일반논문 (Regular Paper)
방송공학회논문지 제25권 제6호, 2020년 11월 (JBE Vol. 25, No. 6, November 2020)
https://doi.org/10.5909/JBE.2020.25.6.965
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

# 반복적인 격자 워핑 기법을 이용한 깊이 영상 초해상화 기술

김 동 신[a)], 양 윤 모[a)], 오 병 태[a)‡]

# Iterative Deep Convolutional Grid Warping Network for Joint Depth Upsampling

Dongsin Kim[a)], Yoonmo Yang[a)], and Byung Tae Oh[a)‡]

## 요 약

깊이 영상은 물체와의 거리 정보를 가지고 있다. 이는 3D 정보를 구성하는데 중요한 역할을 한다. 보통 같은 시점에서 얻은 컬러 영상과 깊이 영상을 함께 사용한다. 그런데 하드웨어 기술의 한계로 인해 깊이 영상은 쌍을 이루는 컬러 영상에 비해 낮은 해상도를 갖는다. 따라서 일반적으로 깊이 영상을 사용할 때 영상의 해상도를 컬러 영상의 해상도에 맞게 업샘플링을 진행한 후 사용한다. 본 논문에서는 깊이 영상의 해상도를 높이기 위해 화소 값을 개선시키는 일반적인 방법이 아닌 화소의 위치를 이동시키는 방법을 제안한다. 제안하는 기법에서는 화소의 위치를 경계 주변에서 경계 중앙으로 이동시키며 이 과정을 여러 단계에 걸쳐 진행하여 블러된 영상을 복원한다. 실험 결과를 통해 제안하는 방법이 기존 방법들에 비해 정량적, 시각적 품질을 모두 개선시켰음을 알 수 있다.

## Abstract

Depth maps have distance information of objects. They play an important role in organizing 3D information. Color and depth images are often simultaneously obtained. However, depth images have lower resolution than color images due to limitation in hardware technology. Therefore, it is useful to upsample depth maps to have the same resolution as color images. In this paper, we propose a novel method to upsample depth map by shifting the pixel position instead of compensating pixel value. This approach moves the position of the pixel around the edge to the center of the edge, and this process is carried out in several steps to restore blurred depth map. The experimental results show that the proposed method improves both quantitative and visual quality compared to the existing methods.

Keyword : Depth map upsampling, Joint filtering, Convolutional neural network

a) 한국항공대학교 항공전자정보공학부(Korea Aerospace University)
‡ Corresponding Author : 오병태(Byung Tae Oh)
E-mail: byungoh@kau.ac.kr
Tel: +82-2-300-0409
ORCID: https://orcid.org/0000-0003-1437-2422

※ 본 연구는 2020년도 정부(교육부)의 재원으로 한국연구재단 기초연구사업(NRF-2019R1F1A1063229)과 경기도 지역협력 연구센터 사업 (GRRC) (2020-B02, 3차원 공간 데이터 처리 및 응용기술 연구)의 지원을 받아 수행되었음.
※ This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1F1A1063229), and by the GRRC program of Gyeonggi Province [2020-B02, Study on Study on 3D point cloud processing and application technology].
· Manuscript received September 14, 2020; Revised October 26, 2020; Accepted October 26, 2020.

## Ⅰ. 서 론

3D 산업이 발전함에 따라 로봇 분야 혹은 자율주행 분야 등에서 다양한 방식의 3D 기술들이 시도되고 있다. 깊이 정보는 이러한 3D 기술에서 내외부적으로 중요한 역할을 한다. 깊이 영상을 얻기 위해서는 크게 수동적 방법과 능동적 방법이 존재한다. 수동적 방법이란 깊이 정보를 간접적으로 얻는 것을 뜻한다. 이 방식의 대표적인 방법으로 스테레오 매칭 방식이 존재하며, 이 방법은 동일한 사물에 대한 시점이 다른 두 영상 간의 쌍안 시차를 통해 깊이 정보를 얻는다. 반대로 능동적인 방법은 레이저 거리계나 Time of Flight (TOF), 구조광 카메라와 같은 특별한 장치를 통해 깊이정보를 직접적으로 얻는다. Microsoft Kinect 와 SoftKinect는 이러한 방식을 사용해 깊이 영상을 얻는 대표적인 카메라이다. 그러나 일반적으로 깊이 영상을 획득했을 때 하드웨어 기술의 한계로 인해 그 해상도나 품질이 컬러 영상에 비해 크게 떨어지는 문제점이 있다.

