

깊이 불연속 정보를 이용한 저해상도 깊이 영상의 업샘플링 방법

강 윤 석^{*}, 호 요 성[°]

Low-Resolution Depth Map Upsampling Method Using Depth-Discontinuity Information

Yun-Suk Kang^{*}, Yo-Sung Ho[°]

요 약

시청자에게 입체감과 몰입감을 줄 수 있는 3차원 영상의 제작을 위해서는 장면의 색상 영상과 함께 깊이 정보가 필요하다. 일반적으로 장면의 깊이를 측정하는 깊이 센서에서 획득된 깊이 영상은 매우 작은 해상도를 가진다. 따라서 색상 영상과 함께 3차원 영상 제작에 이러한 깊이 영상을 사용하기 위해서는 저해상도 깊이 영상의 업샘플링 기술이 필요하다. 본 논문에서는 깊이 불연속 정보를 이용하여 저해상도 깊이 영상을 업샘플링하는 방법을 설명한다. 깊이 영상을 업샘플링할 때 가장 민감하게 다루어야 할 깊이 불연속 부분을 고해상도 색상과 저해상도 깊이 영상으로부터 찾아낸다. 그리고 깊이 불연속 부분을 고려하여 깊이 영상 업샘플링을 위한 에너지 함수를 모델링하고, 신뢰 확산(belief propagation) 방법을 이용하여 해상도가 확대된 깊이 영상을 획득한다. 제안하는 방법은 펠터 기반이나 에너지 함수 기반의 다른 방법들보다 우수한 성능을 나타내었다.

Key Words : Depth upsampling, Depth interpolation, Markov random field, TOF depth sensor, 3D video

ABSTRACT

When we generate 3D video that provides immersive and realistic feeling to users, depth information of the scene is essential. Since the resolution of the depth map captured by a depth sensor is lower than of the color image, we need to upsample the low-resolution depth map for high-resolution 3D video generation. In this paper, we propose a depth upsampling method using depth-discontinuity information. Using the high-resolution color image and the low-resolution depth map, we detect depth-discontinuity regions. Then, we define an energy function for the depth map upsampling and optimize it using the belief propagation method. Experimental results show that the proposed method outperforms other depth upsampling methods in terms of the bad pixel rate.

I. 서 론

3차원 비디오는 입체 영상을 통해 사용자에게 깊이 감과 몰입감을 줄 수 있는 차세대 영상 서비스로 각광 받고 있다^[1]. 양안식(stereo) 영상은 두 시점 간 존재하

는 시차정보로 인해 거리감을 제공하며, 다시점(multi-view) 영상의 경우 양안식 영상보다 더 넓은 시야각과 풍부한 입체감을 제공해 줄 수 있다. 또한 양안식 혹은 다시점 영상과 함께 촬영된 장면의 깊이 정보를 가질 때 3차원 비디오의 다양한 활용이 가능

* First Author : 광주과학기술원 정보통신공학부, 영상통신연구실, yunsuk@gist.ac.kr, 정회원

° Corresponding Author : 광주과학기술원 정보통신공학부, 영상통신연구실, hoyo@gist.ac.kr, 정회원

논문번호 : KICS2013-06-271, 접수일자 : 2013년 6월 30일, 최종논문접수일자 : 2013년 8월 26일

하다.

일반적으로 장면의 깊이 정보는 카메라로부터 장면 까지의 거리가 상대적으로 양자화된 깊이 영상의 형태로, 혹은 정렬된 양안식 영상에서 현재 화소의 위치와 인접 시점에 해당 화소가 맷히는 위치의 화소 간 거리를 값으로 가지는 변위(disparity) 영상의 형태로 사용된다. 이러한 깊이 정보는 3차원 영상 제작 및 3 차원 응용분야의 구현에 있어 필수적인 요소로 사용된다. 따라서, 장면의 정확한 깊이 정보를 획득하는 것이 매우 중요하다^[2,3].

