 경우, 객체의 오른쪽(O_R)에 빈공간이 생성된다. 이러한 빈공간은 전경객체가 카메라에 가깝게 위치할수록 넓게 생성된다. 또한 빈공간은 깊이값이 할당되지 않은 영역이기 때문에 빈공간 주변의 가용 화소들을 이용하여 적절한 깊이값으로 채워야 한다. 본 논문에서 이용하는 하는 깊이영상의 빈공간 채움 방법은 빈공간이 배경의 확장 영역이라는 가정으로 빈공간의 좌/우측 가용 깊이값 중에서 깊이값이 낮은 깊이값으로 대체하는 방법을 이용한다.

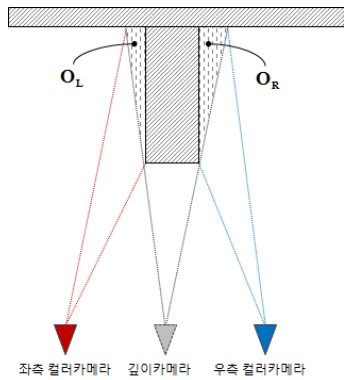


그림 7. 시점이동에 따른 빈공간
Fig. 7. Hole regions due to viewpoint shifting

3.2.3 깊이영상과 컬러영상간의 경계 불일치 조정

앞에서 언급한 3차원 워핑과 빈공간 채움 방법을 이용하면 가운데 위치한 깊이 카메라의 깊이영상을 양쪽의 컬러 카메라에서 획득한 것처럼 보이는 깊이영상을 획득할 수 있다. 하지만, 깊이 카메라에서 획득한 깊이영상은 물체의 반사도 정도와 측정 거리에 따라 깊이값 오류가 포함 될 수 있다. 이러한 깊이값 오류로 인해 컬러영상과 깊이영상간의 객체가 불일치하는 문제가 발생할 수 있다. 그림 8은 이와 같은 객체가 불일치하는 예를 보인 것이다.

www.dbpia.co.kr

그림 8. 깊이영상과 컬러영상간의 객체 불일치
Fig. 8. Boundary mismatches between objects

위와 같은 깊이영상의 불일치값을 보정하기 위하여 그림 9와 같은 방법을 이용할 수 있다. 먼저 시점 이동한 깊이영상을 이용하여 경계영역을 결정한다. 그 다음 경계에서 일정한 거리의 영역을 대상 대상영역으로 설정하고, 최종적으로 컬러영상과 깊이영상을 동시에 고려한 조인트 양방향 필터 (joint bilateral filter)를 이용하여 깊이영상을 보정한다. 영상의 일부분만 깊이값을 보정하는 이유는 대부분의 깊이값 오류가 객체의 경계에서 발생하기 때문이다. 이로 인해 과도한 연산량을 줄일 수 있다.

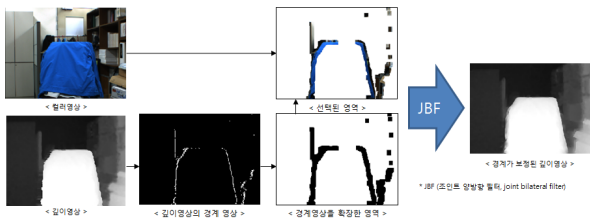


그림 9. 컬러영상과 깊이영상간의 경계 보정
Fig. 9. Object boundary compensation

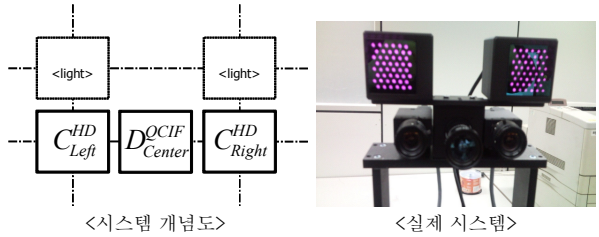
Ⅳ. 실험 결과

4.1 제안한 하드웨어 시스템

앞에서 설명한 시스템은 아래의 그림 10과 같이 세 카메라로 구성된다. 중앙의 저해상도 깊이 카메라와 양쪽에 위치한 고해상도의 컬러 카메라가 1차원 평행형으로 구성된다. 본 연구에서 사용한 깊이 카메라는 적외선 광원(infrared ray light)을 두 대 이용한다. 그림 10(c)는 각 카메라의 가시범위를 나타낸 것으로, 깊이 카메라의 해상도는 작지만 가시범위는 좌/우의 컬러 카메라를 모두 포함할 수 있도록 선택해야 한다. 만약 깊이영상의 가시범위가 두 대의 컬러 카메라의 가시범위를 포함하지 못하면, 깊이영상의 3차원 워핑을 수행한 깊이영상에서 왼쪽 가장자리나 오른쪽 가장자리에 깊이정보가 없는 영역이 발생할 수 있다.

