

가중치 결합 양방향 필터를 이용한 깊이 지도의 업샘플링 방법

오 동 루*, 오 병 태*, 신 지 태^o

Up-Sampling Method of Depth Map Using Weighted Joint Bilateral Filter

Dong-ryul Oh*, Byung Tae Oh*, Jitae Shin^o

요 약

3D 영상의 획득을 위해 이용하는 깊이 지도는 영상의 깊이 정보를 가지고 있다. 그러나 일반적으로 고해상도이며 잡음이 제거된 고화질의 깊이 지도를 카메라로부터 직접 획득하는 것은 어렵다. 그래서 저해상도의 깊이 지도를 획득하고 업샘플링 및 전/후 영상처리를 통해 높은 품질의 고해상도 깊이 지도를 획득하는 기법들이 연구되고 있다. 하지만 기존의 연구는 영상의 질에 큰 변수로 작용하는 에지 부분의 효과적 업샘플링이 미흡하다. 그래서 본 논문은 에지 부분을 차별적으로 고려하는 인지적인 특성을 반영한 영상품질향상 연구에 초점을 맞춰 결합 양방향 필터의 가중치를 적응적으로 조절함으로써 깊이 지도와 합성 영상을 개선한 고해상도의 깊이 지도를 얻는 업샘플링 방법을 제안하였다. 제안 방식을 기존의 방식과 비교하였을 때 PSNR 측면과 주관적 품질에서 이득이 있음을 보였다.

Key Words : Depth map, 3D rendering, Up-sampling, Bilateral filter, Joint bilateral filter

ABSTRACT

A depth map is an image which contains 3D distance information. Generally, it is difficult to acquire a high resolution (HD), noise-removed, good quality depth map directly from the camera. Therefore, many researches have been focused on acquisition of the high resolution and the good quality depth map by up-sampling and pre/post image processing of the low resolution depth map. However, many researches are lack of effective up-sampling for the edge region which has huge impact on image perceptual-quality. In this paper, we propose an up-sampling method, based on joint bilateral filter, which improves up-sampling of the edge region and visual quality of synthetic images by adopting different weights for the edge parts that is sensitive to human perception characteristics. The proposed method has gains in terms of PSNR and subjective video quality compared to previous researches.

※ 이 논문은 2015년도 교육부의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(NRF-2013R1A1A2008593)

※ 이 논문은 2015년도 산업통상자원부의 '창의산업융합 특성화 인재양성사업'의 지원을 받아 연구되었음(과제번호 2015-0742-000)

♦ First Author : Sungkyunkwan University, College of Information and Communication Engineering & Department of Human ICT Convergence, narcoticer@skku.edu, 학생회원

^o Corresponding Author : Sungkyunkwan University, School of Electronic and Electrical Engineering, jtshin@skku.edu, 정회원

* Korea Aerospace University, School of Electronics, Telecommunication and Computer Engineering, byungoh@kau.ac.kr

논문번호 : KICS2015-03-480, Received March 9, 2015; Revised May 6, 2015; Accepted June 9, 2015

I. 서 론

3D 영상의 원리는 사람의 좌우 눈에 투영되는 객체의 시각차인 양안시차의 원리와 같다. 사람은 양안시차에 의해 눈앞의 3차원 공간에 대한 입체감을 인식하는데 이러한 양안시차의 원리를 기반으로 3D 영상을 획득하여 이용자에게 영상의 입체감을 느낄 수 있도록 하는 방법이 스테레오스코픽(stereoscopic)이다. 최초로 고안된 지 오랜 시간이 흘렀지만 이제야 활발히 활용되는 이유는 첫 번째, 이용자를 만족하게 할 3D 콘텐츠가 양적으로 부족하였고 두 번째 어지러움, 눈의 피로 현상 등 시각적 문제의 발생으로 인한 장시간 시청의 어려움, 세 번째 이용자의 기대에 미치지 못하는 낮은 해상도 문제를 들 수 있다. 그러나 이러한 문제들은 디스플레이 하드웨어와 소프트웨어의 발전, 비디오의 저장 및 전송 기술의 진보, UN 산하 전기통신에 관한 국제 표준화 기구 ITU-T (International Telecommunication Union-Telecommunication Standardization Sector, 국제 전기통신 연합-전기통신 표준화 부분)와 MPEG 계열 ISO/IEC의 최신 비디오 압축 표준 작업과 맞물려 데이터의 고효율화, 고속 전송, 고압축률의 부호화 기술, 높은 사양의 기기가 대중화되는 등의 여러 요인으로 인해 다양한 3D 콘텐츠를 실시간으로 제공할 수 있는 발판이 마련됨으로써 해결되고 있다.

현재 3D 영상은 좌우 영상을 분리하여 이용자의 양쪽 눈에 제공하는 방법이 보편화 되어 있다. 하지만 전용안경을 써야 한다는 불편함을 야기하므로 다시점 영상을 통해 3D 영상을 제공하는 방법이 현재 연구되고 있는데 다시점 3D 기술은 안경이 필요 없고 시청자의 움직임에 따라 시차가 존재하는 영상을 보여줌으로써 향상된 입체감과 이용자가 직접 시점을 선택할 수 있는 장점을 가진 것으로 알려졌다. 다시점 영상의 획득을 위해서는 물리적으로 많은 카메라를 이용하는 것이 이상적이지만 그보다 적은 수의 카메라로 영상을 촬영하고 합성하여 새로운 가상시점의 영상을 획득하는 방법이 널리 사용되며 이를 깊이 지도 기반의 렌더링(DIBR : Depth Image Based Rendering)이라 한다^{1,2,4)}.

가상시점은 좌우 한 쌍의 컬러 영상과 그에 상응하는 좌우 한 쌍의 깊이 지도 영상으로 합성되어 표현된다. 영상에서 객체의 거리 정보를 포함하고 있는 깊이 지도는 일반 영상과는 달리 카메라와 객체의 상대적인 거리 차를 그레이 스케일로 구성하며 에지를 경계로 객체 내부에서 대부분 비슷한 값을 가지는 특성을

가진다³⁾. 이때 깊이 지도의 질이 합성 시 영상품질에 큰 영향을 미치기 때문에 높은 품질의 깊이 지도를 획득하는 것은 매우 중요하다^{2,5,6)}.