깊이 정보는 가장 대표적으로 스테레오 정합(stereo matching)을 이용하여 구할 수 있으며^[4], 장비를 이용하여 장면의 깊이를 직접 측정할 수도 있다. Time-of-Flight (TOF) 기술을 이용한 깊이 센서는 적외선 신호의 위상 차를 이용하여 센서로부터 장면 내 물체까지의 거리를 직접 측정하여 깊이 영상을 출력한다. 깊이 센서는 장면의 정확한 깊이 정보를 실시간으로 획득할 수 있지만, 출력 해상도가 작다는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해, 스테레오 정합과 깊이 센서의 정보를 함께 이용하는 방법이 제안되었고^[5], 깊이 영상의 해상도를 업샘플링 하는 여러 가지 기술들도 최근 많은 연구가 수행되고 있다.

깊이 영상의 업샘플링은 방법은 크게 두 가지로 나뉘는데, 필터 기반의 방식과 마르코프 랜덤 필드(Markov random field, MRF) 기반의 방식이다. 필터 기반의 방식으로는 양방향 필터(bilateral filter)를 변형한 결합 양방향 필터를 이용한 업샘플링(joint bilateral upsampling, JBU)이나 깊이 영상 업샘플링을 위한 잡음인자 필터(noise-aware filter for depth upsampling, NAFDU)가 많이 이용된다^[6,7]. 이 방법에서는 깊이 영상의 해상도 확대로 인한 빈 공간을 채울 때, 색상 영상의 정보의 분포 특성을 기반으로 하여 깊이 값을 채우기 때문에, 업샘플링 된 깊이 영상에 색상 영상의 특성이 나타나거나, 주변의 깊이 값을 혼합하기 때문에 경계가 흐려지는 현상이 발생하기도 한다.

한편 MRF 기반의 방법은 색상 영상과 깊이 정보가 포함된 확률 모델을 이용하여 빈 공간에 해당하는 깊이 값을 위한 에너지 함수를 모델링하고, 그로부터 계산되는 비용이 가장 작은 깊이 해당 화소의 깊이 값으로 할당된다^[8]. 에너지 함수의 최적화를 위해서는 그래프 절단(graph cut)이나 신뢰 확산(belief propagation) 방법 등이 주로 이용된다. 그러나 MRF 기반의 방법은 잡음이 주변 영역으로 전파되는 단점

을 가지고 있다.

본 논문에서는 깊이 불연속 정보를 이용하여 저해상도 깊이 영상을 업샘플링하는 방법을 제안한다. 제안하는 방법에서는 고해상도의 색상 영상과 저해상도 깊이 영상을 이용하여 색상 영상에서의 깊이 불연속 구간을 정의하고, 그 구간을 침범하여 인접 화소들의 값에 영향을 미치지 않도록 하고, 또한 현재 화소와 그 주변 영역의 평탄한 정도를 고려하여 MRF 모델의 평활화(smoothness) 항을 설계한다. 최종적으로 신뢰 확산을 이용하여 최적화함으로써 고해상도의 깊이 영상을 생성한다.

본 논문의 구성은 다음과 같다. II장에서는 MRF 모델과 깊이 영상 업샘플링에 대해 소개하고, III장에서 제안하는 방법을 자세히 설명한다. IV장에서 깊이 영상 업샘플링의 실험 결과를 보인 후 V장에서 결론을 맺음으로 논문을 마무리한다.

II. Markov Random Field 모델과 깊이 영상 업샘플링

그림 1과 같이, MRF 모델을 이용하여 깊이 영상의 업샘플링을 수행하기 위해서 사용되는 영상 및 화소 정보는 다음과 같다. 먼저 저해상도의 깊이 영상과 함께 색상 영상이 필요하다. 업샘플링 된 깊이 영상의 화소를 y 라고 지칭하며, z 는 저해상도 깊이 영상으로부터 색상 영상의 위치로 투영된 값을 의미한다. 이 때, 현재 화소 위치 i 에서 자료(data) 항과 평활화 항으로 구성된 깊이 영상의 업샘플링을 위한 에너지 함수 E 는 식 (1)과 같이 정의할 수 있다^[9].