이러한 문제를 해결하기 위한 대표적인 방법으로는 결합(joint) 필터 방식이 있다. 결합 필터의 경우 컬러 영상의 경계정보와 깊이 영상의 경계정보가 유사하다는 가정에서부터 시작한다. 이 방법은 업샘플링에 사용되는 필터를 깊이 영상과 쌍을 이루는 컬러 영상을 통해서 얻는다. 그렇기 때문에 컬러 영상의 경계 값에 대한 구조적인 정보를 효율적으로 깊이 영상으로 전달할 수 있다.

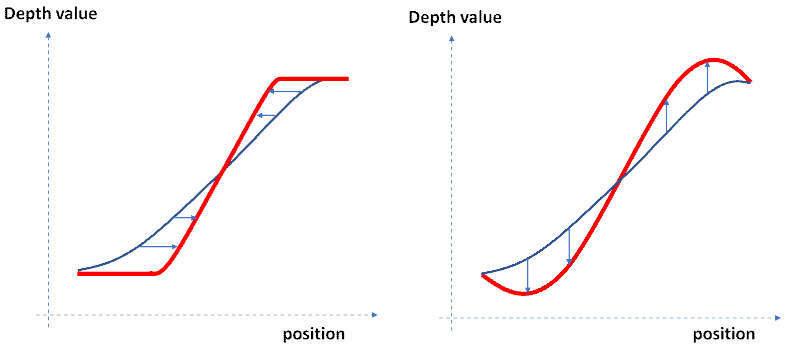
이러한 방법들은 정보를 전달하는데 있어서 보통 두가지 문제가 존재한다. 첫번째로 경계 값들의 구조적인 정보를 전달할 때 깊이 영상에는 존재하지 않고 컬러 영상에만 존재하는 정보들 역시 전달하게 된다. 그런데 커널을 통해 계산하는 방법은 텍스쳐 패턴 같은 화소 값의 작은 변화에 민감하다. 두번째로 커널을 부정확하게 예측함으로써 생기는 오버슈팅과 언더슈팅 현상이 빈번하게 발생한다.

따라서 본 논문에서는 이러한 문제를 해결하기 위해서 픽셀 화소 값을 보상하는 커널 필터링 방식이 아닌 변위 벡터를 통해 픽셀 값의 위치를 변환하는 방식을 제안한다. 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 변위 벡터를 통한 워핑 (warping) 방식에 대해 구체적으로 기술하고, 3장에서는 제안 기술의 성능을 실험을 통해 확인한다. 마지막으로 4장에서는 본 논문에 대한 결론을 맺는다.

## Ⅱ. 제안 기술

### 1. 워핑 방식을 통한 영상 복원

기존 초해상화 기술의 경우 일반적으로 커널을 이용한다. 커널 기반 방법의 경우 그림 1(b)처럼 해당 위치의 왜곡된 픽셀 값을 보상함으로써 그 값을 복원한다. 하지만 이 경우 이상적인 커널을 예측하기 어렵기 때문에 경계 근처에서 오버슈팅 혹은 언더슈팅 현상이 빈번하게 일어나는 문제점이 있다. 반면에 워핑 방식에서는 이러한 현상을 효율적으로 방지할 수 있다.




(a) (b)

그림 1. 블러된 픽셀(파랑)을 블러되지 않은 픽셀(빨강)로 복원하는 방법 비교: (a) 픽셀 위치 이동 (b) 픽셀 값 보상

Fig. 1. Comparison of signal reconstruction from blurred (blue) to deblurred (red): (a) pixel position shift (b) pixel value modification

