

## 고품질 및 고해상도 깊이 영상 구현을 위한 새로운 결합 보간법

김지현    최진욱    류승철    김동현    손광훈

연세대학교 전기전자공학과

khsohn@yonsei.ac.kr

### The new fusion interpolation for high resolution depth image

Kim, Jihyun    Choi, Jinwook    Ryu, Seungchul    Kim, Donghyun    Sohn, Kwanghoon

School of Electrical and Electronic Engineering

Yonsei University

## 요약

3차원 영상 기술은 방송, 영화, 게임, 의료, 국방 등 다양한 기존 산업들과 융합하며 새로운 패러다임을 형성하고 있으며, 고품질 및 고해상도의 3차원 영상 획득에 대한 필요성이 강조되고 있다. 이에 따라, 최근에는 3차원 입체 영상을 제작 하는 방법 중 하나인 2D-plus-Depth 구조에 대한 연구가 활발히 진행되고 있다. 2D-plus-Depth 구조는 Charge-Coupled Device(CCD) 센서 등을 이용한 일반 카메라와 깊이 카메라를 결합한 형태로써 이 구조로부터 얻은 깊이 영상의 해상도를 상향 변환하기 위해서 Joint Bilateral Upsampling(JBU)[1], 컬러 영상의 정보를 활용한 보간법[2] 등의 방법들이 사용된다. 하지만 이 방법들은 깊이 영상을 높은 배율로 상향 변환할 경우 텍스처가 복사되거나 흐림 및 블록화 현상이 발생하는 문제점이 있다. 본 논문에서는 2D-plus-Depth 구조에서 얻은 고해상도 컬러 영상에서 보간 정보를 구하고 이 정보를 저해상도의 깊이 영상에 적용하여 상향 변환된 가이드 깊이 영상을 제작한다. 이 가이드 깊이 영상을 Bilateral Filtering[8]을 이용함으로써 고품질의 고해상도 깊이 영상을 획득한다. 실험 결과 제안하는 방법으로 해상도를 상향 변환을 할 경우에 기존의 보간법들에 비해 깊이 영상의 특성을 잘 보존함을 확인할 수 있고, 가이드 깊이 영상에 필터링을 처리한 결과가 JBU의 결과보다 향상됨을 확인할 수 있다.

## 1. 서론

3차원 입체 영상은 텔레비전, 핸드폰 등의 제품들과 방송, 영화, 게임, 의료 등 기존 산업들에 응용되면서 차세대 신성장 동력 산업으로 급부상하고 있다. 최근에는 3D 텔레비전, 3D 핸드폰, 3D 게임기가 대중화되면서 3차원 입체감을 더욱 생생하게 체험할 수 있는 양질의 3D 콘텐츠에 대한 요구가 증가하고 있다. 3차원 영상에서 입체감을 느낄 수 있는 이유는 깊이 정보가 존재하기 때문인데 깊이 정보가 정확할수록 입체감을 더욱 효과적으로 느낄 수 있다. 또한 3D 디스플레이의 해상도가 확대되고 있기 때문에 이에 맞는 고해상도 깊이 영상이 필요하다. 깊이 정보를 획득하는 방법은 크게 두 가지로 나눌 수 있다. 여러 대의 카메라로부터 획득한 영상들을 이용해서 깊이 정보를 추정하는 스테레오 매칭(Stereo matching)등의 수동적 방식과 깊이 카메라로 직접 물체의 깊이 정보를 획득하는 능동적 방식이 있다. 수동적 방식의 경우 텍스처가 없는 영역, 반복되는 패턴이 있는 영역, 가려진 영역 등에서 깊이 정보의 오류가 발생하고, 계산 비용도 크다는 단점이 있다. 능동적 방식에서는 레이저 스캐닝, Time of Flight(ToF) 센서 등을 이용하여 깊이 영상을 직접 획득한다. 주로 레이저 스캐닝 방법보다 비용이 저렴하고 동적인 환경에서 사용이 가능한 ToF 센서가 많이 활용되고 있다. 능동적 방식은 수동적 방식에서 문제점이 발생하는 부분에서