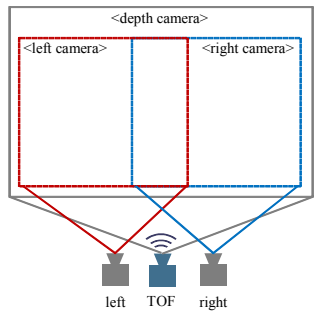
4.2 제안한 소프트웨어 시스템

제안한 카메라 시스템은 한 대의 깊이 카메라로 두 시점의 깊이영상을 생성하고, 획득한 두 시점의 컬러



<시스템 개념도> <실제 시스템>

(a) 카메라 시스템 설계 (b) 실제 카메라시스템



(c) 각 카메라의 가시 범위(field-of-view, FOV)

그림 10. 병합형 3차원 카메라 시스템
Fig. 10. Combined 3D camera system

영상과 깊이영상을 이용하여 다시점의 중간영상을 생성하기 때문에 복잡도가 매우 높다. 이러한 시스템을 촬영부터 다시점 생성까지 다중으로 처리하기 위해서는 구현과정에서 윈도우 프로그래밍의 멀티스레드 (multi-thread) 기법을 적용했다. 이외의 병렬처리를 위한 프로그래밍 기법은 적용하지 않았다. 그림 11은 멀티스레드를 이용한 소프트웨어 구현을 보인 것으로 각 스레드를 컬러영상의 획득과 깊이영상의 획득/워핑/다시점합성 등의 단계로 나누어 여러 스레드가 동시에 작업을 수행하도록 구현했다. Tread_1은 두 대의 컬러 카메라에서 동시에 컬러영상을 획득하고 RGB 형식의 영상으로 저장하는 역할을 수행한다. Thread_2는 TOF 카메라에서 획득한 깊이데이터를 식 (5)를 이용하여 8비트(N=8)의 회색톤의 영상으로 변환한다. Thread_3은 영상 정렬을 통해 획득된 변환행렬을 이용하여 두 시점의 컬러영상을 정렬하는 역할을 수행한다. Thread_4와 Thread_5는 Thread_2를 통해 처리된 깊이영상을 왼쪽과 오른쪽 시점으로 각각 워핑하는 역할을 수행한다. 이 과정에서 빈공간 채움과 객체 경계 불일치 문제를 함께 처리한다. 마지막으로 Thread_6에서 Thread_15까지는 10개의 중간영상을 합성하기 위한 스레드이다. 각 스레드는 컬러 카메라와 동일한 해상도의 두 시점의 컬러영상과 깊이영상을 이용하여 중간 시점의 영상을 각각 생성한다. 기본적으로 이들의 역할은 같으나 합성하는 위치가 다르다.

www.dbpia.co.kr

이와 같이 다중 스레드를 이용한 시스템을 구축할 경우엔 각 스레드간 정확한 동기화가 필요하다. 본 연구에서는 스레드간 동기화를 제한적으로 수행했다. 컬러영상을 획득하는 Thread_1과 Thread_3는 연속적으로 수행되므로 각 스레드의 작업완료 여부를 인지하여 수행하도록 구현했다. 그리고 깊이영상을 획득하고 보정하는 Thread_2는 컬러영상획득과는 독립적으로 수행하고, Thread_4와 Thread_5는 동시에 수행하도록 동기화를 수행했다. 이때 컬러영상 획득과 깊이영상간의 동기화가 이루어지지 않기 때문에 약간의 화면 불일치가 발생할 수 있으나, 기본적으로 각 카메라에서의 처리속도가 빠르기 때문에 이러한 불일치는 눈으로 인지하기 어려울 정도로 적다. 마지막으로 다시점 영상 생성을 위한 다중 스레드는 각 프레임별로 동시에 수행할 수 있도록 동기화를 수행했다.