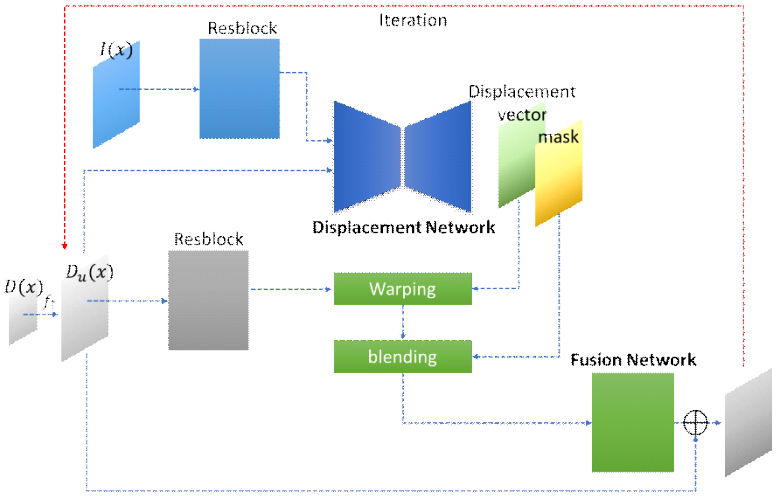



그림 2. 세부 시스템 구조도
Fig. 2. Detailed system structure

영상 복원에 있어서 워핑 방식을 사용하는 논문은 대표적으로 [1-3]이 존재한다. 이러한 접근 방식들은 블러링 과정 중 실제 경계 값 주변의 픽셀 값들을 경계 바깥쪽으로 이동시킴으로써 왜곡을 발생시킨다는 가정을 기반으로 한다. 그러므로 블러링 문제를 해결하기 위해서 멀리 떨어진 픽셀 값들을 그림 1(a)처럼 경계 값 주변으로 이동시키는 방식을 사용한다. 이 과정에서 변위벡터를 계산해야 하며, 이상적인 변위벡터를 구하지 못한 경우에 커널 방식과 동일하게 오버슈팅 혹은 언더슈팅 현상이 일어날 수 있다. 하지만 변위벡터의 경우 그 값을 직접적으로 제한할 수 있고, 컬러 영상에서 경계 정보를 직접적으로 가져와 사용할 수 있기 때문에 일반적인 커널 기법과 달리 그러한 문제를 효율적으로 방지할 수 있다.

이러한 기술을 초해상화에 사용하기 위해서 먼저 저해상도 영상을 쌍선형 보간법과 같은 전통적인 방식을 사용해 고해상도 영상으로 변환한다. 변환된 고해상도 영상은 일종의 블러된 영상이기 때문에 해당 영상에 워핑 기법을 적용해 고품질의 고해상도 영상을 획득한다.

## 2. 반복 워핑 기반 제안 기술

제안하는 방식에서는 격자 워핑 방식을 깊이 영상 업샘플링에 적용한다. 전통적인 격자 워핑 방식은 블러된 영상을 복원하기 위해 만들어졌다. 따라서, 격자 워핑 기술을 사용해 초해상화를 하기 위해 먼저 블러된 영상을 획득한다. 블러된 영상의 경우 목표로 하는 해상도에 맞게 원본 깊이 영상을 업샘플링해 획득하며, 해당 깊이 영상에 워핑 기법을 적용해 복원 과정을 진행한다.

워핑 방식을 통해 블러된 영상을 복원하는데 있어서, 워핑에 사용되는 변위 벡터를 정확하게 계산하는 것은 매우 중요하다. 기존 격자 워핑 방식처럼 블러된 깊이 영상만을 사용해 복원을 진행할 경우 변위 벡터의 방향과 크기를 정확하게 계산하기 어렵다. 따라서 제안 기술에서는 변위벡터를 깊이 영상만을 사용해 계산하는 것이 아니라 주어진 고해상도 컬러 영상을 같이 사용하며, 고해상도 컬러 영상에는 블러되지 않은 경계 정보가 존재하기 때문에 이를 통해 변위벡터의 방향과 크기를 효율적으로 계산할 수 있다.

$$D_u(x) = f_t(D(x)) \tag{1}$$

그림 2는 제안 시스템의 전반적인 구조를 보여주고 있다. 먼저 저해상도 깊이 영상 를 고해상도 컬러 영상 와 같은 해상도를 가질 수 있도록 업샘플링 하여 를 만든다. 이를 수식으로 나타내면 다음과 같다.