도 정확한 깊이 정보를 얻을 수 있고, 실시간으로 획득이 가능하다는 장점이 있다. 하지만 반사가 심한 영역이나 야외에서는 정확도가 떨어지고 획득되는 깊이 영상의 해상도가 낮다는 단점이 있다. 이런 수동적 방식과 능동적 방식의 단점들을 상호 보완하기 위해서 깊이 카메라와 CCD 카메라 같은 일반 카메라를 결합한 2D-plus-Depth 구조 형태가 활발히 연구되고 있다[5][6].

본 논문에서는 2D-plus-Depth 구조를 기반으로 일반 카메라에서 얻은 컬러 영상과 깊이 카메라에서 얻은 깊이 영상을 결합한 보간법을 제안한다. 제안한 보간법으로 고해상도 가이드 깊이 영상을 제작한 후 Bilateral Filtering을 통해 고품질의 깊이 영상을 획득한다.

## 2. 본론

컬러 영상이 색상 정보가 다양하고 물체의 텍스처가 많은 반면, 깊이 영상은 물체의 형태만을 나타내며 평활한 영역이 대부분이라는 특성을 갖는다. 즉, 깊이 영상은 물체와 물체간의 경계 영역과 물체의 평활한 영역으로 구분되는데 이 영역들을 얼마나 잘 보존하느냐에 따라서 깊이 영상의 품질이 결정된다. 따라서 깊이 영상을 상향 변환할 때에는 깊이 영상의 특성을 보존하는 것이 중요하다. 영상의 해상도를 상향 변환하는 방법 중에는 보간법이 있다. 보간법은 선형 보간법인 이웃