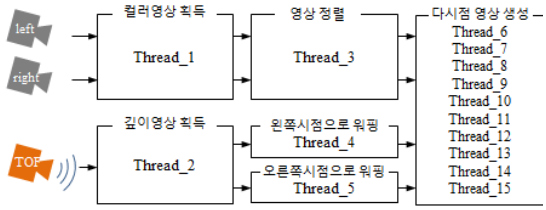


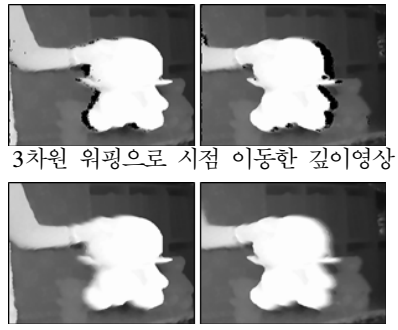
그림 11. 다중 스레드를 이용한 소프트웨어
Fig. 11. Software implementation using multi-thread

4.3 카메라 시스템을 이용한 영상 획득 결과

그림 12은 제안하는 병합형 깊이 카메라 시스템을 이용하여 촬영한 원본 영상을 보인 것이다. 왼쪽의 그림이 깊이 카메라를 이용하여 획득한 원본 깊이영상이고, 나머지 두 영상이 컬러 카메라로 획득한 컬러영상이다. 그림 13는 앞에서 획득한 깊이영상을 이용하여 각 컬러 카메라의 시점으로 시점 이동한 결과를 보인 것이다. 그림 13(a)는 시점 이동으로 드러나는 빈공간(hole)을 보인 것으로, 검정색으로 표시되어있다. 그림 13(b)는 앞에서 제안한 배경의 깊이값을 이용하여 빈공간을 채운 결과를 보인 것이다.

(a) 깊이영상 (b) 왼쪽 컬러영상 (c) 오른쪽 컬러영상

그림 12. 제안한 카메라를 이용하여 촬영한 원본
Fig. 12. Captured images using proposed camera



(a) 3차원 워핑으로 시점 이동한 깊이영상

(b) 빈공간을 채운 깊이영상

그림 13. 3차원 워핑으로 시점이동한 깊이영상
Fig. 13. 3D warped depth images

4.4 다시점 영상 생성

제안한 카메라 시스템의 최종 단계는 두 시점의 컬러영상과 각 시점의 깊이영상을 이용한 다시점 영상의 생성이다. 그림 14는 10개의 중간시점을 추가적으로 생성한 12시점의 다시점 영상을 보인 것이다. 그림의 1번 위치의 영상이 왼쪽 원본영상을 나타내고, 12번 위치의 영상이 오른쪽 원본영상이다. 나머지 영상은 앞에서 설명한 다시점 영상 생성 방법을 이용하여 합성한 영상이다. 시점이동으로 인한 빈공간은 인페인팅 기법을 이용하여 채웠다[18]. 1번부터 12번으로 갈수록 가운데의 인형의 위치가 바뀌는 것을 알 수 있다.

4.5 카메라 시스템의 각 부분별 처리속도 측정

제안한 시스템의 복잡도를 측정하기 위하여 각 부분의 처리속도를 측정했다. 측정을 위해 사용한 컴퓨터는 Zeon E5630@2.53GHz 프로세서와 32GB DDR2 메모리, Windows 7 64-bit로 구성되었다. 이러한 장비를 이용하여 각 기술의 부분별로 표 1과 같이 측정했다.

두 대의 컬러 카메라를 이용하여 VGA(640x480) 해상도의 컬러 영상을 획득하는 프레임율은 약 30 fps이다. 이 이 과정에서 영상 정렬도 함께 수행된다. 깊이 카메라를 이용한 깊이영상의 획득은 약 34 fps의 프레임율을 보였다. 중앙 시점의 깊이영상을 양쪽의 컬러 카메라의 시점으로 시점 이동하는 속도는 약 57 fps의 프레임율을 보였다. 이로써 다시점 영상 생성을 위한 입력데이터를 획득하기 위한 과정은 실시간으로 획득할 수 있었다. 반면 다시점 영상 생성을 위한 시간은 다소 복잡도가 높았다. 총 10시점의 중간영상을 동시에 생성함에 있어서 약 7 fps의 프레임율을 보였다. 특히 시점 이동에 따른 빈공간 채움 과정에서 많은 시간을 소비했다.