$f_t$는 업샘플링 함수를 뜻하며 $D_u(x)$는 저해상도 깊이 영상 $D(x)$를 $I(x)$와 같은 해상도를 갖도록 업샘플링한 블러

된 영상을 뜻한다. 제안하는 네트워크는 $D_u(x)$와 $I(x)$를 네트워크의 입력으로 하며 복원된 깊이 영상을 출력으로 한다. 네트워크 내부에서 변위벡터를 예측해 워핑을 진행하며, 출력 깊이 영상을 다시 입력 깊이 영상으로 하여 해당 과정을 반복한다. 이러한 반복과정을 통해 변위벡터를 점진적으로 예측할 수 있으며, 각 반복과정마다 신뢰도 마스크를 사용해 워핑된 깊이 영상에서 실제로 사용할 부분을 결정한다. 변위벡터를 점진적으로 예측하면 각 반복 과정에서 작은 크기의 변위벡터를 예측하며, 워핑 역시 조금씩 진행된다. 따라서 변위벡터를 한 번에 찾는 방법보다 경계 중앙을 더욱 안정적으로 찾을 수 있다. 또한 각 반복 과정에서 신뢰도 마스크를 사용한 혼합기법을 사용하였다. 각 마스크는 워핑된 깊이 영상에서 유효한 부분을 판단한다. 이를 신뢰도라 하며, 신뢰도에 따라 워핑된 깊이 영상과 워핑되지 않은 깊이 영상을 혼합한다. 이러한 과정은 변위벡터를 잘못 예측했을 경우를 방지할 수 있으며, 네트워크의 안정성을 향상시킬 수 있다.

## 3. 네트워크 구조

제안하는 기술은 그림 2과 같이 딥러닝 네트워크로 구성된다. 그림 중앙에 위치하는 네트워크는 변위벡터 네트워크(displacement network)이며, 고해상도 컬러 영상과 블러된 깊이 영상으로부터 변위벡터를 찾는다. 아래에 있는 네트워크는 합성 네트워크(fusion network)이며, 변위벡터 네트워크 에서 전달받은 변위벡터를 사용해 깊이 영상을 재구성한다. 두 네트워크는 서로 결합되어 있기 때문에, end-to-end 방식으로 학습이 가능하다. 각 네트워크의 세부적인 정보는 아래와 같다.

### 3.1 변위벡터 네트워크

변위벡터 네트워크 는 각 화소의 위치에서 변위벡터를 예측한다. 앞서 기술했듯이 블러된 영상은 경계 중심부의 픽셀이 경계 외부로 이동되어 생긴 것으로 가정한다. 따라서 블러되지 않은 영상과 블러된 영상간의 변위벡터를 찾아야 하며, 이 과정은 optical flow를 얻는 과정과 비슷하다. 그러므로 그림 2의 변위벡터 네트워크 는 r광류를 찾는 기술들 중에서 state-of-the-art에 속하는 FlowNetS 구조를 적용하였다.

이전 논문들에서는 워핑 기법을 적용하기 전에 경계 중심의 위치를 찾았다[2,3]. 이러한 접근법은 경계주변에서 경계 중심으로 워핑하는 방향을 결정할 수 있었고, 전체적인 성능에도 많은 영향을 주었다. 하지만 제안 기술은 컬러 영상을 참조하며, 컬러 영상으로부터 경계 중심을 바로 결정할 수 있기 때문에 해당 정차를 포함하지 않는다.

### 3.2 합성 네트워크

합성 네트워크 는 일반적인 GridNet[4] 구조를 사용하였으며, 변위벡터 네트워크 에서 변위벡터와 마스크를 받아 워핑과 혼합 과정을 거친 특징맵을 입력으로 받는다. 또한 입력 받은 특징맵을 재구성하여 깊이 영상을 출력한다. 워핑 과정의 경우 변위벡터를 사용해 깊이 영상의 각 화소들을 이동시키는 것을 의미하며, 해당 과정은 특징 영역에서 이루어진다. 또한 제안하는 네트워크는 반복적인 구조를 사용하여 워핑을 진행한다. 따라서 해당 반복 시점에서 이동된 깊이 영상 중 유효한 부분을 마스크를 통해서 결정한다. 이를 이동되지 않은 깊이 영상과 혼합하며 이 과정을 혼합이라고 한다.

### 3.3 Residual block

제안 기술에서는 워핑기법을 화소영역이 아닌 특징영역에서 하기 위해 각 네트워크의 입력으로 들어가기 전에 몇 개의 잔차 블록들로 이뤄진 층을 지난다. 컬러 영상의 특징을 추출하기 위한 잔차 블록보다 깊이 영상의 특징을 추출하기 위한 잔차 블록을 두껍게 구성하였다.