현재 깊이 지도를 획득하는 방법으로는 수동적 방법과 능동적 방법이 존재한다. 수동적 방법의 예로 스테레오 카메라를 이용한 스테레오매칭 방법이 있다. 이 방법은 두 대로 구성된 카메라로 물체를 촬영하고 삼각법을 이용하여 거리정보를 얻는데 알고리즘 복잡도가 높다는 단점을 가지며 텍스처가 없는 영역, 반복되는 영역, 폐색영역에서 문제를 가지게 된다. 능동적 방법에는 레이저를 이용하여 물체와의 거리를 측정하는 방법, ToF(Time of Flight) 센서 기반 적외선을 이용하는 방법이 개발되어 사용 중이다. 이 방법은 해상도가 낮고, 빛에 의한 반사나 야외 등에서 잡음이 발생한다는 단점을 지니지만, 실시간으로 깊이 지도를 획득할 수 있다는 면에서 다양하게 활용되고 있다^{11,8,9)}.

능동적 방법의 단점을 보완하기 위해 저해상도 깊이 지도를 획득한 후 업샘플링과 영상처리기술을 이용하여 고효율 깊이 지도를 얻는 기법들이 연구되고 있으며 기존 가장 보편적인 업샘플링 방법으로 Nearest neighbor 보간법, Bilinear 보간법, Bicubic 보간법 등이 있다. 업샘플링 된 깊이 지도는 합성 영상의 객관적 화질은 물론 사람이 느끼는 영상의 인지적인 품질에 대해서 큰 변수로 작용하는데 특히 영상의 품질을 결정하는데 중요한 에지의 보존과 잡음의 제거가 기존 방법들에 의해 효과적으로 이루어지지 않으므로 이러한 문제를 개선한 양방향 필터(Bilateral Filter)가 업샘플링을 포함하여 잡음 제거(Denoising), 깊이 정보 복원(Depth reconstruction) 등 영상처리 분야 전반에서 광범위하게 사용되고 있다^{10,11)}.

양방향 필터는 가우시안 필터 두 개의 가중치를 곱한 가중치 함수라고 소개할 수 있다. 이 단순 곱으로 이루어진 결괏값은 구하고자 하는 화소의 평활화 정도를 정하게 된다. 가우시안 필터는 각각 영역 필터(Domain filter)와 범위 필터(Range filter)라는 이름을 가지는데 윈도우 안의 화소 값들과 해당하는 각각의 영역 필터 가중치와 범위 필터 가중치의 곱의 총합에 대한 표준화로 현재 화소에 적용되는 값이 정해지는 것이다. 관련 식은 다음과 같다.

$$I_{out}(X) = \frac{1}{k} \int \int \exp\left(-\frac{1}{2}\left(\frac{\|X-m\|}{\sigma_d}\right)^2\right) \exp\left(-\frac{1}{2}\left(\frac{|f(X)-f(m)|}{\sigma_r}\right)^2\right) f(m) dm \quad (1)$$

이때, k 는 Normalization factor 이며 σ_d 와 σ_r 은 표준편차, $f(m)$ 은 필터를 적용하는 이미지도이다. 또한, \mathbb{X} 는 현재 구하려는 화소의 위치, 즉 윈도우 내 중앙이고 m 은 인접 화소의 위치이다. 양방향 필터를 깊이 지도에 적용할 경우 식 (1)에서 적분 내 첫 번째 부분은 영역 필터로 가우시안 분포의 $\|\mathbb{X} - m\|$ 에 현재 값을 구하려는 화소와 인접 화소의 거리를 유클리디언 거리 측정법으로 계산하고 그에 따른 가중치를 구한다. 동시에 두 번째 부분인 범위 필터는 영역 필터가 화소의 유클리디언 단순 거리를 이용한 것과 달리 $|f(\mathbb{X}) - f(m)|$ 에서 현재 값을 구하려는 화소 값과 인접 화소 값의 차를 이용하여 가중치를 구한다. 이 두 필터의 가중치는 각 화소에 대해 이루어지며 위 식의 결과로써 화소 값이 정해진다. 그 결과 현재 값을 구하려는 화소와 물리적 거리가 가깝고 화소 값의 차가 작은 인접 화소는 현재 값을 구하려는 화소에 큰 가중치로 영향을 미치고 그 반대의 경우 적은 가중치로 현재 값을 구하려는 화소에 영향을 주게 된다. 이는 화소 값이 급격히 변화되는 에지와 에지가 아닌 부분, 평탄한 화소 값을 가지는 같은 객체 내 부분에 따라 에지가 보존되고 평활화로 잡음이 제거되도록 적절한 가중치가 계산되어 화소 단위로 적용되는 것이다.

결합 양방향 필터(Joint bilateral filter)^[12,13]는 기본적인 모양이나 필터링 절차는 양방향 필터와 같다. 다른 점은 깊이 지도에 적용할 경우 식 (1)의 두 번째 부분인 범위 필터 $|f(\mathbb{X}) - f(m)|$ 에 양방향 필터는 깊이 지도의 화소 값을 사용한 것과 달리 결합 양방향 필터는 현재 필터링하는 깊이 지도에 상응하는 컬러 프레임의 화소 값의 차를 이용하여 깊이 지도의 디테일, 에지 보존, 잡음 제거를 수행한다. 이 방법은 가이드 필터(Guided filter)의 한 방법으로 이해될 수 있다. 2장에서 결합 양방향 필터를 개선하여 제안하는 가중치 결합 양방향 필터는 식 (1) 양방향 필터의 범위 필터 $|f(\mathbb{X}) - f(m)|$ 에 해당하는 컬러 프레임의 화소 값을 이용하므로 $|C(\mathbb{X}) - C(m)|$ 으로 기재한다.

결합 양방향 필터를 바탕으로 Kopf 등^[14]은 결합 양방향 업샘플링(Joint bilateral upsampling)을 고안하였으며 이는 업샘플링 후에 발생하는 모자이크 현상을 개선하는 등 기존 방법들과 비교하여 향상된 업샘플링 능력을 보인다. 최근에도 이를 기반으로 깊이 지도의 질 향상을 위해 연구가 진행되고 있다. 하지만 이러한 방법들은 여전히 업샘플링을 수행할 때 에지를 원래 화소 분포 특성에 가깝게 보존하는 능력이 미흡하며 사람의 인지적인 특징을 고려하지 않은 업샘

플링 방법들이다. 그래서 본 연구는 깊이 지도를 보간한 뒤 결합 양방향 필터를 기반으로 정적인 배경(static background)보다 움직임이 있는 동적 객체(dynamic object)에 더욱 민감하게 반응하는 사람의 인지적인 특징^[16]을 고려한 가중치 결합 양방향 필터를 적용하여 깊이 지도의 품질 향상과 중간 시점 합성 영상에 존재하는 잡음을 개선, 최종적으로 깊이 지도의 PSNR과 합성 영상의 시각적인 품질을 개선한 영상 시퀀스에 특화된 깊이 지도 업샘플링 기법을 제안한다.