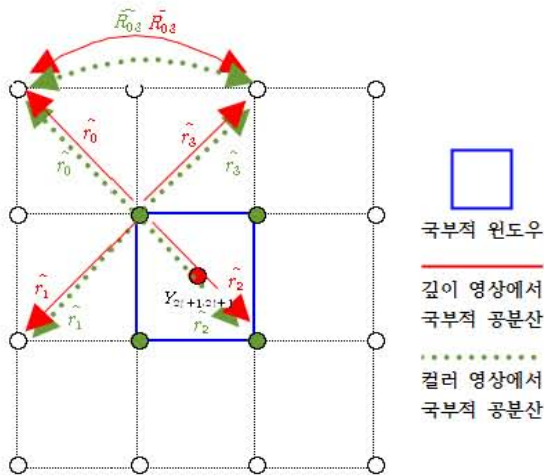


그림1. [2]에서의 컬러 영상과 깊이 영상의 결합 보간법

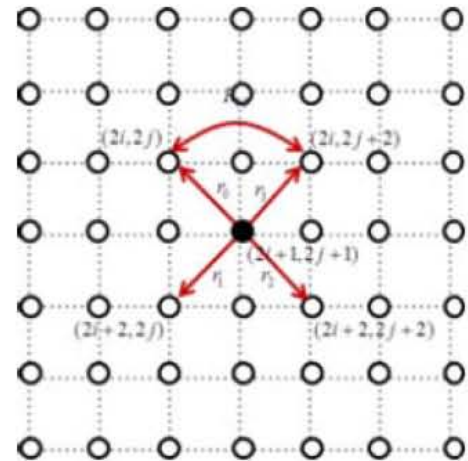


그림2. 제안하는 새로운 결합 보간법

보간법(Nearest Neighbor Interpolation), 양선형 보간법(Bilinear Interpolation), 고등차수 보간법(Bicubic Interpolation)[7]등이 대표적이다[9]. 이 밖에도 경계 영역의 방향을 고려한 공분산 기반 적용 보간법[3], 영상의 기울기 값을 이용해서 상황 변환하는 방법[4] 등 국부적 영역의 특징을 이용한 비선형 보간법이 있다. 깊이 영상의 정보가 한정적이기 때문에 이런 보간법들을 깊이 영상에 적용할 경우 일반 영상에 적용했을 때보다 성능이 떨어진다. [2]에서는 그림 1과 같이 컬러 영상과 깊이 영상에서 공분산 기반 적용 보간법[3]을 기반으로 구한 공분산 값들을 결합하여 깊이 영상의 상황변환에 적합한 깊이 영상 보간법을 제안하였다. 공분산 기반 적용 보간법은 경계 영역의 정보를 활용하여 영상의 해상도를 상황 변환하는 대표적인 보간법이다. 경계 영역의 정보는 Wiener filtering theory에 따라서 공분산을 이용해서 구하고 그 값을 식 (1)과 같이 가중치  $\alpha$ 로 표현한다.

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

$R = [R_{kl}], (0 \leq k, l \leq 3)$ 과  $\vec{r} = [r_k], (0 \leq k \leq 3)$ 은 고해상도의 국부적 공분산 행렬들이다. 저해상도 깊이 영상만 있을 경우  $R, r$ 을 구할 수 없기 때문에 기하학적 쌍대성(Geometric duality)을 기반으로 저해상도의 영상으로부터 공분산 행렬들을 식 (2)와 같이 계산한다.

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

$\vec{y}$ 는 국부적 윈도우 안의 모든 화소 값을 포함한 벡터이고, 행렬  $C$ 의  $k$ 번째 열벡터는 데이터 벡터  $\vec{y}$ 의 원소  $y_k$ 와 대각선 방향으로 이웃한 화소들이다. 식 (2)에서 구한 공분산 행렬들을 식 (1)에 적용하여 가중치  $\alpha$ 를 구한다. 이렇게 구한 가중치들은 식 (3)에서 채워야 할 화소의 값을 계산하는데 활용된다. 식 (3)에서  $Y^d$ 는 깊이 영상에서의 화소 값이다.

$$Y_{2i+1, 2j+1}^{depth} = \sum_{k=0}^1 \sum_{l=0}^1 \alpha_{2k+l} Y_{2(i+k), 2(j+l)}^d \quad (3)$$

$$Y_D = \begin{cases} w Y_{dw}^{depth} + (1-w) Y_{dw}^{depth} & (d < th_1) \\ w Y_{dw}^{depth} + (1-w) Y_{cw}^{depth} & (th_1 \leq d < th_2) \\ w Y_{cw}^{depth} + (1-w) Y_{cw}^{depth} & (d \geq th_2) \end{cases} \quad (4)$$

where,  $d = |Y_{dw}^{depth} - Y_{cw}^{depth}|$

[2]에서는 깊이 영상과 컬러 영상에서 식 (1), (2)처럼 각각의 공분산으로 구한 가중치들로 식 (3)과 같이 보간된 화소 값  $Y_{dw}^{depth}, Y_{cw}^{depth}$ 를 얻는다. 평활한 영역과 경계 영역을 구분하여 평활한 영역에서는 깊이 영상에서 구한 보간 값이, 경계 영역에서는 컬러 영상에서 구한 보간 값이 더 많이 적용되도록 식 (4)와 같이 보간 값들을 결합한다. 경계 영역의 방향 특성을 고려해서 빈 화소들을 채워주기 때문에 경계 영역을 더 정확하게 보존한다는 장점이 있다[2][3]. 하지만 [2]에서 제안한 방법은 보간하기 위해 필요한 파라미터의 수가 많고, 컬러 영상과 깊이 영상에서 공분산을 구하기 때문에 계산량이 많다는 단점이 있다.