### 3.4 Feedback system

제안하는 네트워크는 되먹임 구조로 이루어져 있다. 그림 2에서 보는것과 같이 변위벡터 네트워크 에서 마스크를 예측하는데, 이는 현재 이동된 화소들의 신뢰도를 의미한다. 따라서 마스크는 해당 반복 구간에서 이동된 경계의 위치 정보를 가지고 있다. 그러므로 마스크를 이용한 혼합과정으로 만들어진 깊이 영상은 경계의 위치 정보를 잘 가지고 있으며, 다음 반복 과정에서 변위벡터 네트워크 가 경계의 중심을 효과적으로 찾을 수 있게 해준다. 그림 3에서 보는것과 같이 마스크를 시각화 하였을 때 반복 과정을 거치

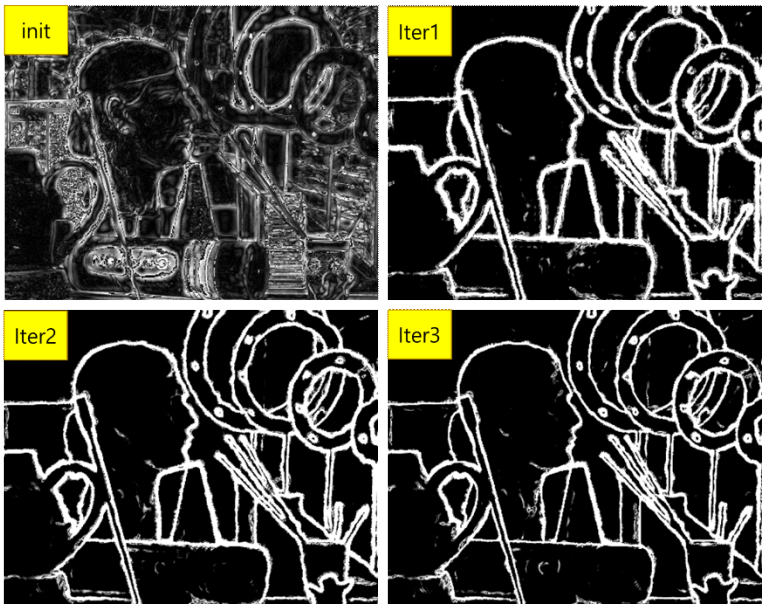



그림 3. 마스크를 시각화 한 영상
Fig. 3. Visualization of mask

면서 경계의 위치 정보가 점점 더 날카로워지는 것을 볼 수 있다.

## Ⅲ. 실험결과 및 분석

### 1. 실험 세부사항

제안하는 방법을 실험하여 검증하는데 있어서 각각 목표하고자 하는 배율을 따로 학습하였다. 또한 $\beta_1$= 0.9, $\beta_2$= 0.999 로 하는 Adam optimizer를 사용하였다[5]. 학습률은 $1e^{-4}$으로 하였으며 매 5 epoch 마다 5로 나누었다. 모든 실험은 pytorch를 사용한 end- to-end 방식을 사용하였다. 또한 반복횟수의 경우 출력 영상의 품질과 네트워크의 복잡도를 고려하여 총 3번 진행하였다.

### 2. 데이터셋

공정한 비교를 위해서 결과 비교에는 네 개의 유명한 데이터셋을 사용하였다. 처음 두개의 데이터셋은 학습집합과 평가집합으로 나누어 사용하였으며 나머지 두개의 데이터셋은 평가에만 사용하였다. 세부적인 사항은 다음과 같다.

1) Sintel dataset[6] : 이 데이터셋은 컴퓨터 그래픽 영상이다. 컬러 영상과 깊이 영상을 모두 제공한다. 영상들의 해상도는 1024 x 438 이다. 컬러 영상과 깊이 영상의 경계는 일치한다. 네트워크 학습에는 총 1000개의 컬러 깊이 영상 쌍이 사용되며 평가에는 총 300개의 컬러 깊이 영상 쌍이 사용하였다.
2) NYU v2 dataset[7] : 이 데이터셋은 Microsoft Kinect로 촬영된 RGB/D 영상으로 구성 되어있다. 각 영상들은 640 x 480의 해상도를 가지며 학습을 위해 800개의 영상을 평가를 위해 600개의 영상을 사용하였다.
3) Lu dataset[8] : 6개의 RGB/D 영상이 제공된다. 각 영상들은 640 x 480의 해상도를 가지며 ASUS Xtion Pro 카메라를 통해 촬영되었다. 이 데이터셋은 평가에만 사용하였다.
4) Middlebury dataset[9]. 2001-2006 데이터셋에서 총 30개의 RGB/D 영상으로 구성되며 평가에만 사용하였다.