본 논문의 구성은 2장에서는 제안하는 업샘플링 기법에 관해 설명하며 3장에서는 실험결과를 보인다. 4장에서는 향후 과제와 함께 결론을 맺는다.

II. 제안한 깊이 지도 업샘플링 기법

기존 깊이 지도의 업샘플링 방법은 원래 화소 분포 특성에 가까운 값으로 에지를 보존하지 못한다. 이와 관련하여 결합 양방향 업샘플링이 좋은 효과 보이는 이유는 깊이 지도의 업샘플링을 수행할 때 화소 값을 계산하는데 영향을 미치는 가중치를 가이드 영상을 참조하여 구하기 때문이다. 가이드 영상은 저해상도인 깊이 지도에 상응하는 고해상도의 컬러 프레임이다. 하지만 결합 양방향 업샘플링 또한 두 가우시안 가중치 함수의 값에만 의존하기 때문에 4배 이상의 업샘플링을 수행할 경우 블록 현상이 발생하고 경계가 드러나지 않는 부정확한 깊이 지도를 얻게 된다^[7]. 이러한 문제는 에지 부분이 에지가 아닌 부분보다 깊이 지도의 품질에 큰 손실을 준다. 또한, 깊이 지도의 업샘플링 후 이를 이용한 합성 영상에 나타나는 오차는 깊이 지도와 컬러 프레임 간에 동일 객체가 일치되어 합성되지 못하는 이유로 생겨난다. 그래서 본 연구에서는 보간한 깊이 지도에 동적 객체 에지 가이드 영상인 가중치 지도(Weight map), 에지 가중치(Edge weight value), 개선 영역 필터(Domain filter)로 기존 결합 양방향 필터를 변경한 가중치 결합 양방향 필터를 적용하여 기존 업샘플링 방법들의 단점인 업샘플링을 수행할 때 에지가 원래 화소 분포 특성에 가깝게 잘 보존되지 않는 것과 합성 후 깊이 지도와 컬러 프레임의 경계 불일치 때문에 일어나는 경계 잡음, 계단형 잡음을 개선하고자 한다. 본 연구에서 제안하는 방법은 반복적이지 않은 한 번의 필터링으로 수행된다. 결합 양방향 필터를 변경하여 제안하는 가중치 결합 양방향 필터의 식은 다음과 같다.

$$I_{out}(\mathbb{X}) = \frac{1}{k} \int \int \exp\left(-\frac{1}{2} \left(\frac{\|\mathbb{X} - m\|}{\sigma_d} \times \epsilon_{(\sigma_r, N)}\right)^2\right) \exp\left(-\frac{1}{2} \left(\frac{|C(\mathbb{X}) - C(m)|}{\sigma_r} * W_{(C(m))}\right)^2\right) f(m) dm \quad (2)$$

본 연구의 전체 업샘플링 흐름도는 다음과 같다.

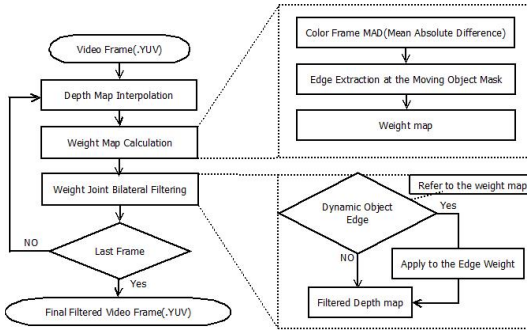


그림 1. 제안한 깊이 지도 업샘플링 흐름도
Fig. 1. Flow chart of the proposed upsampling

위 흐름도에서 보간(Interpolation)은 Bilinear나 Bicubic 기법이 컬러 프레임의 보간에서 좋은 성능을 보이지만 회색 톤이며 에지를 제외한 객체 내에 존재하는 화소들은 대부분 평탄한 값들로 분포되는 깊이 지도의 특성상 Nearest neighbor 기법이 적합하여 사용한다.

2.1 가중치 지도(Weight map) 생성

가중치 지도는 정적인 배경보다 동적 객체에 민감하게 반응하는 사람의 인지적 특성을 고려하여 동적 객체들의 에지를 추출한 동적 객체 에지 지도이며 이를 참조하여 동적 객체의 에지 부분에 아래 2.2에서 설명할 에지 가중치를 적용한다. 본 연구에서 깊이 지도는 품질 면에서 에지가 매우 중요하므로 잡음제거를 위한 평활화 시 다른 부분에 비해 차별적으로 가중치를 주는데 특히 사람이 민감한 동적 객체 에지를 원래 깊이 지도의 화소 분포 특성에 가깝게 보존하기 위해 가중치 지도를 참조하는 것이다. 가중치 지도를 참조해 에지 가중치를 동적 객체의 에지 부분에 적용하므로 가중치 지도는 에지 가이드 영상이라 할 수 있는데 이는 기존 가이드 필터에 해당하는 결합 양방향 필터가 가이드 영상인 교해상도의 컬러 프레임을 참조하듯이 저해상도 깊이 지도에 상응하는 컬러 프레임으로 획득한 에지 가이드 영상을 참조하여 깊이 지도의 동적 객체 에지 부분의 가중치를 계산함으로써 동적 객

체 에지를 보존하는 업샘플링의 과정이기 때문이다. 가중치 지도는 깊이 지도에 해당하는 컬러 프레임의 MAD(Mean Absolute Difference)를 이용하여 구한다.

$$MAD = \frac{\sum_{p=1}^N |I_{p-1} - I_p|}{N} \quad (3)$$

I_{p-1} 은 이전 컬러 프레임의 화소 값을 I_p 는 현재 컬러 프레임의 화소 값을 나타내며 MAD의 결과 화소 값의 변화가 있을 시 그 부분의 화소들이 속해 있는 윈도우를 동적 객체 마스크(Moving Object Mask)로 정의한다. 문턱 값(Threshold)은 다음과 같이 MAD의 결과가 2 이상일 경우이다.

$$T_p = \begin{cases} 255 & \text{if } MAD \geq 2 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

MAD 수행 뒤 문턱 값 이상 화소들이 속한 동적 객체 마스크의 에지를 검출하여 합하면 가중치 지도가 생성된다. 다음은 가중치 지도 획득 흐름도와 그 예다.

