

논문 2019-14-01

# 딥러닝 기반 3차원 라이다의 반사율 세기 신호를 이용한 흑백 영상 생성 기법

(Deep Learning Based Gray Image Generation  
from 3D LiDAR Reflection Intensity)

김 현 구, 유 국 열, 박 주 현, 정 호 열\*

(Hyun-Koo Kim, Kook-Yeol Yoo, Ju H. Park, Ho-Youl Jung)

**Abstract :** In this paper, we propose a method of generating a 2D gray image from LiDAR 3D reflection intensity. The proposed method uses the Fully Convolutional Network (FCN) to generate the gray image from 2D reflection intensity which is projected from LiDAR 3D intensity. Both encoder and decoder of FCN are configured with several convolution blocks in the symmetric fashion. Each convolution block consists of a convolution layer with  $3 \times 3$  filter, batch normalization layer and activation function. The performance of the proposed method architecture is empirically evaluated by varying depths of convolution blocks. The well-known KITTI data set for various scenarios is used for training and performance evaluation. The simulation results show that the proposed method produces the improvements of 8.56 dB in peak signal-to-noise ratio and 0.33 in structural similarity index measure compared with conventional interpolation methods such as inverse distance weighted and nearest neighbor. The proposed method can be possibly used as an assistance tool in the night-time driving system for autonomous vehicles.

**Keywords :** Artificial intelligence, Deep learning, Fully convolution network, Image generation, LiDAR sensor

## 1. 서 론

라이다 (LiDAR) 센서는 물체까지의 거리를 측정할 수 있고, 물체의 매질에 따른 반사율 (reflection intensity) 정보도 동시에 얻을 수 있다. 라이다 센서 데이터는 햇빛이나 그림자와 같은 주변의 밝기에 영향을 받지 않기 때문에 주간 및 야간에 상관

없이 일정한 데이터를 얻을 수 있는 장점을 가지고 있다 [1-6]. 그러나 라이다 센서로 획득한 3차원 점-구름 (point-cloud) 데이터는 라이다의 스캔 각도에 따라 데이터가 희소한 (sparse) 특성이 있기 때문에 2차원 영상으로 변환 시 조밀하지 않아 정확하게 카메라로 획득한 영상과 정확히 매칭 시키기 힘들다. 라이다 센서의 3차원 점-구름 데이터로부터 2차원 영상을 생성하는 것은 주간 및 야간에 상관없이 일관된 영상을 제공해주기 때문에, 무인 자동차 응용 분야에서 매우 유용하다 [5, 6].

최근 딥러닝을 이용하여 제한된 정보를 가진 영상으로부터 고해상도의 정밀한 영상을 복원하거나 또는 다른 특성을 가진 영상을 생성시키는 연구가 진행되고 있다 [7-14]. 특히, 입력 데이터가 극단적으로 적은 스케치 영상으로부터 텍스처 정보를 채워 넣는 방법 [11, 12]과 열화상 카메라 센서 데이터로부터 컬러 영상을 생성하는 방법 [13, 14]도 연구되고 있다. 이 연구들은 공통적으로 영상 생성 모델로써 인코더-디코더 구조를 가진 FCN (Fully

\*Corresponding Author (hoyoul@yu.ac.kr)

Received: Nov. 8 2018, Revised: Nov. 30 2018,  
Accepted: Dec. 3 2018.

H-K. KIM, K-Y. Yoo, J. H. Park, H-Y. Jung:  
Yeungnam University

※ 이 성과는 2018년도 정부 (과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. NRF-2018R1A2B6005105).

※ 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2018-2016-0-00313).

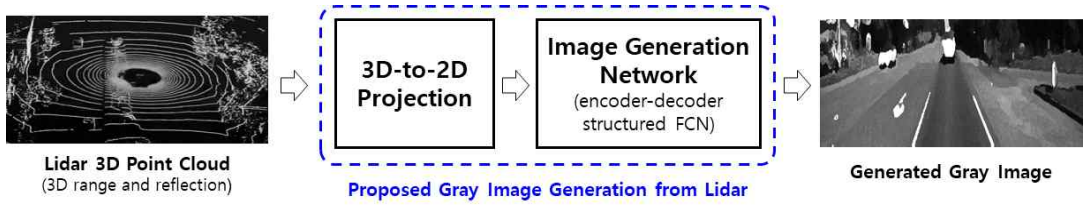


그림 1. 3차원 라이다의 반사율 세기 정보 기반 흑백 영상 생성을 위한 제안 방법의 구성

Fig. 1 Proposed gray image generation method from 3D LiDAR reflection intensity

Convolutional Network 모델) [7, 8]을 기본적으로 사용한다.