www.dbpia.co.kr

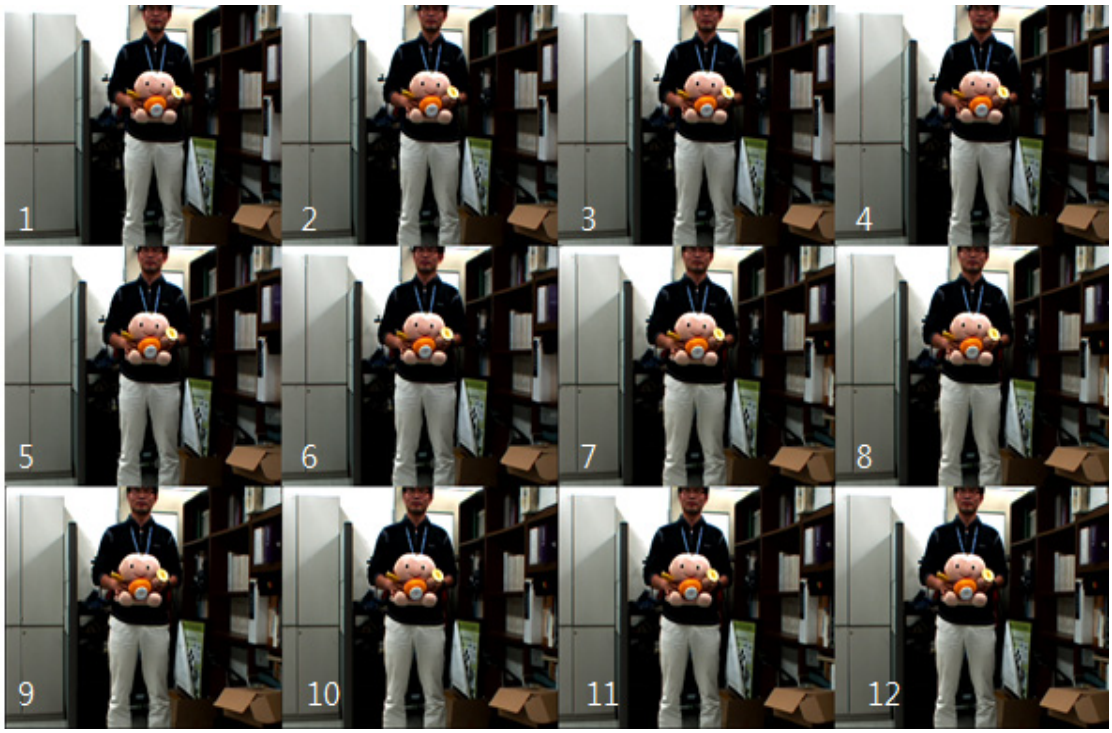


그림 14. 2시점의 영상을 12시점의 다시점영상으로 확장한 다시점 영상
Fig. 14. 12 view images from 2 captured images

표 1. 각 부분별 처리속도 측정 결과
Table 1. Estimated time of each process

구분	세부 기술	평균 속도(ms)	전체 속도(ms)	프레임율 (fps)
컬러영상 획득	컬러영상 획득	17.89	33.54	29.82
	영상 정렬	15.65		
깊이영상 획득	깊이영상 획득	28.04	29.07	34.40
	깊이영상 정렬	1.03		
깊이영상 시점 이동	깊이영상의 시점 이동	17.20	17.63	56.73
	깊이영상의 빈공간 채움	0.42		
다시점 영상 생성	깊이영상을 이용한 시점 이동	42.01	144.15	6.94
	각 시점별 빈공간 채움	102.14		

그림 15는 다중 스레드를 이용한 각 시점별 중간영상 생성 속도와 하나의 스레드로 10개 시점을 생성했을 때의 프레임율을 비교한 것이다. 다중 스레드를 사용하면 시점마다 다소 차이가 발생하지만 대체적으로 중간 시점으로 갈수록 프레임율이 낮아짐을 알 수 있다. 이는 중간 시점으로 갈수록 빈공간의 영역이 넓어져 복잡도가 증가하기 때문이다. 이를 통해 실시간 다시점 영상 생성을 위해서는 보다 빠르고 효과적인 빈공간 채움 방법이 필요함을 알 수 있다. 그리고 하나의 스레드로 10개 시점을 생성했을 때 약 1.44 fps의 속도를 보였다. 이를 통해 다중 스레드를 이용한 다시점 영상 생성 방법이 더 효율적임을 알 수 있었다.