이를 개선하기 위해서 새로운 결합 보간법을 제안한다. 본 논문에서는 고해상도 컬러 영상과 저해상도 깊이 영상의 경계 영역이 동일하다는 가정 하에 알고리즘을 제안한다. 새로운 결합 보간법은 일반 컬러 영상에서 경계 영역의 방향 정보를 획득하고 이 정보를 깊이 영상의 정보에 결합한 것이다. 그림 2와 같이 보간 가중치  $\alpha$ 를 구하기 위해서 식 (1)의 국부적 공분산 행렬  $R, r$ 을 고해상도인 컬러 영상에서 구한다.

$$R_{03} = \frac{1}{(2T+1)^2} \sum_{k=i-T}^{i+T} \sum_{l=j-T}^{j+T} Y_{2k, 2l}^c Y_{2k, 2l+2}^c$$

$$r_0 = \frac{1}{(2T+1)^2} \sum_{k=i-T}^{i+T} \sum_{l=j-T}^{j+T} Y_{2k, 2l}^c Y_{2k+1, 2l+1}^c \quad (5)$$

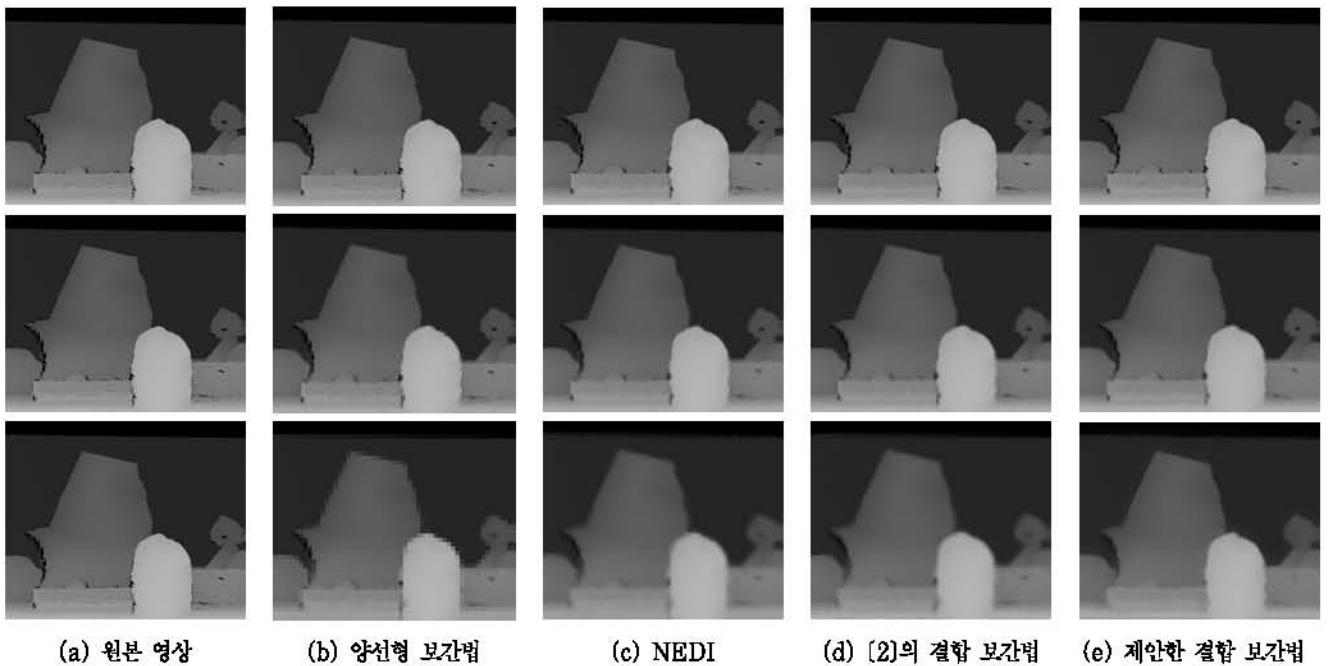


그림3. 양선형 보간법, NEDI, [2]의 결합 보간법과 제안한 새로운 결합 보간법의 결과 비교

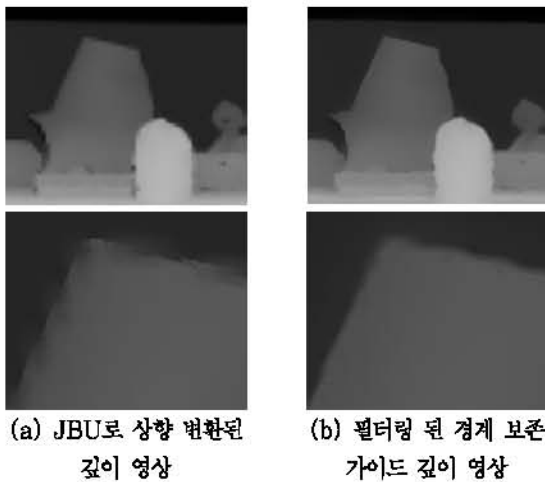


그림 4. JBU와 제안한 방법의 64배 상향 변환 비교

식 (5)에서  $Y^c$ 는 컬러 영상에서의 화소 값이고,  $R_{03}$ 은 식 (1)의 행렬  $R$ 의 원소,  $r_0$ 은 행렬  $r$ 의 원소이다. 컬러 영상의 경우 상향 변환할 깊이 영상에 비해 해상도가 크기 때문에 기하학적 쌍대성을 기반으로 공분산 값들을 계산하지 않아도 된다. 식 (5)와 같이 공분산 행렬  $R, r$ 의 원소들을 간단하게 계산 후 이 값들을 식 (1)에 대입하여 보간 가중치  $\alpha$ 를 계산한다. 고해상도의 컬러 영상은 경계 영역에 대한 방향과 보간 값을 예측하는 가중치를 계산할 정보가 많기 때문에 공분산 기반 적응 보간법을 [3] 기반으로 구하는 방법들 보다 정확한 가중치 값을 계산한다. 상향 변환 배율이 높아져도 컬러 영상에서 정확한 가중치를 구할 수 있기 때문에 확대된 깊이 영상의 품질이 향상된다. 깊이 영상의 상향변환에서 중요한 점은 경계 영역과 평활한 영역을 보존해주는 것이다. 경계 영역은 뚜렷하게 보존하는 것이 중요하지만 평활한 영역의 경