$$E = \sum_{i \in L} k(y_i - z_i)^2 + \sum_i \sum_{j \in N(i)} w_{ij}(y_i - y_j)^2 \quad (1)$$

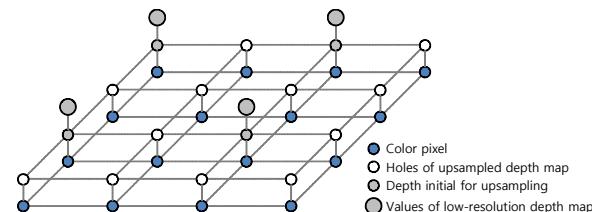


그림 1. MRF 모델을 이용한 깊이 업샘플링 구조
Fig. 1. Depth upsampling using MRF model

상수 k 는 자료 항의 크기를 조절하며, j 는 화소 i 와 상하좌우로 인접한 영역인 $N(i)$ 에 속하는 화소들을 의미한다. 그리고 화소 i 와 j 의 관계를 바탕으로 정의되는 평활화 항의 가중치 w_{ij} 는 평활화 항의 크기를 조절한다. 자료 항에서는 저해상도 깊이 영상의 값

이 투영되어 존재하는 위치에서만 후보 값들과의 비용 계산이 수행되며, 그 이외의 영역에서는 평활화 항만 계산된다. 이와 같이 구성된 에너지 함수로부터 각 화소에서 깊이 값들의 후보군들에 대한 비용 계산이 신뢰 확산이나 그래프 절단과 같은 전역 최적화 방법을 통해 이루어지고, 각 화소에서 최적의 깊이 값을 찾아낸다.

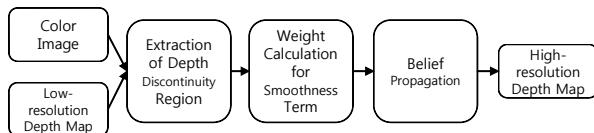


그림 2. 제안하는 방법의 흐름도
Fig. 2. Flow chart of the proposed method

III. 깊이 불연속 정보를 이용한 깊이 영상 업샘플링 방법

본 장에서는 깊이 불연속 정보를 이용하여 MRF 모델 기반의 깊이 영상을 업샘플링 방법을 설명한다. 제안하는 방법의 전체 흐름은 그림 2에 나타나 있다.

II장에서 소개한 것처럼, 에너지 함수의 평활화 항의 영향력은 가중치 w_{ij} 에 의해 결정되는데, 일반적으로 깊이 값의 변화가 작은 평탄한 영역에서는 그 영향력을 크게, 그리고 깊이 불연속 구간과 같이 깊이 값의 변화가 큰 영역에서는 그 영향력을 작게 가지도록 가중치를 설계해야 한다. 따라서 제안하는 방법에서는 식 (2)와 같이 가중치를 깊이 불연속 정보에 기반한 색상 차이의 가중치 w_{cij} 와, 깊이 영역의 변화도를 고려한 가중치 w_{di} 의 곱으로 정의하고, 각 가중치는 다음과 같이 계산한다.

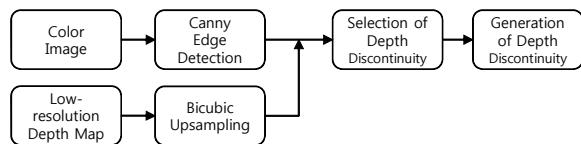


그림 3. 깊이 불연속 영역 생성 방법
Fig. 3. Generation of depth-discontinuity region

$$w_{ij} = w_{c,ij}w_{d,i} \quad (2)$$

먼저, 깊이 불연속 정보를 고려한 색상 차이 가중치를 업샘플링에 이용할 수 있다. 색상 영상에서 깊이의 불연속 영역을 찾기 위해서는 먼저 영상의 경계(edge) 영역을 구해야 한다. 그리고 색상 영상의 경계 영역 중 깊이 불연속 영역이라고 판단되는 부분들을 남기고 나머지 부분들을 제거하면 깊이 불연속 영역이 생

성된다. 이 과정은 그림 3과 같은 순서로 수행된다.