표 1. RMSE 지표를 통한 정량적 품질 비교
Table 1. Quantitative performance comparisons in terms of RMSE

| Dataset | Lu | | NYU v2 | | Sintel | | Middlebury | |
|---|---|---|---|---|---|---|---|---|
| Methods | 4x | 8x | 4x | 8x | 4x | 8x | 4x | 8x |
| DJF[10] | 2.39 | 3.94 | 1.73 | 2.60 | 3.62 | 4.88 | 1.82 | 3.32 |
| PAC[11] | 1.39 | 3.42 | 1.52 | 2.75 | 3.72 | 5.42 | 1.61 | 3.54 |
| DKN[12] | 1.14 | 2.42 | 1.08 | 1.91 | 2.92 | 4.55 | 1.32 | 2.35 |
| GWN[13] | 1.04 | 2.24 | 0.91 | 1.60 | 2.66 | 3.81 | 1.27 | 2.07 |
| Ours | 0.91 | 1.67 | 0.80 | 1.37 | 2.46 | 3.12 | 1.23 | 1.84 |

표 2. MAE 지표를 통한 정량적 품질 비교
Table 2. Quantitative performance comparisons in terms of MAE

| Dataset | Lu | | NYU v2 | | Sintel | | Middlebury | |
|---|---|---|---|---|---|---|---|---|
| Methods | 4x | 8x | 4x | 8x | 4x | 8x | 4x | 8x |
| DJF[10] | 0.83 | 1.25 | 0.85 | 1.18 | 1.02 | 1.19 | 0.99 | 1.60 |
| PAC[11] | 0.54 | 1.45 | 0.63 | 1.29 | 0.72 | 1.50 | 0.81 | 1.80 |
| DKN[12] | 0.24 | 0.59 | 0.37 | 0.75 | 0.31 | 0.78 | 0.59 | 0.97 |
| GWN[13] | 0.23 | 0.60 | 0.33 | 0.63 | 0.28 | 0.53 | 0.58 | 0.92 |
| Ours | 0.21 | 0.49 | 0.28 | 0.54 | 0.25 | 0.32 | 0.56 | 0.83 |

## 3. 정량적 비교

제안 기술을 평가하기 위해서 모델 기반 방법을 포함한 여러 우수한 기술들을 비교군으로 사용하였다. 위에서 언급한 4개의 테스트 데이터셋을 사용하여 각각 4배와 8배의 배율에 대한 평가를 진행하였다. 표 1과 2는 각각 root mean square error(RMSE) 지표와 mean absolute difference error(MAE)를 평가 지표로 사용한 결과를 표기하였으며, 가장 좋은 값을 빨간색으로 두번째로 좋은 값을 파란색으로 표기하였다. 표 1과 표 2를 통해 RMSE 관점과 MAE 관점에서 모두 제안 기술이 가장 좋은 성능을 보였음을 알 수 있다.

## 4. 정성적 비교

제안 기술의 시각적 비교를 위해서 그림 4와 같이 비교를 진행하였다. 몇 개의 데이터셋에서 사진을 뽑아 특정한 부분을 확대하여 비교하였으며, 대부분의 경우 제안 기술이 깊이 영상의 구조를 효과적으로 복원하였음을 알 수 있다. 예를 들어 그림 4의 첫번째 줄을 보면, 제안 기술이 가장 날카로운 경계를 가지고 있으며 물체의 전체적인 구조 역시 잘 보존됨을 확인할 수 있다. 또한 물체가 겹쳐진 경우인 다섯 번째 줄을 보면 제안 기술이 물체를 가장 명백하게 분리했음을 알 수 있다. 또한 그림 5는 그림 4보다 작은 해상도를 가진 영상을 사용한 결과이다. 같은 크기의 영역에서는 저해상도 영상에는 고해상도 영상보다 더욱 복잡한 패턴이 존재한다. 따라서 일반적으로 저해상도 영상을 업샘플링 하는 것이 고해상도 영상을 업샘플링 하는 것 보다 어려운 작업이다. 그림 5의 두 번째 줄과 세 번째 줄을 보면 다른 비교 기술들에 비해 제안 기술이 깊이 영상의 경계를 더욱 날카롭게 복원함을 볼 수 있다. 반면, 다섯 번째 줄과 같이 다운샘플링을 하는 과정에서 작은 물체들이 사라지는 경우 다른 기술들과 동일하게 제안 기술 역시 사물의 세부적인 부분들을 복원하지는 못하는 것을 확인할 수 있다. 하지만 제안 기술은 물체가 사라지는 경우가 아니라면 경계를 어느정도 복원할 수 있다.