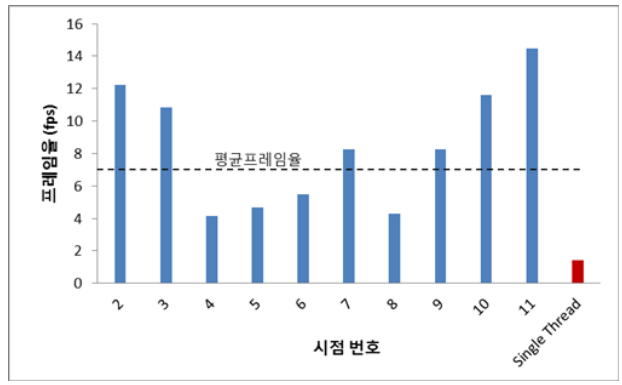


그림 15. 중간영상의 시점별 속도 비교
Fig. 15. Comparison of frame rates

www.dbpia.co.kr

Ⅵ. 결 론

본 논문에서는 저해상도의 깊이 카메라 한 대와 고해상도의 스테레오 카메라를 이용하여 다시점 영상 생성 시스템을 제안했다. 제안한 카메라 시스템은 고해상도의 컬러영상과 같은 시점을 가지는 정확한 고해상도의 깊이영상을 획득하고, 이를 이용하여 다시점 영상을 생성한다. 깊이 카메라를 중간에 설치해서 두 컬러 카메라간 거리가 10.5cm로 다소 크지만, 깊이영상을 이용한 영상 합성을 이용하여 가장 시청하기 편안한 시점을 사용자가 직접 선택할 수 있도록 시스템을 구현했다. 다시점의 영상을 생성하기 위해서는 각 컬러 시점의 깊이영상이 필요하기 때문에 중앙시점에서 획득한 깊이영상을 양쪽의 컬러시점으로 시점을 이동했다. 두 시점의 컬러영상은 편안한 3차원 영상을 제공하기 위한 영상 정렬을 실시간으로 수행하고, 깊이영상의 시점 이동 또한 실시간으로 수행한다. 빠른 데이터의 처리를 위하여 다중 스레드를 이용하여 소프트웨어를 구현했다. 이로 인하여 두 시점의 컬러영상의 획득과 깊이영상의 획득은 실시간으로 획득할 수 있었다. 하지만, 10개의 중간 시점을 생성하는 데는 다소 시간이 오래 걸려 약 7fps의 프레임율의 성능을 보였다. 이는 시점 이동에 따른 빈공간을 채우는 과정이 복잡도가 다소 높기 때문이다. 본 논문에서 제안한 시스템은 빠른 속도로 다시점 영상을 효율적으로 생성할 수 있기 때문에 스테레오 영상 재현뿐만 아니라 다시점 영상 재현에도 효율적으로 이용할 수 있다.

감사의 글

본 연구는 지식경제부, 방송통신위원회 및 한국산업기술평가관리원의 산업원천기술개발사업의 일환으로 수행하였음. [반응형 광시야각 초다시점 3D광학계 원천기술 개발]

참 고 문 헌

[1] C. Wheatstone, "On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision," *Philos. Trans. R. Soc.*, vol. 54, pp. 196-199, 1838.

[2] Yo-Sung Ho, "Realistic 3D Content Producing using Multi-view and Depth Cameras," *The Magazine of the IEEK*, vol. 38, pp. 136-141, 2011.

[3] A. Smolic, K. Müller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "Intermediate View Interpolation based on Multiview Video plus Depth for Advanced 3D Video Systems," in *IEEE ICIP*, pp. 2448-2451, Oct. 2008.

[4] C. Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV," in *Stereoscopic Displays and Virtual Reality Systems XI,* pp. 93-104, Jan. 2004.

[5] ISO/IEC JTC1/SC29/WG11, "Study Text of ISO/IEC 14496-10:200x/FPDAM 1 Constrained Baseline Profile and supplementary enhancement information," in *MPEG output document N10540*, April 2009.

[6] ISO/IEC JTC1/SC29/WG11, "Draft Call for Proposals on 3D Video Coding Technology," in *MPEG output document N11679*, Oct. 2010.