우 매끄럽게 보존하여야 한다. 따라서 [3]에서 계산의 복잡성을 낮추기 위해 제안한 방법처럼 본 논문에서도 경계 영역에서는 새로운 결합 보간법을 적용하고 평활한 영역은 양선형 보간법으로 처리한다. 영역의 구분은 깊이 영상에서 보간될 화소의 이웃한 네 개의 픽셀들의 분산으로 결정한다. 분산이 일정한 기준치 보다 클 경우 경계 영역으로 구분하고, 그렇지 않은 경우 평활한 영역으로 정의 한다. 구분된 영역에 따라 각각의 보간법으로 빈 홀들을 채워 경계 보존된 가이드 깊이 영상을 만든다. 영상의 품질을 향상시키기 위해서 가이드 깊이 영상에 Bilateral Filtering을 처리한다. 제안한 새로운 결합 보간법은 [2]에서 제안한 방법과 다르게 경계 영역과 평활한 영역을 구분하는 파라미터와 공분산을 계산하기 위한 윈도우 크기만 고려하면 된다. 컬러 영상과 경계 영역에서만 공분산을 계산하기 때문에 복잡도가 낮아진다는 장점이 있다.

### 3. 구현 및 결과

본 논문에서는 Middlebury의 "Midd2" 컬러 영상과 깊이 영상을 사용하였다. 제안한 알고리즘의 성능을 알아보기 위해서 양선형 보간법, NEDI, [2]의 결합 보간법, JBU로 저해상도의 깊이 영상을 전체 4배, 16배, 64배로 상향 변환한 결과 영상을 비교 하였다. 영상들의 원본 크기는 448\*368로 확대 배율에 따라 축소하여 사용하였다. 그림 3에서 첫 번째 행은 가로 방향과 세로 방향으로 각각 2배씩 (4배) 확대한 결과 영상이고, 두 번째 행은 4배씩 (16배) 그리고 세 번째 행은 8배씩 (64배) 확대한 영상이다. 결과 영상을 주관적으로 살펴보면, 양선형 보간법의 경우 평활한 영역은 상향 변환 배율이 높아져도 보존이 되지만 경계 영역에서는 블러화 현상이 심해짐을 확인할 수 있다. NEDI는 4배와 16배로 상향 변환을 할 때에는 경계 영역을 매끄럽게 보존해주지