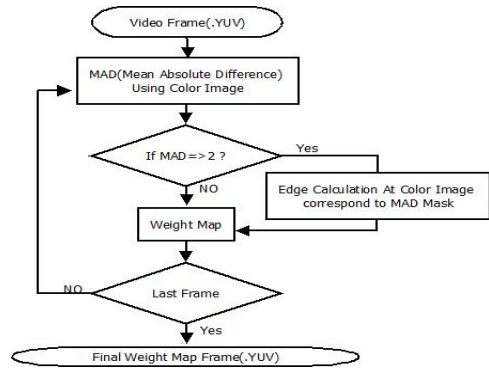


그림 2. 가중치 지도 획득 흐름도
Fig. 2. Flow chart to obtain the weight map

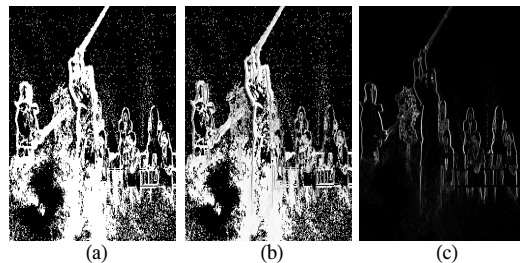


그림 3. (a) MAD (b) 동적 객체 에지추출 (c) 가중치 지도
Fig. 3. (a) MAD (b) Edge extraction of the moving object (c) Weight map

위 그림 2는 그림 1의 첫 번째 그림상자에 해당하는 가중치 지도의 획득 흐름도이며 그림 3은 Kendo 시퀀스에 대한 가중치 지도를 획득하는 각 단계로 (a)는 MAD, (b)는 동적 객체 마스크의 에지 검출, (c)는 가중치 지도이다.

2.2 에지 가중치 (Edge weight value)

결합 양방향 필터의 범위 필터 가중치는 가우시안 분포의 특징에 따라 분자인 $|C(\mathbb{X}) - C(m)|$ 이 작을 수록 커지고 반대로 클수록 작아지는데 이는 인접 화소의 값이 현재 계산하는 화소의 값에 영향을 미치는 평활화 정도를 정한다. 이러한 성질을 이용해 본 연구에서는 가중치 지도를 참조하여 현재 계산하는 화소가 MAD 값 기준으로 동적 객체 에지일 때 인접 화소 중 에지 바깥에 있는 인접 화소가 동적 객체 에지 화소 값을 결정하는데 큰 영향을 주지 않도록 에지 가중치를 범위 필터에 적용하여 에지 바깥에 있는 인접 화소의 가중치를 작게 조정함으로써 동적 객체의 에지 화소 값이 에지 바깥에 있는 인접 화소에 의해 덜 평활화 되도록 에지를 보존한다.

식 (5) 에지 가중치 $W_{(C(m))}$ 을 식 (2)와 같이 기존 결합 양방향 필터에 적용하여 본 논문에서 제안하는 가중치 결합 양방향 필터를 이용한 업샘플링을 수행한다.

$$W_{(C(m))} = \epsilon_{(\sigma_r, N)} * C(m) \tag{5}$$

에지 가중치 식 (5)에서 $C(m)$ 은 가중치 지도를 참조하여 깊이 지도의 화소가 동적 객체 에지일 경우 그 위치의 화소 좌표에 해당하는 컬러 프레임의 화소 값으로써 깊이 지도의 동적 객체 에지 부분의 화소 값을 계산하기 위해 에지 가중치를 적용할 때 해당 컬러 프레임의 화소 값을 적용, 원래 영상의 화소 분포에 맞는 영상을 얻고자 함이다. 이때 $\epsilon_{(\sigma_r, N)}$ 은 $4N^2/\sigma_r$ 으로 정의되는데 N 은 윈도우의 크기, 4 는 실험을 진행하며 경험적으로 얻은 최적 값이다. $\epsilon_{(\sigma_r, N)}$ 은 다음과 같은 이유로 고안하여 적용된다.

기존 결합 양방향 필터는 영역 필터와 범위 필터의 표준편차가 적절히 주어지지 않으면 에지치 않은 잡음이 생성되고 그에 따라 화질이 손상된다. 만일 범위 필터의 표준편차 σ_r 을 고정하고 영역 필터의 표준편차 σ_d 을 변경 하였을 때 작다면 작은 잡음조차 개선하지 못하는 약한 평활화를 수행하며 너무 크다면 강한 평활화를 통해 모든 객체가 뭉개지는 현상을 가져

온다. 반면 영역 필터의 표준편차 σ_d 을 고정하고 범위 필터 σ_r 을 변경하였을 경우 σ_r 이 크다면 범위 필터는 윈도우 내 모든 화소에 거의 같은 가중치를 부여하므로 필터링에 영향을 주지 못해 범위 필터는 없는 것이나 마찬가지인 결과를 얻게 되고 σ_r 이 작다면 컬러 프레임에만 치중되어 평활화되는 결과로 좋지 못한 화질을 가져오게 된다^{10,11,18}. 그러므로 깊이 지도의 에지 보존과 잡음을 제거 하는 데 두 표준편차가 적절한 값을 가져야 좋은 품질의 깊이 지도를 얻을 수 있다. 식 (2)에서 영역 필터에 적용된 $\epsilon_{(\sigma_r, N)}$ 값은 동적 객체 에지일 경우 적용되는 에지 가중치와 상관없이 항상 적용된다.

2.1에서 2.2에 설명하는 내용을 종합하면 사람이 민감하게 인지하는 동적 객체의 에지를 검출한 가중치 지도를 참조하여 동적 객체의 에지 부분과 아닌 부분을 구별하고 차별적으로 가중치를 주어 잡음 제거 시 주변 화소 값에 의해 평활화되는 동적 객체의 에지 부분 화소 값이 덜 평활화 되도록 한다. 그래서 가우시안 분포 성질을 이용하여 동적 객체의 에지 부분 화소 값의 계산에 관여하는 에지 바깥에 있는 인접 화소 값들의 가중치를 작게 조정하기 위해 에지 가중치를 사용하였다. 즉, 기존 결합 양방향 필터는 잡음제거를 위해 에지 부분 또한 어느 정도 평활화가 진행되는 데 반해, 제안 방법은 잡음 제거를 수행하면서 동적 객체의 에지 부분에 대한 평활화를 최소화하여 원래 영상에 가까운 깊이 지도 화소 분포 특성을 보존함으로써 기존 방법을 개선하는 것이다.