[7] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms," *IJCV*, vol. 47, pp. 7-42, April 2002.

[8] D. Scharstein and R. Szeliski, "High-accuracy Stereo Depth Maps using Structured Light," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. I/195-I/202, June 2003.

[9] R. Kaufmann, M. Lehmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, T. Oggier, N. Blanc, P. Seitz, G. Gruener, and U. Zbinden, "A time-of-flight line sensor - Development and application," *Proceedings of SPIE - The International Society for Optical Engineering*, pp. 192-199, 2004.

[10] M. Lindner, A. Kolb, and K. Hartmann, "Data-fusion of PMD-based distance-information and high-resolution RGB-images," *International Symposium on Signals, Circuits and Systems*, pp. 121-124, 2007.

[11] B. Huhle, S. Fleck, and A. Schilling,

www.dbpia.co.kr

"Integrating 3D Time-of-Flight Camera Data and High Resolution Images for 3DTV Applications," *Proceedings of 3DTV-CON*, 2007.

[12] Y. M. Kim, D. Chan, C. Theobalt, and S. Thrun, "Design and calibration of a multi-view TOF sensor fusion system," *IEEE CVPR*, 2008.

[13] U. Hahne and M. Alexa, "Depth Imaging by Combining Time-of-Flight and On-demand Stereo," *Lecture Notes in Computer Science*, vol. 5742 , ed, pp. 70-83, 2009.

[14] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of time-of-flight depth and stereo for high accuracy depth maps," *IEEE CVPR*, 2008.

[15] E. K. Lee and Y. S. Ho, "Generation of Multi-view Video using a Fusion Camera System for 3D Displays," *IEEE Transactions on Consumer Electronics*, vol. 56, pp. 2797-2805, Nov. 2010.

[16] R. Y. Tsai, "Versatile Camera Calibration Technique for High-accuracy 3D Machine Vision Metrology using Off-The-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, pp. 323-344, Aug. 1987.

[17] Y. S. Kang, C. Lee, and Y. S. Ho, "An Efficient Rectification Algorithm for Multi-view Images in Parallel Camera Array," *Proceedings of 3DTV-CON*, pp. 61-64, 2008.

[18] A. Telea, "An Image Inpainting Technique based on The Fast Marching Method," *Journal Graphics Tools*, vol. 9, pp. 25-36, May 2004.

이 천 (Cheon Lee) 정회원

2005년 한국항공대학교 항공전자공학과 졸업(학사)
2007년 광주과학기술원 정보통신공학과 졸업 (석사)
2007년~현재 광주과학기술원 정보통신공학과 박사과정
<관심분야> 3차원 TV, 실감방송, 3차원 비디오 부호화, MPEG 표준

송 혁 (Hyok Song) 정회원

1999년 광운대학교 제어계측공학과 졸업(학사)
2001년 광운대학교 전자공학과 졸업(석사)
2012년 광운대학교 전자공학과 졸업(박사)
2000년~현재 전자부품연구원 멀티미디어IP연구센터 책임연구원
<관심분야> 영상인식, 3D영상처리, 보안시스템

최 병 호 (Byeongho Choi) 정회원

1991년 한양대학교 전자공학과 졸업(학사)
1993년 한양대학교 전자공학과 졸업(석사)
2010년 중앙대학교 첨단영상대학원 영상공학과 졸업(박사)
1993년~1997 LG전자 비디오 연구소 주임연구원
1997년~현재 전자부품연구원 멀티미디어IP연구센터 센터장
<관심분야> 3D, Vision, 반도체

호 요 성 (Yo-Sung Ho) 종신회원

1981년 서울대학교 공과대학 전자공학과 졸업(학사)
1983년 서울대학교 대학원 전자공학과 졸업(석사)
1989년 Univ. of California, Santa Barbara, Dept. of Electrical and Computer Engineering.(박사)
1983년~1995년 한국전자통신연구소 선임연구원
1990년~1993년 미국 Philips 연구소, Senior Research Member
1995년~현재 광주과학기술원 정보통신공학과 교수
<관심분야> 디지털 신호처리, 영상 신호 처리 및 압축, 디지털 TV와 고선명 TV, 멀티미디어 시스템, MPEG 표준, 3차원 TV, 실감방송

www.dbpia.co.kr