

고품질의 고해상도 깊이 영상을 위한 컬러 영상과 깊이 영상을 결합한 깊이 영상 보간법

김지현 최진욱 손광훈

연세대학교 전기전자공학과

khsohn@yonsei.ac.kr

Depth Image Interpolation using Fusion of color and depth Information

Kim, Ji-Hyun Choi, Jin-Wook Sohn, Kwang-Hoon

Department of Electrical and Electronic Engineering

Yonsei University

요약

3D 콘텐츠를 획득하는 여러 가지 방법 중 2D-plus-Depth 구조는 다시점 영상을 얻을 수 있는 장점 때문에 최근 이에 관한 연구가 활발히 진행되고 있다. 이 구조를 통해서 고품질의 3D영상을 얻기 위해서는 무엇보다 고품질의 깊이 영상을 구현하는 것이 중요하다. 깊이 영상을 얻기 위해서 Time-of-Flight(ToF)방식의 깊이 센서가 활용되고 있는데 이 깊이 센서는 실시간으로 깊이 정보를 획득할 수 있지만 낮은 해상도와 노이즈가 발생한다는 단점이 있다. 따라서 깊이 영상의 특성을 보존하는 상환 변환을 하여야지만 고품질의 3D 콘텐츠를 제작할 수 있다. 주로 깊이 영상의 해상도를 높이기 위해서 Joint Bilateral Upsampling(JBU) 방식이 사용되고 있다. 하지만 이 방식은 4배 이상의 고 해상도 깊이 영상을 획득하는 데에는 적합하지 않다. 따라서 고해상도의 깊이 영상을 얻기 위해서 보간법을 수행하여 가이드 영상을 만든 후 Bilateral Filtering(BF)을 처리함으로써 영상의 품질을 향상시킨다. 본 논문에서는 2D-plus-Depth 구조에서 얻은 컬러 영상과 깊이 영상을 결합한 보간법을 통해서 깊이 영상의 특성을 살린 가이드 영상을 구현하는 방법을 제안한다. 실험 결과는 제안 방법이 기존 보간법보다 경계 영역 및 평활한 영역에서 깊이 영상의 특성을 잘 보존하는 것을 보여준다.

1. 서론

최근 3차원 영상 산업이 각광받으면서 고품질의 3D 콘텐츠에 대한 요구가 증가하고 있다. 또한 어느 자리에서도 실감 있는 입체감을 느낄 수 있는 영상을 원하고 있다. 입체감을 나타내기 위해서는 깊이 정보가 필요하다. 물체의 깊이 정보를 획득하는 방법으로는 크게 수동적 방식과 능동적 방식으로 구분할 수 있다. 수동적 방식의 대표적인 예로는 여러 카메라로 부터 얻은 영상을 이용하는 스테레오 매칭 등이 있다. 이 방식은 텍스처가 없는 영역, 반복되는 패턴이 있는 영역, 가려진 영역 등에서 문제가 발생한다. 능동적 방식으로는 레이저를 이용하여 물체를 스캔하는 방식, ToF 센서 방식 등이 있다. 능동적 방식은 수동적 방식과는 반대로 텍스처가 있는 영역에서 정확도가 떨어진다. 그리고 ToF 방식의 깊이 센서를 통해서 얻은 깊이 영상은 해상도가 낮고, 반사가 심한 물체나 빛의 양이 많은 야외 등에서 노이즈가 발생한다는 단점이 있다. 하지만 깊이 영상을 실시간으로 획득할 수 있기 때문에 다양하게 활용이 되고 있다. 수동적 방식과 능동적 방식의 단점들은 보완하고 장점들은 수용하기 위해서 두 가지의 방식을 결합하여 깊이 영상의 품질 및 해상도를 개선하는 2D-plus-Depth 구조에 대한 연구가 활발히 진행되고 있다 [1, 4, 6]. 또한 이 구조는 깊이 센서에 여러 대의 Charge-Coupled Device(CCD) 카메라

를 이용하기 때문에 다시점 3D 영상을 얻을 수 있다.