아래 그림 4는 기존 결합 양방향 필터와 제안하는 방법에 따라 화소 값을 계산한 예이다. 그림 4의 왼쪽 그림은 보간한 깊이 지도 내 임의의 부분을 확대한 것으로서 오른쪽 원 화소 값은 왼쪽 그림에서 네모 칸이 있는 부분에 해당하는 5 x 5의 윈도우이다. 빗살 부분은 가중치 지도에서 참조할 수 있는 동적 객체 에지 부분인데 설명을 위해, 보간된 깊이 지도와 가중치 지도를 겹친 것이다(여기서 깊이 지도와 가중치 지도가 일치하지 않는 이유는 2.3에서 설명하며 이를 가리켜 경계 잡음이라 함). 그림 4의 왼쪽 네모 칸 중앙의 a 화소는 가중치 지도를 참조하여 찾아낸 깊이 지도 내 동적 객체의 에지 화소이며 오른쪽 원 화소 값에서 240, 245, 250과 10, 15, 20은 에지를 기준으로 다른 객체에 속해 있거나 혹은 에지를 기준으로 구분된 객체와 배경의 화소 값이다. 여기서 기존 방법 적용 가중치, 제안 방법 적용 가중치는 각 방법의 개념에 따른 것이다. 기존 방법 적용 화소 값, 제안 방법 적용

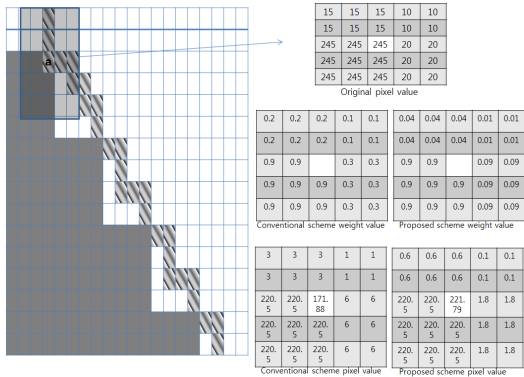


그림 4. 기존 방법과 제안한 방법의 개념 적용 예시 (에지 가중치)
 Fig. 4. An example of the proposed method and the conventional method (edge weight value)

화소 값의 윈도우 중앙 흰색 칸의 값이 최종적으로 구해진 값인데 제안한 방법을 이용하여 구한 화소 값이 원 화소 값의 주변 에지와 에지 안쪽에 있는 화소 값들에 가까움을 알 수 있다(중앙 주변의 화소 값들은 중앙 화소 값을 계산하기 위해 가중치가 적용된 값일 뿐 그 자체로 화소 값을 의미하지는 않음). 이는 잡음을 제거하기 위해 평활화를 수행하되 동적 객체 에지가 평활화되어 깊이 지도의 원래 화소 분포 특성이 변경되는 것을 최소화할 수 있음을 보여준다(원래 화소 값은 정수, 업샘플링에서는 올림 혹은 반올림을 통해 값이 정해짐).

2.3 개선 영역 필터(Domain filter)

기존 업샘플링 방법을 적용하여 획득한 깊이 지도로 가상시점 합성을 했을 때 합성영상에서 III. 실험 결과의 그림 8 (b)와 그림 8 (c)의 객체의 에지 부분의 검은 잔상과 같은 오차가 발생한다. 이러한 것을 경계 잡음(Boundary noise)이라 하는데 이는 합성 시 컬러 프레임과 이에 해당하는 깊이 지도의 화소 간 불일치로 객체의 경계가 일치되지 않아 발생하는 잡음이다 [17].

예를 들어 아래 그림 6의 왼쪽 그림은 임의의 가상시점 합성 영상 내 일부분을 확대한 것으로서 제일 어두운 부분이 저해상도의 깊이지도가 보간된 즉, 업샘플링된 깊이 지도이며 중간 밝기가 깊이 지도에 상응하는 고해상도의 컬러 프레임인데 보는 것과 같이 객체의 경계가 서로 일치되지 않고 깊이 지도의 경계가 더 안쪽에 위치하여 경계 잡음이 존재한다고 할 수 있다. 원래 깊이 지도의 화소 분포 특성에 맞다면 이러한 경계 잡음이 발생하지 않으나 이는 기존 업샘플링

2√2	√5	2	√5	2√2
√5	√2	1	√2	√5
2	1	X	1	2
√5	√2	1	√2	√5
2√2	√5	2	√5	2√2

(a)

1	1	1	1	1
1	0	0	0	1
1	0	X	0	1
1	0	0	0	1
1	1	1	1	1

(b)

그림 5. (a) 기존 영역 필터 유클리디언 거리 예시 (b) 개선 영역 필터 거리 예시
 Fig. 5. (a) An example of an existing domain filter euclidean distance map (b) An example of improved domain filter distance

기법을 사용해 업샘플링한 깊이 지도의 객체 에지 부분 화소 값이 원래 깊이 지도의 화소 분포 특성과 다르게 보간되거나 객체의 에지 경계 부분과 아닌 부분을 효과적으로 구별하지 않는 잡음제거 방법이 적용되어 화소 값이 변경, 소실되는 것을 원인으로 들 수 있다. 그래서 본 논문에서는 기존 결합 양방향 필터의 영역 필터 가중치를 결정하는 $\|X - m\|$ 을 다음과 같이 조정한다.

그림 5 (a)와 같이 결합 양방향 필터의 영역 필터 유클리디언 거리 $\|X - m\|$ 에 그림 5 (b)와 같은 거리를 대신 적용함으로써 합성 영상의 경계 잡음 위치에 해당하는 깊이 지도 내 객체의 에지와 접한 경계 부분 화소 값들이 객체의 에지와 객체의 에지 안쪽 화소의 가중치를 크게 받아 값이 결정되도록 한다. 그림 5 (b)의 거리는 $(\|X - m\| / 2)$ 와 같이 구하며 이는 기존 유클리디언 거리를 2로 나누고 가우스 함수를 적용한 것이다.

기존 결합 양방향 필터를 이용하여 합성 영상의 경계 잡음 위치에 해당하는 깊이 지도 내 객체의 에지와 접한 경계 부분 화소 값을 계산할 때, 영역 필터의 유클리디언 거리인 물리적 거리가 가중치를 결정하였으나 본 연구에서는 최근접 화소인 객체의 에지와 인접 화소(최근접 화소보다 먼)인 객체의 에지 안쪽 화소들의 가중치를 크게 받도록 조정하여 화소 값을 결정하기 위해 최근접 화소와 중앙 화소의 거리, 인접 화소와 중앙 화소 간의 거리를 유클리디언 거리보다 가깝게 조정하였다. 그림으로써 경계 잡음 위치에 해당하는 깊이 지도 내 객체의 에지와 접한 경계 부분 화소 값이 최근접 화소인 객체의 에지 화소 값, 인접 화소

인 객체의 에지 안쪽 화소 값과 더 큰 공간적 상관도를 갖게 되어, 객체의 에지와 접한 경계 부분 화소 값을 결정하는데 객체의 에지와 객체의 에지 안쪽 화소가 미치는 가중치가 커진다. 그래서 그림 6 왼쪽과 같이 깊이 지도와 컬러 프레임 간에 가상시점 합성을 했을 때 발생하는 경계 잡음 위치에 해당하는 깊이 지도 내 객체의 에지와 접한 경계부분의 화소들의 값이 객체의 에지와 객체의 에지 안쪽 화소 값들과 비슷해지며 컬러 프레임 객체의 경계에 비해 안쪽에 위치한 깊이 지도의 객체 경계가 컬러 프레임에 가까워 지도록 한다. 즉, 경계 잡음이 있는 깊이 지도 내 객체의 에지와 접한 부분의 화소 값을 객체의 에지와 에지 안쪽 값과 최대한 비슷한 값으로 채워 컬러 프레임과 경계를 맞추는 것이라 할 수 있다.