그림 4(a)에 나타난 영상에 대하여 Canny 경계 검출 방법을 적용한 결과가 그림 4(b)에 나타나 있다. 색상 영상의 차이를 이용하였기 때문에, 실제 깊이는 변화가 없는 영역에서도 경계 영역이 검출된 것을 볼 수 있다. 이제 깊이 불연속 영역의 검출을 위해서 저해상도 깊이 영상을 bicubic 방법을 이용하여 업샘플링한다. 그 결과 깊이 영상에서의 경계 부분, 즉 깊이 불연속 부분에 흐려짐 현상이 나타난 결과를 그림 4(c)에서 볼 수 있다. 이제 식 (3)을 이용하여 그림 4의 (b)와 (c)를 이용하여 깊이 불연속 영역 맵 D_d 를 찾아낼 수 있다. 깊이 불연속 맵의 계산은 색상 영상에서 경계로 검출된 화소에 대해서만 수행한다.

$$D_d(i) = \begin{cases} 1 & \text{if } |\min_{i \in N_d(i)} D_{\text{bicubic}}(i) - \max_{i \in N_d(i)} D_{\text{bicubic}}(i)| > th_d \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

D_{bicubic} 은 저해상도 깊이 영상을 bicubic 방법으로 업샘플링한 영상을 말하고, 영역 $N_d(i)$ 는 화소 i 를 기준으로 일정 크기의 정사각 영역을 설정한 것이다. th_d 는 깊이 불연속 영역을 허용하는 문턱 값으로 사용된다. 이렇게 구해진 깊이 불연속 맵이 그림 4(d)에 나타나 있다. 그림 5는 사용된 문턱 값에 따른 깊이 불연속 영역을 보여준다. 문턱 값이 작을수록 깊이 불연속 영역 주변의 잡음이 많이 발생하고, 너무 큰 문턱 값에서는 오히려 불연속 영역이 사라지는 결과를 보여준다.



그림 4. 깊이 불연속 영역 생성 결과
Fig. 4. Result of depth-discontinuity region generation



그림 5. th_d 값의 변화(4, 6, 8, 10)에 따른 깊이 불연속 영역
Fig. 5. Depth-discontinuity region according to the threshold value thd

이제 깊이 불연속 정보를 고려한 색상 차이 가중치를 정의하기 위해서 세 가지 경우를 고려한다. 첫 번째로, 현재 화소 i 및 인접 화소 집합 $N(i)$ 가 깊이 불연

속 구간과 만나지 않을 때 색상 차이 가중치는 식 (4)와 같이 정의된다.

$$w_{c,ij} = \exp\left(-\frac{\sqrt{\sum_{I=RGB}(I(i)-I(j))^2}}{2\sigma_{c,ij}^2}\right) \quad (4)$$

먼저, 식 (4)에서 I 는 색상 영상의 화소 값을 나타내며, R, G, B 각 채널에서 식이 계산된다. 는 가중치의 크기를 조절하는 역할을 한다.

두 번째로, 그림 6(a)와 같이 현재 화소가 깊이 불연속 영역일 때 가중치는 다음과 같이 정의된다. 깊이 불연속 영역은 색상의 경계 영역과 일치하기 때문에, 그 영역은 두 영역의 색상이 혼합되어 나타나는 특성이 있다. 따라서 색상의 차이를 더 명확히 하기 위해서는 현재 화소 i 의 인접 영역 $N(i)$ 를 식 (5)와 같이 수정하고, 색상 차이 가중치는 식 (6)과 같이 정의한다. 식 (6)의 j_m 은 식 (5)의 $N'(i)$ 에 속하여 m 개 화소 값에 대한 평균값을 취하여 현재 화소의 색상 값과의 차이를 구하여 가중치에 이용하게 된다.