본 논문에서는 이 구조를 바탕으로 깊이 영상과 CCD 카메라에서 얻은 컬러 영상을 결합하여 저해상도의 깊이 영상을 고해상도로 보간 하는 방법을 제안한다.

2. 본론

고품질의 깊이 영상을 구현하기 위해서 깊이 센서와 CCD 카메라를 결합하여 사용한다. ToF방식의 깊이 센서에서는 저해상도의 깊이 영상을 얻지만 CCD 카메라는 상대적으로 높은 해상도의 컬러 영상을 제공하기 때문이다. 고품질의 3D 영상을 제작하기 위해서는 깊이 영상의 특성을 잘 드러나도록 즉, 물체간의 경계는 보존되고 그 외 영역은 평활하도록 깊이 영상을 상환 변환하는 것이 필요하다. 최근 JBU 방식을 이용해서 저해상도의 깊이 영상의 크기를 상향 변환하는 연구들이 진행되고 있다. 하지만 이 방식은 4배 이상의 고해상도 깊이 영상을 얻는 데에는 적합하지 않다. 원하는 깊이 영상의 크기가 커질수록 필요한 정보의 양이 증가하기 때문이다. JBU 방식으로 깊이 영상의 해상도를 높일수록 블록 현상이 발생하고 경계가 드러나지 않는 부정확한 깊이 영상을 얻게 된다[2]. 따라서 [2]에서는 깊이 영상을 살릴 수 있는 보간법을 이용하여 상환 변환한 후 Bilateral Filtering을 통해서 4배 이상의

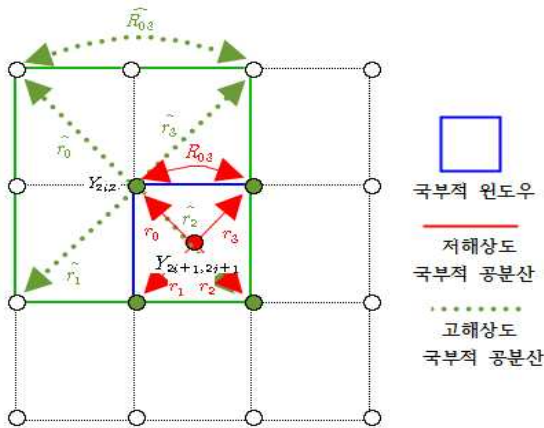


그림1. 기존 NEDI 방법

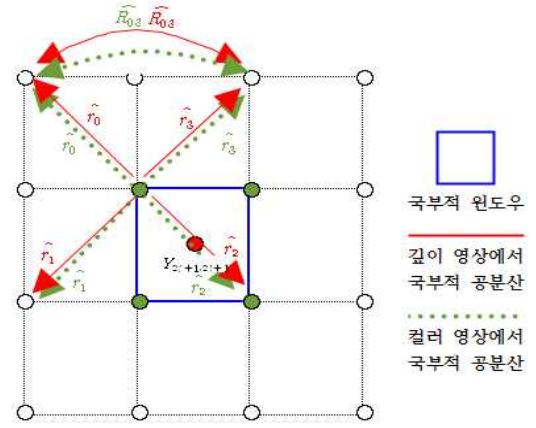


그림2. 제안한 컬러 영상, 깊이 영상 결합 보간법

고해상도 깊이 영상을 얻었다. 하지만 [2]에서 제안한 보간법은 확대 비율이 커질수록 경계가 보존이 되지 않는 문제가 있다. 또한 깊이 영상의 특성 때문에 기존 보간법들을 깊이 영상에 적용할 경우 문제가 발생한다. 따라서 본 논문에서는 깊이 영상의 특성을 살리는 깊이 영상 맞춤 보간법을 제안한다. 제안한 방법은 New Edge Directed Interpolation(NEDI)를 기반으로 고해상도의 컬러 영상의 정보와 저해상도의 깊이 영상의 정보를 결합하여 보간을 수행한다.