아래 그림 6의 a 화소를 예로 들면 깊이 지도 내 a 위치의 화소와 같이 객체의 에지와 접한, 경계 잡음 위의 화소 값을 객체의 에지와 안쪽 화소 값들과 비슷한 값을 가지도록 하기 위해 최근접 화소인 에지 화소 값의 가중치에 크게 영향을 받도록 하여 화소 값을 계산함으로써 가상시점 합성 시 기존보다 1~2화소가량 경계 잡음을 개선하는 것이다. 또한, a 화소의 값이 주변 객체의 에지의 화소 값들과 비교적 가까운 값을 갖게 되어 영상에서 곡선을 표현하는 화소의 계단 크기 (step size)가 작아지므로 가장자리가 곡선인 객체가 기존보다 부드럽게 표현된다. 연속적인 필터링으로 그림 6의 왼쪽 영상의 화소 b~m 또한 같은 절차를 거쳐 값을 가지게 되며 a 화소 값을 결정하는데 기여한다. 아래 그림 6에서 왼쪽 그림의 깊이 지도는 보간을 적용한 것이며 원 화소 값은 그림 6 왼쪽 5 x 5 윈도우의 깊이 지도 화소 값을 가정한 것이다. 기존 방법 적용 가중치, 제안 방법 적용 가중치는 각 방법의 개념에 따른 것이다. 기존 방법 적용 화소 값, 제안 방법 적용 화소 값의 윈도우 가운데 흰색 칸이 최종적으로 구한 값인데 약 3 정도 기존 방법에 의해 상승하였음을 알 수 있다(중앙 주변의 화소 값들은 중앙 화소 값을 계산하기 위해 가중치가 적용된 값일 뿐 그 자체로 화소 값을 의미하지는 않음). 이는 a 주변 화소 값들이 먼저 필터링이 수행되지 않았을 경우를 계산한 결과이며 왼쪽 위부터 오른쪽으로 차례대로 적용되는 공간 필터의 특성상 그림 6의 b~m에 먼저 제안한 방법을 적용하고, a 화소 값을 구하기 위해 가중치를 계산하면 최근접 화소인 에지와 값이 더 가까워진다. 그 결과 결합 양방향 필터 및 기존 업샘플링 방법들보다 경계잡음, 계단형 잡음을 개선하는 것을 실험을 통해 알 수 있다.

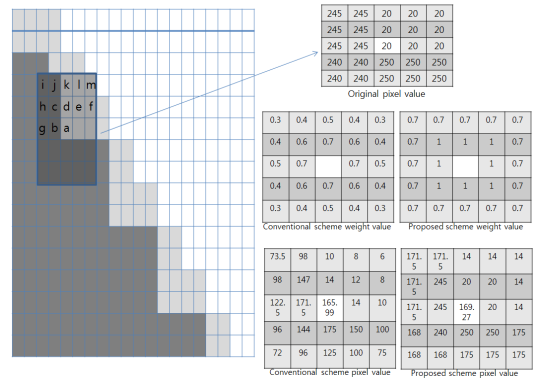


그림 6. 기존 방법과 제안한 방법의 개념 적용 예시 (개선 영역 필터)
 Fig. 6. An example of the proposed method and the conventional method (Modified domain filter)

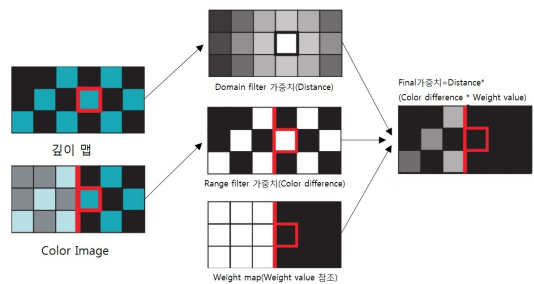


그림 7. 가중치 계산 방법
 Fig. 7. The weight calculation method

위 그림 7은 본 논문에서 제안한 가중치 결합 양방향 필터의 간략한 흐름이다. 가중치 지도를 참조하여 현재 필터링이 진행되는 깊이 지도 내 화소가 동적 객체 에지일 경우에 에지 가중치는 범위 필터의 가중치 계산에 적용되며 개선 영역 필터는 영역 필터 가중치 계산에 항상 적용되어 최종적으로 화소 값이 구해진다. 이는 깊이 지도의 에지, 그중에서도 사람이 민감하게 반응하는 동적 객체의 에지 부분이, 잡음을 제거하기 위한 평활화의 영향을 최소한으로 받도록 가중치를 조정하고 경계 잡음이 있는 위치의 화소 값을 인접 객체의 화소 값들과 비슷하게 함으로써 기존보다 동적객체의 에지 보존이 잘 이루어지며 이를 이용한 합성 영상의 잡음이 개선되도록 하는 것이다.

III. 실험 결과

3D video 표준 Multi view video-plus-depth 시퀀스를 사용하였다. 실험에 사용한 시퀀스는 Dancer(1920 x 1088), Lovebird(1024 x 768), Newspaper(1024 x 768), Gt_fly(1920 x 1088),

Poznan_hall2(1920 x 1088)이다. Aliasing 현상을 고려하여 저역통과필터를 이용해 깊이 지도를 다운 샘플링 하였고 파라미터는 $\sigma_d = 0.5$, $\sigma_r = 0.1$, 윈도우 사이즈는 5로 하였으며 각 시퀀스 당 100프레임의 YUV 시퀀스를 이용하여 실험하였다.

아래 표 1을 보면 업샘플링 비율에 상관없이 기존의 방법들보다 확실히 개선된 PSNR 수치를 확인할 수 있는데 무엇보다 Poznan_hall 수치의 개선이 눈에 띈다. Poznan_hall은 깊이 지도를 관찰해보면 Dancer, Gt_fly와는 달리 객체의 세세한 부분보다 큰 에지를 기준으로 화소 값이 평탄한 값으로 이루어진 것을 알 수 있으며 카메라와 객체가 동시에 움직이는 시퀀스이다. 이는 본 논문의 업샘플링 과정에서 동적 객체의 에지 참조를 위한 가중치 지도를 획득할 때 객체만 움직이는 영상의 가중치 지도보다 더 많은 에지가 검출되어 더 많은 동적 객체 에지에 에지 가중치가 적용됨으로써 다른 시퀀스에 비해 PSNR이 눈에 띄게 개선된 것으로 보인다.