$$N'(i) = N(x,y) = \begin{cases} (x+1,y), (x+2,y), \dots, (x+m,y), \\ (x-1,y), (x-2,y), \dots, (x-m,y), \\ (x,y+1), (x,y+2), \dots, (x,y+m), \\ (x,y-1), (x,y-2), \dots, (x,y-m) \end{cases} \quad (5)$$

$$w_{c,ij} = \exp\left(-\frac{\sqrt{\sum_{I=RGB}(I(i)-\frac{1}{m}\sum_{p=1}^m I(j_p))^2}}{2\sigma_{c,ii}^2}\right) \quad (6)$$

마지막으로, 그림 6(b)와 같이 현재 화소 i 의 인접 화소가 깊이 불연속 영역일 때에는, 불연속 영역에 속한 인접 화소 너머에 있는 영역의 값들이 영향을 주지 못하도록 식 (7)과 같이 색상 차이 가중치를 정의한다. 식 (7)의 C_{trunc} 는 매우 큰 상수로써 가중치를 0에 가깝도록 만들어준다.

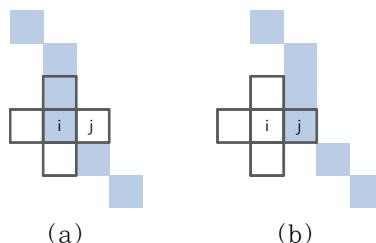


그림 6. 현재 화소와 깊이 불연속 영역과의 관계
Fig. 6. Relation between the current pixel and depth-discontinuity

$$w_{c,ii} = \exp(-C_{trunc}) \quad (7)$$

다음으로 깊이 영역의 변화도를 고려한 가중치를 구할 때, 현재 화소 i 에서 인접한 깊이 영역의 변화도를 이용한다. 앞서 설명한 것처럼, 어떤 영역에 대한 깊이 분포가 큰 편차를 가진다면 그 영역은 깊이 불연속 구간일 확률이 크다. 따라서 가중치를 통해 평활화 항의 영향을 조절해 주어야 한다.

깊이 영역의 변화도를 살펴보기 위해서는 그림 4(c)에서 구한 bicubic으로 업샘플링된 저해상도 깊이 영상을 사용한다. 현재 화소 i 에서 인접 영역 $N_v(i)$ 는 i 를 중심으로 한 변의 길이가 $n+1$ 개의 화소로 되어있는 정방형 영역이다. n 은 업샘플링하고자 하는 스케일의 두 배의 값을 가진다. 이 영역 내에서 bicubic 방법으로 업샘플링 된 깊이 영상의 깊이 값들의 분산을 취하고 그 값을 이용하여 가중치 $w_{d,i}$ 는 식 (8)과 같이 정의된다.

$$w_{d,i} = \exp\left(-\frac{\text{Var}(N_v(i))}{\sigma_d^2}\right) \quad (8)$$

식 (8)에 의해 정의되는 가중치는 영역 $N_v(i)$ 의 분산이 클 경우, 즉 영역이 깊이 불연속이라고 판단될 때 0에 가까운 값을 가지면서 평활화 항의 가중치를 매우 작게 만든다. 반면, 영역 내 깊이 값의 분산이 작을 경우에는 1에 가까운 값을 가지면서 평활화 항이 앞서 구한 가중치 $w_{c,ij}$ 에 의해 조절되도록 하는 효과를 가진다.

IV. 실험 결과 및 분석

실험을 위해 Middlebury 웹사이트에서 제공하는 테스트 영상 Cones, Teddy, Venus를 사용하였다. 그림 7은 각각의 색상 영상과 동일한 해상도의 원본 깊이 영상(ground truth)를 보여준다.