NEDI는 기하학적 쌍대성(Geometric duality)을 기반으로 저해상도의 영상으로부터 구한 국부적 공분산을 이용하여 가중치를 계산한 후 이를 이웃한 4개의 화소에 적용하여 고해상도의 영상으로 보간 하는 방법이다[3].

$$Y_{2i+1, 2j+1} = \sum_{k=0}^1 \sum_{l=0}^1 \alpha_{2k+l} Y_{2(i+k), 2(j+l)} \quad (1)$$

여기서 가중치는 classical Wiener filtering theory에 따라서 결정된다.

$$\vec{\alpha} = R^{-1} \vec{r} \quad (2)$$

$R = [R_{kl}], (0 \leq k, l \leq 3)$ 과 $\vec{r} = [r_k], (0 \leq k \leq 3)$ 은 고해상도 영상의 국부적 공분산이다. 저해상도의 화소 값만 알기 때문에 이 값들은 구할 수가 없다. 따라서 기하학적 쌍대성을 기반으로 저해상도의 화소 값으로 저해상도의 국부적 공분산 \hat{R}, \hat{r} 을 구하여 가중치를 구한다. 저해상도의 국부적 공분산은 저해상도 영상에 국부적 윈도우를 적용하여 classical covariance method를 이용하여 구할 수 있다.

$$\hat{R} = \frac{1}{M^2} C^T C, \quad \hat{r} = \frac{1}{M^2} C^T \vec{y}$$

$$y^T = [Y_{2i, 2j} \quad Y_{2i+2, 2j} \quad Y_{2i+2, 2j+2} \quad Y_{2i, 2j+2}]$$

$$C = \begin{bmatrix} Y_{2i-2, 2j-2} & Y_{2i+2, 2j-2} & Y_{2i+2, 2j+2} & Y_{2i-2, 2j+2} \\ Y_{2i-2, 2j} & Y_{2i+2, 2j} & Y_{2i+2, 2j+4} & Y_{2i-2, 2j+4} \\ Y_{2i, 2j-2} & Y_{2i+4, 2j-2} & Y_{2i+4, 2j+2} & Y_{2i, 2j+2} \\ Y_{2i, 2j} & Y_{2i+4, 2j} & Y_{2i+4, 2j+4} & Y_{2i, 2j+4} \end{bmatrix} \quad (3)$$

그림 1에서 행렬 \vec{y} 는 국부적 윈도우 안의 모든 화소 값을 포함하는 데이터 벡터이고, 행렬 C 의 k번째 열벡터는 데이터 벡터 \vec{y} 의

원소 y_k 의 대각선 방향으로 이웃하는 네 개의 이웃 화소로 이루어진다.

$$\vec{\alpha} = (C^T C)^{-1} (C^T \vec{y}) \quad (4)$$

구해진 가중치는 식(1)에 적용하여 고해상도의 화소 값을 구할 수 있다[3]. 이 보간법을 깊이 영상에 적용할 경우 경계 부분에서 주관적으로 우수한 성능을 발휘하지만 평활한 영역에서는 기존의 영상을 왜곡시키는 문제가 있다. 평활한 영역에서는 국부적 윈도우에 속하는 모든 픽셀들이 비슷한 값을 가지기 때문에 가중치를 구하는데 어려움이 있기 때문이다.

이러한 점을 보완하기 위해서 깊이 영상과 컬러 영상을 결합한 깊이 영상 맞춤 보간법을 제안한다. 컬러 영상은 깊이 영상보다 상대적으로 많은 정보를 가지고 있기 때문에 식(4)를 이용해서 가중치 구하는 것이 용이하다. 두 영상을 결합하기 위해서 깊이 영상에서의 경계와 컬러 영상에서의 경계가 동일하다는 가정을 세운다.