오른쪽 그림 8 (a)의 경우 남자의 머리에 잡음이 발생, 그림 8 (b)와 그림 8 (c)에서는 경계 잡음과 계단형 잡음, 그림 8 (d)에서도 머리 부분에 잡음이 존재함을 알 수 있다. 이와 비교하여 제안한 방법을 적용한 그림 8 (e)를 살펴보면 합성영상의 잡음은 물론 기존 방법들의 경계 잡음과 계단형 잡음을 개선했음을 볼 수 있다.

아래 그림 9 (a)의 원 영상의 네모 부분은 잡음이 발생하여 시각적으로 매우 좋지 않고 중앙 부분에 세로 라인 잡음이 발생하는 것을 볼 수 있다. 이와 비교하여 그림 9 (e)의 제안한 방법을 보면 원 영상의 잡

표 1. PSNR을 이용한 깊이 지도의 객관적 화질 평가 비교 (100프레임 당 평균)

Table 1. Objective visual quality comparison of the depth map using PSNR metric (Average per 100 frame)

Sequence	Up-sampling rate	Bilinear	Bicubic	Joint bilateral	Propose method
Poznan hall	x 4	47.52	47.76	47.80	49.08
	x 8	43.43	43.60	44.85	45.84
	x 16	39.57	39.70	41.55	42.10
Love bird1	x 4	43.21	43.47	43.81	44.65
	x 8	39.63	39.83	40.89	41.53
	x 16	36.42	36.58	37.89	38.23
Dancer	x 4	38.56	38.77	39.48	40.12
	x 8	34.99	35.12	36.41	36.92
	x 16	31.64	31.74	33.19	33.48
News paper	x 4	35.10	35.37	36.12	36.88
	x 8	31.22	31.41	32.77	33.32
	x 16	27.81	27.96	29.52	29.81
Gt_fly	x 4	44.55	44.75	45.18	45.86
	x 8	41.51	41.72	42.81	43.46
	x 16	38.27	38.50	40.23	40.70

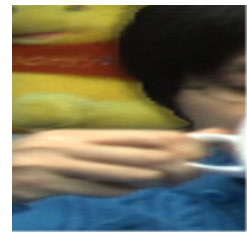
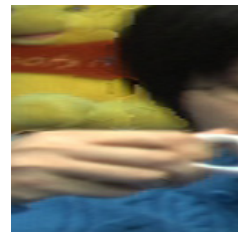
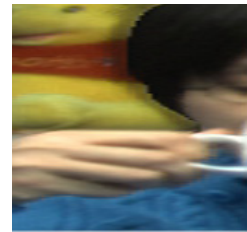
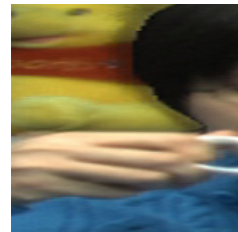
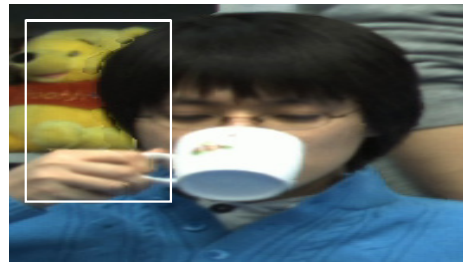


그림 8. Newspaper 영상을 이용한 합성 영상의 시각적 품질 비교 (85번째 프레임) (a) original (b) Bilinear (c) Bicubic (d) Joint bilateral upsampling (e) Proposed scheme
Fig. 8. Visual quality comparison using Newspaper (85th Frame) (a) original (b) Bilinear (c) Bicubic (d) Joint bilateral upsampling (e) Proposed scheme

음이 워낙 심해 완전하진 않으나 그림 9 (b), 그림 9 (c), 그림 9 (d)와 같은 기존 방법과 비교하면 잡음을 최소화함을 알 수 있다. 이는 확대한 영상이 아닌 원 크기의 영상에서는 눈에 거슬리지 않을 만큼 자연스럽다. 표 1과 그림 8, 그림 9에서 업샘플링율이 증가할수록 기존 화소의 변경, 소실로 인해 화질이 저하되는 기존 업샘플링 방법들보다 효과적인 업샘플링 방법임을 수치와 시각적으로 보여주고 있다.

이러한 실험 결과 외에 업샘플링 비율을 2로 하여 실험한 최신 업샘플링 연구^[15]와 비교하여 시퀀스 Dancer(1920 x 1088), Lovebird1(1024 x 768), Newspaper(1024 x 768), Gt_fly(1920 x 1088)를 PSNR을 이용하여 비교했을 때 기존 업샘플링 방법 대비 약 0.2 %의 미미한 개선 효과를 보였다. 하지만 본 연구가 제안하는 방법은 업샘플링 비율이 2인 결과보다 그 이상의 비율로 업샘플링 하였을 때 더 확실한 효과를 위의 표 1에서 보인다. 실제 저해상도의 깊

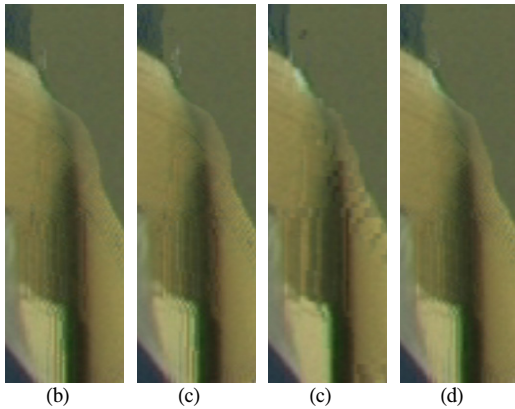


그림 9. Poznan_hall2 영상을 이용한 합성 영상의 시각적 품질비교 (73번째 프레임) (a) original (b) Bilinear (c) Bicubic (d) Joint bilateral upsampling (e) Proposed scheme
 Fig. 9. Visual quality comparison using Poznan_hall2 (73th Frame) (a) original (b) Bilinear (c) Bicubic (d) Joint bilateral upsampling (e) Proposed scheme

이 지도의 업샘플링 또한 컬러 프레임에 상응하는 해상도를 얻기 위해 2보다 더 큰 비율로 업샘플링을 수행해야 하므로 본 논문의 실험결과와는 의미 있다고 할 수 있다. 제안하는 방법은 VGA, QVGA, CIF, QCIF 등의 포맷을 이용한 결과 기존 업샘플링 기법들보다 수치상으로 뛰어난 효과를 보였으며 위 표에서 다루지 않은 16배 업샘플링 그 이상에서도 효과를 보인다.