그림 4의 여덟 번째 줄과 그림 5의 아홉 번째 줄을 보면 블러링 과정에 의해서 물체가 겹쳐지는 경우에도 어느정도

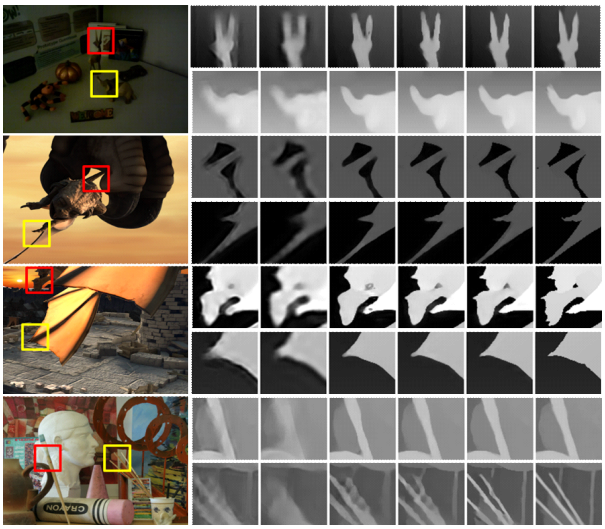


그림 4. 8배 업샘플링한 환경에서 Lu, Temple3, Temple2 그리고 Art 깊이 영상의 시각적인 비교: 왼쪽부터 DJF[10], PAC[11], DKN[12], GWN[13], 제안기술, 원본

Fig. 4. Visual comparisons for the 8x upsampled Lu, Temple 3, Temple 2, and Art depth maps by various methods: DJF[10], PAC[11], DKN[12], GWN[13], Ours, and ground-truth(from left to right)

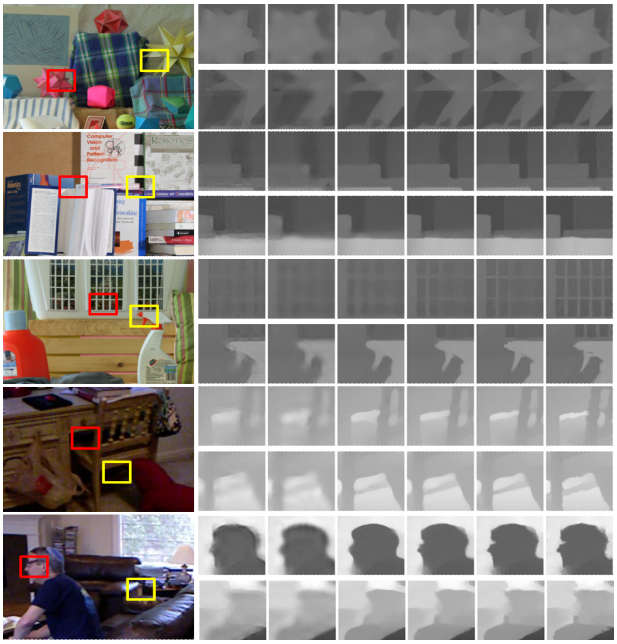


그림 5. 8배 업샘플링한 환경에서 Middlebury와 NYU v2 깊이 영상의 시각적인 비교: 왼쪽부터 DJF[10], PAC[11], DKN[12], GWN[13], 제안기술, 원본

Fig. 5. Visual comparisons for the 8x upsampled Middlebury and NYU v2 depth maps by various methods: DJF[10], PAC[11], DKN[12], GWN[13], Ours, and ground-truth(from left to right)

복원할 수 있음을 볼 수 있다. 기존의 변위벡터를 사용하는 기술인 GWN[13]의 결과와 비교했을 때 역시 제안 기술이 더 우수함을 볼 수 있다. 이를 통해 반복적인 구조를 사용해변위벡터를 점진적으로 예측하는 기술이 변위벡터를 한 번에 예측하는 방법보다 안정적임을 알 수 있다. 그러나 제안 기술은 컬러 영상에 의존한다는 약점이 있다. 예를 들어 제안 기술의 경우 깊이 영상의 경계 부분을 복원할 때 컬러 영상의 구조를 복제하는 경향을 띈다. 이 현상은 실제 영상과 컴퓨터 그래픽 영상을 비교하면 더욱 잘 나타나며 그림 4의 두 번째 줄과 세 번째 줄의 결과를 보면 그 차이를 확인할 수 있다. 두 번째 줄의 결과 영상은 세 번째 줄의 결과 영상에 비해서 매우 흐린 것을 볼 수 있으며 이는 세 번째 줄의 컬러 영상이 두 번째 줄의 컬러 영상에 비해서 더욱 날카로운 경계를 가지고 있기 때문이다.