$$\vec{\alpha}_c = (C_c^T C_c)^{-1} (C_c^T \vec{y}_c),$$

$$\vec{\alpha}_d = (C_d^T C_d)^{-1} (C_d^T \vec{y}_d) \quad (5)$$

그림 2처럼 NEDI 방법을 기반으로 식(5)과 같이 저해상도인 깊이 영상과 고해상도인 컬러 영상에서 각각 가중치를 구한 뒤 보간된 화소의 값 I_{dw}, I_{cw} 을 식(2)와 같이 구한다. 그리고 경계 영역과 평활한 영역을 깊이 영상의 국부적 윈도우에서 가장 큰 픽셀 값과 가장 작은 픽셀 값의 차이로 구분한 후 영역의 특성에 맞게 I_{dw} 와 I_{cw} 을 결합한다. 경계 영역인 경우, 컬러 영상과 깊이 영상의 경계 영역이 동일하기 때문에 식 (5)와 같이 컬러 영상에서 구한 화소 값과 깊이 영상에서 구한 화소 값을 결합한다.

$$I_i = w I_{dw} + (1-w) I_{cw} \quad (0 \leq w \leq 1) \quad (6)$$

이때 파라미터 w값은 I_{dw} 값을 더 많이 적용할 수 있도록 값을 정한다. 평활한 영역인 경우, 깊이 영상에서는 평활한 영역이지만 컬러 영상에서는 그렇지 않을 수 있다. 또한 앞에서 언급한 것과 같이 가중치 값을 계산할 수 없는 깊이 영상의 평활한 영역에서 구한 I_{dw} 값은 정확하지 않을 수 있다. 이러한 점들을 모두 고려하기 위해서 I_{dw} 와 I_{cw} 의 차이 값을 비교하여 보간 할 화소의 값을 계산한다.

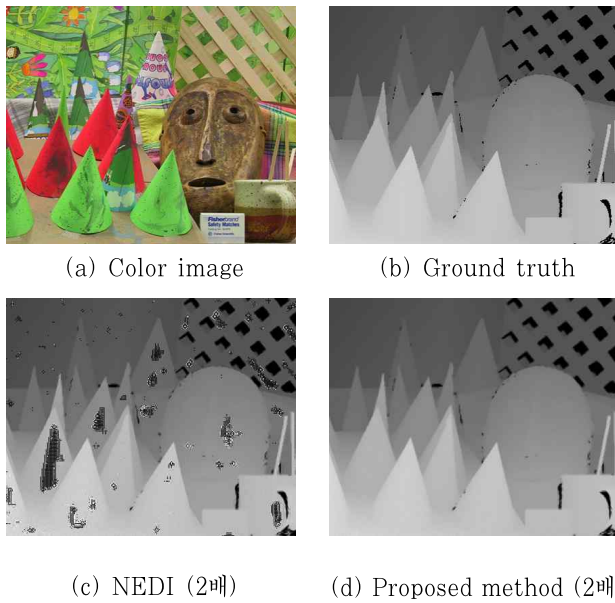


그림3. 깊이 영상에서의 NEDI와 제안한 방법 비교

$$d = |I_{dw} - I_{cw}| \text{ 일 때,}$$

$$I_d = \begin{cases} wI_{dw} + (1-w)I_{cw} & (d < th_1) \\ wI_{dw} + (1-w)I_{cw} & (th_1 \leq d < th_2) \\ wI_{cw} + (1-w)I_{cw} & (d \geq th_2) \end{cases} \quad (7)$$

이 때 파라미터 w 의 값을 작게 설정하여 I_{cw} 값이 더 많이 적용될 수 있도록 한다. 또한 파라미터 th_1, th_2 의 값을 실험적으로 정해진다. 깊이 영상을 상향 변환하는 것이기 때문에 I_{dw} 와 I_{cw} 의 차이 값이 작을 경우 깊이 영상에서 구한 I_{dw} 값으로 보간 한다. I_{dw} 값이 정확하지 않는 경우 I_{dw} 와 I_{cw} 의 차이 값이 크기 때문에 컬러 영상에서 구한 I_{cw} 값으로 보간 한다.