IV. 결 론

본 연구는 고해상도, 높은 품질의 깊이 지도와 합성 영상 획득을 위해 결합 양방향 필터를 개선해 잡음 제거와 함께 동적 객체의 에지 화소 값을 보존하는 업샘플링 방법을 제안하고자 저해상도의 깊이 지도에 상응하는 고해상도의 컬러 프레임을 이용해 동적 객체의 에지를 검출하여 획득한 가중치 지도를 깊이 지도 업샘플링 시 동적 객체 에지 보존을 위한 가이드 영상

으로 참조, 깊이 지도 내에서 동적 객체의 에지와 아닌 부분을 구별하여 차별적으로 가중치를 적용하였으며 경계 잡음과 계단형 잡음을 개선함으로써 합성 영상의 품질을 향상시키고자 하였다. 그 결과 기존 방법들을 적용한 깊이 지도의 평균 PSNR이 Bilinear은 38.23dB, Bicubic은 38.42dB, 결합 양방향 업샘플링은 39.50dB인 것에 비교하여 제안하는 방법은 40.13dB로 개선되었다. 또한, 합성 영상의 시각적 품질 측면에서도 더욱 자연스러운 합성 영상을 얻을 수 있음을 위에서 보았다. 특히, 카메라와 객체가 동시에 움직이는 실사 프레임으로 이루어진 영상에 대한 4배 이상의 업샘플링에서 효과적이었다는 점과 반복적이지 않은 한 번의 공간 필터링으로 깊이 지도의 에지 보존과 잡음 제거, 가상시점 합성 시 잡음을 개선했다는 이점을 가진다. 이러한 결과는 사람이 인지하는 가상시점 합성 영상의 품질을 향상해 줄 것으로 기대된다. 추후 연구로 능동적, 수동적 방법으로 직접 획득한 깊이 지도에 적용하여 검증과 보완하는 연구가 필요하다.

References

- [1] W. S. Cheong, K. J. Yun, G. Bang, G. S. Lee, and N. H. Hur, "A trend of multiview 3DTV technologies and standardization," *Etri. Electron. Telecommun. Trends*, vol. 29, issue. 3, pp. 17-26, Jun. 2014.
- [2] Y. S. Ho, "Improve quality of the depth-map upsampling technology," *Broadcasting & Technol. Mag.*, vol. 215, pp. 158-163, Nov. 2013.
- [3] R. Cheng and K. Nahrstedt, "Empirical study of 3d video source coding for autostereoscopic displays," in *Proc. ACM. Conf. Multimedia*, pp. 573-576, Augsburg, Germany, Sept. 2007.
- [4] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in *Proc. SPIE. Conf. Stereoscopic Displays and Virtual Reality Systems XI*, pp. 93-104, San Jose, CA, USA, May 2004.
- [5] K. Muller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," in *Proc. IEEE.*, vol. 99, no. 4, pp. 643-656, Apr. 2011.
- [6] S. B. Lee and Y. S. Ho, "Real-time eye

contact system using a kinect depth camera for realistic telepresence,” *J. KICS.*, vol. 37C, no. 4, pp. 277-282, May 2012.

[7] J. H. Kim, J. W. Choi, and K. H. Son, “Depth map resolution and quality enhancement based on edge preserving interpolation,” in *Proc. KSBE Conf.*, pp. 8-10, Seoul, Korea, Nov. 2011.

[8] J. W. Choi, D. B. Min, B. S. Ham, and K. H. Sohn, “Spatial and temporal up-conversion technique for depth video,” in *Proc. IEEE Conf. Image Processing*, pp. 3525-3528, Cairo, Egypt, Nov. 2009.

[9] X. Zhu, X. Song, and X. Chen, “Image guided depth map upsampling using anisotropic TV-L2,” *IEEE Signal Process. Lett.*, vol. 22, pp. 318-321, Mar. 2015.

[10] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. IEEE Conf. Computer Vision*, pp. 839-846, Bombay, India, Jan. 1998.

[11] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, “Bilateral filtering : Theory and applications,” *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, 2008.

[12] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital photography with flash and no-flash image pairs,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 664-672, Aug. 2004.

[13] E. Eisemann and F. Durand, “Flash photography enhancement via intrinsic relighting,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 673-678, Aug. 2004.

[14] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, “Joint bilateral upsampling,” *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007.

[15] J. H. Kim, G. G. Jeon, and J. C. Jeong, “Joint-adaptive bilateral depth map upsampling,” *Elsevier, Signal Processing-Image Commun.*, vol. 29, no. 4, pp. 506-513, Apr. 2014.

[16] T. W. Kim, J. H. Kim, M. W. Park, and J. T. Shin, “Hybrid down-sampling method of depth

map based on moving objects,” *J. KICS.*, vol. 37A, no. 11, pp. 918-926, Nov. 2012.

[17] M. S. Ko and J. S. Yoo, “Boundary noise removal and hole filling algorithm for virtual viewpoint image generation,” *J. KICS.*, vol. 37A, no. 8, pp. 679-688, Aug. 2012.

[18] T. W. Bae and Y. T. Kim, “Small target detection method using bilateral filter based on surrounding statistical feature,” *J. Korea Multimedia Soc.*, vol. 16, No. 6, pp. 756-763, Jun. 2013.

오 동 루 (Dong-ryul Oh)



2013년 8월 : 동국대학교 정보통신공학과 졸업
 2013년 8월~현재 : 성균관대학교 정보통신대학 성균융합원 휴먼 ICT융합학과 석사과정
 <관심분야> 3차원 영상처리, 스테레오 매칭

오 병 태 (Byung Tae Oh)



2003년 8월 : 연세대학교 전기전자공학부 학사
 2009년 8월 : Univ. of southern California (USC), Dept. of Electrical Eng. 석사 및 박사
 2013년 3월~현재 : 한국항공대학교 항공전자공학부 조교수
 <관심분야> 3차원 영상시스템, 영상처리, 영상압축

신 지 태 (Jitae Shin)



1986년 2월 : 서울대학교 전기공학과 학사
 1988년 2월 : KAIST 원자력공학과 석사
 2001년 5월 : Univ. of southern California (USC), Dept. of Electrical Eng. 석사 및 박사
 2001년 8월~2002년 2월 : 경희대학교 정보통신전문대학원 교수
 2002년 3월~현재 : 성균관대학교 정보통신대학 교수
 <관심분야> 영상통신네트워크, 영상처리