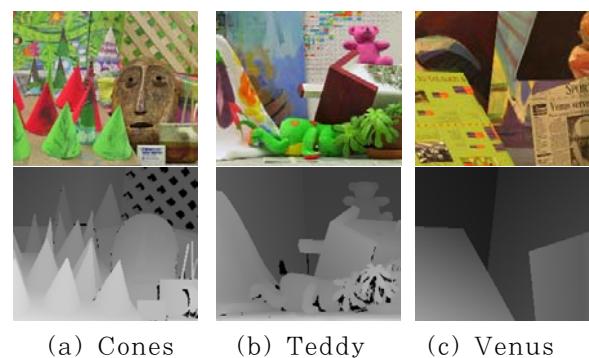


그림 7. 실험에 사용된 영상 및 원본 깊이 영상
Fig. 7. Color and ground truth depth images for test

업샘플링의 실험을 위해 깊이 영상의 원본 해상도에 대하여 각각 2배, 4배, 8배로 다운샘플링을 하였다. 다운샘플링 방법은 n 배 다운샘플의 경우 가로와 세로에 대해 n 번 째 화소마다 한 화소씩 추출하여 다운샘플링 된 영상을 만들었다. 그리고 다운샘플링 된 영상에 대해 원본 영상의 해상도로 업샘플링을 수행한 결과를 표 1에 나타내었다. 업샘플링에 사용된 값 th_d , $\sigma_{c,ij}^2$, σ_d^2 은 각각 10, 50, 10을 사용하였다.

표 1은 업샘플링 된 깊이 영상과 원본 깊이 영상의 각 화소를 비교하였을 때 각 화소의 깊이 값의 차이가 1보다 큰 화소의 전체 대비 비율인 bad pixel rate (BPR)을 나타낸다. 즉, BPR 값이 작을 수록 원본 깊이 영상에 가깝게 업샘플링 되었다고 할 수 있다. 제안하는 방법과 함께 bicubic 업샘플링 방법, MRF 모델을 이용한 방법 [8], 필터 기반인 JBU [6]에 대한 결과도 나타내었다.

표 1. 업샘플링 결과 및 비교

Table 1. Results of depth upsampling

Image	Scale	Bad pixel rate (%)			
		Bicubic	MRF	JBU	Proposed
Cones	2	11.70	7.34	6.25	5.08
	4	21.15	8.78	7.76	6.23
	8	36.79	12.68	11.10	10.98
Teddy	2	9.42	7.74	6.50	6.63
	4	17.08	9.53	7.58	8.58
	8	28.76	14.05	13.50	12.66
Venus	2	2.21	0.77	0.98	0.42
	4	4.48	0.84	1.58	0.50
	8	8.90	1.86	2.51	1.51

표 1의 결과에서 보듯이, 제안하는 방법은 다른 업샘플링 방법에 비해 실험에 사용된 영상의 대부분의 스케일에 대해서 우수한 결과를 나타내었다. 색상 차이의 가중치만을 이용한 기존 MRF 모델 기반의 방법과 비교해 보았을 때, 깊이 불연속 영역을 고려하였기 때문에 경계 영역에서 더 우수한 성능을 나타내었음을 알 수 있다. 그림 8은 사용한 세 종류의 영상에 대하여 제안한 방법으로 업샘플링된 깊이 영상을 보여준다. 그림 9는 표 1에서 비교한 방법들로 4배 업샘플링된 Cones 영상과 제안하는 방법으로 얻은 결과 중 일부를 보여준다. 객체의 경계 영역에서 제안하는 방법이 더 우수한 결과를 나타낸 것을 알 수 있다.