3. 구현 및 결과

제안한 방법은 다음과 같은 환경에서 수행되었다. 모든 방법들은 Visual Studio 2008에서 구현되었으며, 3D 영상 처리 분야에서 주로 사용되는 “Cone”영상을 이용하였다. 원본 영상은 448*368 해상도로 원본 영상의 가로, 세로 크기의 1/2로 축소된 영상을 기존의 NEDI 방법과 제안한 방법으로 보간 하였다. 그림 2의 (a)에서 NEDI 방식으로 깊이 영상을 상향 변환을 한 경우 평활한 영역에서 정확한 값을 갖지 못하는 것을 확인할 수 있다. 반면에 본 논문에서 제안한 NEDI방식을 기반으로 컬러 영상과 깊이 영상을 결합한 방법은 경계 영역과 평활한 영역 모두가 정확하게 상향변환이 되었다는 점을 그림 2의 (b)에서 확인할 수 있다. 그리고 원본 영상의 가로, 세로 크기의 1/4로 축소된 영상을 [2]의 보간법과 본 논문에서 제안한 컬러 영상과 깊이 영상을 결합한 깊이 영상 맞춤 보간법으로 확대하였다. 그림 3의 왼쪽 영상들은 [2]와 본 논문에서 제안한 방법으로 상향 변환을 한 것이고, 오른쪽 영상은 왼쪽 영상의 한 부분을 확대한 것이다. 그림 3의 (a)와 (b)를 비교하면

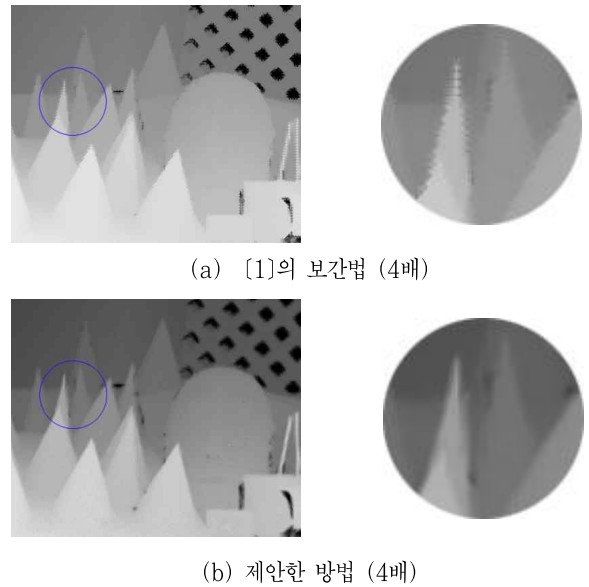


그림4. [2] 보간법과 제안한 방법 비교

[2]의 방법보다 본 논문에서 제안한 방법이 경계 영역을 잘 보존한다는 점을 확인할 수 있다.

4. 결론 및 향후 연구 방향

본 논문에서는 NEDI 방식을 기반으로 컬러 영상과 깊이 영상을 결합한 보간법을 제안하였다. 기존 NEDI를 깊이 영상에 적용할 경우 평활한 영역에서 문제가 발생하였지만 제안한 방법으로 깊이 영상을 상향 변환을 경계와 평활한 영역 모두가 보존이 되는 결과 값을 얻을 수 있었다. 또한 [2]의 보간법보다 정확하게 경계를 보존하였다. 하지만 본 논문에서 제안한 보간법은 파라미터의 수가 많아 복잡하고 확대 비율이 높아질수록 정확도가 다소 떨어지는 문제점이 있다. 따라서 이를 보완하기 위한 방안들을 연구할 계획이다.

5. 참고문헌

- [1] J. Zhu, L. Wang, R. Yang, J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," *IEEE Proc. CVPR*, 2008.
- [2] 김지현, 최진욱, 손광훈, "경계 보존 보간법을 이용한 깊이 영상의 해상도 및 품질 개선," 2011년도 한국방송공학회 하계 학술대회.
- [3] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1571 - 1527, Oct. 2001.
- [4] J. Choi, D. Min, B. Ham and K. Sohn, "Spatial and temporal up-conversion technique for depth video," *in Proc. IEEE Conf Image Processing*, pp. 3525-3528, 2009.
- [5] J. Kopf, MF. Cohen, D. Lischinski and M. Uyttendaele, "Joint bilateral upsampling," *ACM SIGGRAPH*, 2007.