만 64배의 높은 배율로 상향 변환을 하는 경우에는 경계 영역에서 흐림 현상이 두드러지게 나타난다. [2]의 보간법은 양선형 보간법에서 나타나는 블러화 현상이 나타나지 않고, 64배 상향 변환에서는 NEDI보다 흐림 현상이 줄어들었다. 하지만 확대 배율이 높아질수록 흐림 현상이 증가되는 것은 문제가 된다. 반면에 본 논문에서 제안한 새로운 결합 보간법은 깊이 영상의 특성을 잘 보존하였다. 64배로 상향 변환을 하여도 양선형 보간법에서 생긴 블러화 현상이 나타나지 않고, NEDI와 논문 [2]에서 생기는 흐림 현상이 줄어들었음을 확인할 수 있다. 그림 4의 (a)는 JBU로 깊이 영상을 64배 상향 변환한 결과 값이고, 그림 4의 (b)는 제안한 새로운 결합 보간법으로 만든 가이드 깊이 영상을 필터링한 결과이다. JBU로 상향 변환된 영상에서는 텍스처가 묻어나고 경계 영역에서 블러화 현상이 발생하는 것을 볼 수 있다. 반면에 제안한 보간법으로 만든 가이드 깊이 영상에 필터링한 결과 영상은 경계가 주관적으로 향상되었음을 확인할 수 있다. 즉, 깊이 영상을 고평배율로 상향 변환을 할 경우 사용할 정보가 부족하기 때문에 JBU방식을 사용하기 보다는 우선적으로 경계 보존 가이드 영상을 제작하고 이후에 필터링을 처리하는 것이 고품질 및 고평배율도 깊이 영상을 구현하기에 적합하다는 것을 알 수 있다.

#### 4. 결론 및 향후 연구 방향

본 논문에서는 고평배율의 컬러영상을 이용해서 저해상도의 깊이 영상을 상향 변환하는 알고리즘을 제안하였다. 기존 보간법들을 깊이 영상에 적용할 경우 경계 영역에서 블러화 현상과 흐림 현상이 발생하여 깊이 영상의 특성을 잘 살리지 못하는 문제점이 발생하였다. [2]에서 제안하는 방법은 파라미터의 수가 많아 복잡도가 크다는 단점이 있다. 하지만 본 논문에서 제안한 새로운 결합 보간법은 컬러 영상에서 정확한 경계 영역의 정보들을 활용하기 때문에 경계 영역과 평활한 영역 모두를 보존한다. 높은 배율로 확대할 경우에도 다른 방법들에 비해 주관적 성능이 뛰어난 것을 확인하였다. 또한 [2]의 방법보다 필요한 파라미터의 수가 줄어들므로 복잡성이 낮아졌다. 하지만 높은 배율로 상향 변환을 할 경우 흐림 현상이 어느 정도 발생하는 문제점이 있다. 따라서 높은 주파수를 일정하게 유지할 수 있는 방안들을 연구할 계획이다.

#### 5. 참고문헌

- [1] J. Kopf, M.F. Cohen, D. Lischinski and M. Uyttendaele, "Joint bilateral upsampling," *ACM SIGGRAPH*, 2007.
- [2] 김지현, 최진욱, 손광훈, "고품질의 고평배율도 깊이 영상을 위한 컬러 영상과 깊이 영상을 결합한 깊이 영상 보간법," 2011년도 한국방송공학회 추계 학술대회, 2011
- [3] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1571 -1527, 2001.
- [4] J. Sun, Z. Xu, and H. Shum, "Image super-resolution using gradient profile prior," in *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2008.
- [5] 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.
- [6] 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.
- [7] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust. Speech Signal Process.* 1978.
- [8] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, 1998.
- [9] T.M. Lehmann, C. Gonner, K. Spitzer, "Survey: interpolation methods in medical image processing", *IEEE Transactions on Medical Imaging*, 1999.