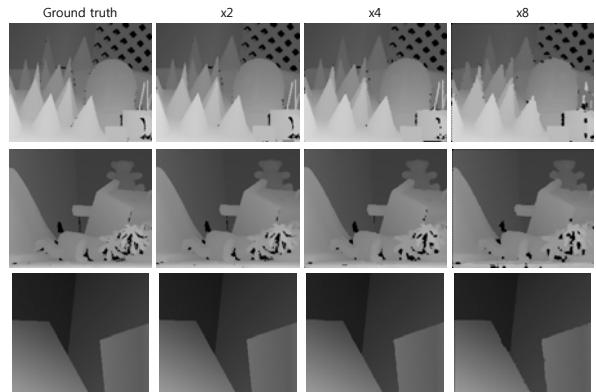


그림 8. 제안하는 방법의 깊이 업샘플링 결과

Fig. 8. Depth upsampling results of the proposed method

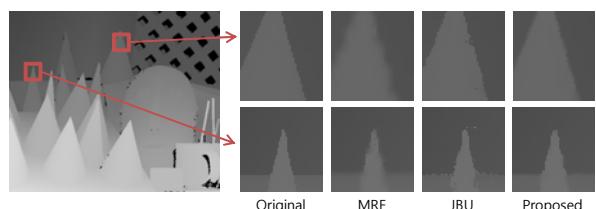


그림 9. Cones 영상에서의 결과 비교

Fig. 9. Comparison of upsampled 'Cones'

V. 결 론

본 논문에서는 MRF 모델에 기반하여 깊이 불연속 영역을 고려한 깊이 영상 업샘플링 방법을 제안하였다. 기존의 MRF 모델 기반이나 필터 기반의 깊이 영상 업샘플링 방법들은 깊이 불연속 영역을 고려하지 않기 때문에 경계 영역이 흐려짐으로 인해 업샘플링의 성능이 저하되는 문제점이 있었다. 제안하는 방법에서는 색상 영상의 경계 영역들을 대상으로 깊이 불연속 영역을 찾고, 해당 영역을 고려하여 에너지 함수의 평활화 항을 위한 색상 차이의 가중치를 결정하고, 또한 깊이 영역의 분산 값을 이용한 가중치로 한번 더 평활화 항의 영향을 조절한다. 그리고 신뢰 확산 방법을 이용하여 에너지 함수를 최적화함으로써, 업샘플링 된 깊이 영상을 원본 깊이 영상과 비교해 보았을 때, 기존의 방법들보다 우수한 성능을 나타내었다.

References

- [1] C. Fehn, R. Barre, and S. Pastoor, "Interactive 3DTV - concepts and key technologies," *Proc. IEEE*, vol. 94, no. 3, pp. 524-538, Mar. 2006.
- [2] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a

- new approach on 3D-TV," *Proc. SPIE Stereoscopic Displays Virtual Reality Syst.*, vol. 5921, pp. 93-104, May 2004.
- [3] A. Smolic and P. Kauff, "Interactive 3-D video representation and coding technologies," *Proc. IEEE*, vol. 93, no. 1, pp. 99-110, Jan. 2005.
- [4] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 787-800, July 2003.
- [5] Y. S. Kang and Y. S. Ho, "Generation of high-resolution disparity map using multiple cameras and low-resolution depth camera," in *Proc. KICS 2012 Fall Conf.*, pp. 287-288, Seoul, Korea, Nov. 2012.
- [6] J. Kopf, M. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral up-sampling," *ACM Trans. Graphics*, vol. 26, no. 3, pp. 1-5, Aug. 2007.
- [7] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, "A noise-aware filter for real-time depth upsampling," in *Proc. ECCV Workshop Multi-camera Multi-modal Sensor Fusion Algorithms Applicat.*, pp. 1-12, Marseille, France, Oct. 2008.
- [8] J. Diebel and S. Thrun, "An application of markov random fields to range sensing," *Advances in Neural Inform. Process. Syst.*, vol. 18, pp. 291-298, Dec. 2006.

강 윤 석 (Yun-Suk Kang)



2007년 한국항공대학교 전자공학과 졸업(학사)

2008년 광주과학기술원 정보통신공학부 졸업(석사)

2008년~현재 광주과학기술원 정보통신공학부 박사과정

<관심분야> 디지털 영상처리, 다시점 영상 획득 및 처리, 3차원 TV, 실감방송

호 요 성 (Yo-Sung Ho)



1981년 서울대학교 전자공학과 학사

1983년 서울대학교 전자공학과 석사

1989년 Univ. of California, Santa Barbara, Department 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, 영상신호 처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV,

MPEG 표준, 3차원 TV, 실감 방송