## Ⅳ. 결 론

본 논문에서는 영상 워핑 기법을 통해 깊이 영상을 업샘플링하는 새로운 방법을 제안한다. 고해상도 컬러 영상과 깊이 영상의 정보들을 통해 변위벡터를 예측하여 영상을 재구성하며, 변위 벡터를 점진적으로 찾음으로써 워핑 과정에서 생길 수 있는 오차를 줄였다. 또한 신뢰도 기반 마스크 기법을 통해 경계의 위치 정보를 효율적으로 전달하며 이는 반복 구조와 더불어 제안 기술의 안정성을 높였다. RMSE 지표를 통한 비교표와 MAE 지표를 통한 비교표를 통해 제안 기술이 존재하는 다른 기술들에 비해서 우수한 성능을 보임을 알 수 있으며 시각적 결과 역시 제안 기술의 우수성을 보여준다. 하지만 제안하는 기술의 성능은 컬러 영상과 깊이 영상의 유사성에 의존한다는 한계가 존재한다. 따라서 향후 연구에서 이러한 한계가 다뤄질 예정이다.

## 참 고 문 헌 (References)

[1] N. Arad and C. Gotsman, ‘‘Enhancement by image-dependent warping,’’ IEEE Trans. Image Process., vol. 8, no. 8, pp. 1063 – 1074, Aug. 1999.

[2] A. Krylov, A. Nasonov, and Y. Pchelintsev, ‘‘Single parameter post-processing method for image deblurring,’’ in Proc. 7th Int. Conf. Image

Process. Theory, Tools Appl., pp. 1 - 6 Nov. 2017.
[3] A. Nasonova and A. Krylov, ‘‘Deblurred images post-processing by Poisson warping,’’ IEEE Signal Process. Lett., vol. 22, no. 4, pp. 417 - 420, Apr. 2015.
[4] D. Fourure, R. Emonet, E Fromont, D. Muselet, A. Tremeau, C. Wolf, "Residual conv-deconv grid network for semantic segmentation." arXiv preprint arXiv:1707.07958, 2017.
[5] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980, 2014.
[6] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, ‘‘A naturalistic open source movie for optical flow evaluation,’’ in Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer, pp. 611 - 625, 2012.
[7] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, ‘‘Indoor segmentation and support inference from RGBD images,’’ in Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer, pp. 746 - 760, 2012.
[8] H. Hirschmuller and D. Scharstein, ‘‘Evaluation of cost functions for stereo matching,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1 - 8, Jun. 2007.
[9] S. Lu, X. Ren, and F. Liu, ‘‘Depth enhancement via low-rank matrix completion,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 3390 - 3397, Jun. 2014.
[10] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang, ‘‘Joint image filtering with deep convolutional networks,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 41, no. 8, pp. 1909 - 1923, Aug. 2019.
[11] H. Su, V. Jampani, D. Sun, O. Gallo, E. Learned-Miller, and J. Kautz, ‘‘Pixel-adaptive convolutional neural networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., pp. 11166 - 11175, Jun. 2019.
[12] K. Beomjun, J. Ponce, H. Bumsub, ‘‘Deformable kernel networks for guided depth map upsampling,” arXiv preprint, arXiv:1910.08373, 2019.
[13] Y. Yoonmo, K Dongsin, O. Byung Tae, "Deep Convolutional Grid Warping Network for Joint Depth Map Upsampling." IEEE Access, vol. 8, pp. 147580-147590, Aug 2020

## 저 자 소 개

### 김 동 신

- 2020년 2월 : 한국항공대학교 전자및항공전자공학 학사
- 2020년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사과정
- ORCID : https://orcid.org/0000-0002-0301-3275
- 주관심분야 : 영상처리, 비디오압축, 강화학습

### 양 윤 모

- 2015년 2월 : 한국항공대학교 전자및항공전자공학 학사
- 2015년 3월 ~ 2017년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2017년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : http://orcid.org/0000-0003-2816-1685
- 주관심분야 : 영상처리, 영상포렌식, 기계학습

### 오 병 태

- 2003년 8월 : 연세대학교 전기전자공학부 학사
- 2009년 8월 : Univ. of Southern California (USC), Dept. of Electrical Eng. 석사 및 박사
- 2009년 ~ 2013년 : 삼성종합기술원 전문연구원
- 2013년 ~ 현재 : 한국항공대학교 항공전자정보공학부 부교수
- ORCID : http://orcid.org/0000-0003-1437-2422
- 주관심분야 : 영상처리, 비디오압축, 영상포렌식, 3차원영상