본 연구에서는 딥러닝을 이용하여 3차원 라이다의 반사율 세기 정보로부터 흑백 영상 생성하는 효율적인 기법을 제안한다. 제안 기법은 라이다의 3차원 점-그룹 데이터의 반사율 세기 신호를 2차원 반사율 영상으로 변환하고, 영상 생성 네트워크를 사용하여 2차원 흑백 영상으로 생성한다. 이를 위해 영상 생성 네트워크는 관련 연구들 [7-14] 중에서 낮은 복잡도로 효과적으로 영상을 발생시키는 것으로 알려진 인코더-디코더 구조를 가진 FCN 모델을 적용한다. 특히, 입력 데이터인 2차원 반사율 영상은 생성하고자 하는 2차원 흑백 영상에 비해 픽셀 정보가 희소하므로, 인코더-디코더 구조를 가진 FCN 모델에서 인코더와 디코더 네트워크의 깊이를 조절하는 네트워크 모델을 제안하고 실험을 통해 성능을 분석한다.

제안된 방법은 라이다 데이터로부터 흑백 영상을 생성하는 방법으로 카메라로 촬영한 영상 데이터와 유사한 영상을 생성하는 방법이다. 따라서 제안 방법은 보간법 (interpolation)을 이용하여 라이다 센서의 반사율 데이터의 해상도를 개선하는 기존의 방법들 [5, 6]과는 다른 새로운 차원의 영상 생성 기법이라 할 수 있다. 더욱이, 참고문헌 [5]와 [6]에서 최고 성능을 나타낸다고 하는 역거리 가중 (inverse distance weighted) 보간법과 최근접 이웃 (nearest neighbor) 보간법을 사용하더라도, 생성된 2차원 반사율 영상은 카메라로 획득한 영상에 비해 많은 잡음과 낮은 화질로 인해 기존의 카메라 영상기반 물체인식 알고리즘에는 적용하기 어렵다. 본 연구에서는 제안된 방법의 성능평가를 위하여 이러한 기존의 보간법 방법들을 구현하여 그 성능을 비교한다.

이 논문의 나머지 부분은 다음과 같이 구성된다. II장에서는 제안하는 3차원 라이다의 반사율 세기 정보 기반 흑백 영상 생성을 위한 방법을 소개한다.

III장에서는 실험을 위한 데이터 및 실험 환경을 설명하고 인코더-디코더 구조를 가진 FCN 모델에서 인코더와 디코더 네트워크의 깊이에 따른 실험 결과를 분석한다. 또한 보간 기법을 사용하는 기존의 방법들과 성능 비교를 한다. IV장에서는 결론을 도출한다.

## II. 제안하는 방법

제안하는 3차원 라이다의 반사율 세기 정보 기반 흑백 영상 생성을 위한 방법은 그림 1과 같이 3D-to-2D 투영 방법과 영상 생성 네트워크의 2가지 단계로 구성된다. 3D-to-2D 투영 방법에서는 반사율 세기 정보를 가진 3차원 라이다의 점-그룹 데이터를 희소한 픽셀 정보를 가진 2차원 반사율 영상으로 변환 한다. 그리고 인코더-디코더 구조를 가진 FCN 모델을 사용하여 2차원 반사율 영상으로부터 카메라로 획득한 영상과 유사한 화질을 갖는 흑백 영상을 생성한다.

### 1. 제안된 라이다-to-흑백 영상 생성 방법

2차원 반사율 영상은 식 (1)을 통해 반사율 정보가 있는 라이다 3차원 점-그룹 좌표  $[X, Y, Z]$ 에 라이다-카메라 투영 행렬인  $P_L^C$ 를 곱하여 매핑 할 카메라의 2차원 영상 좌표  $[u, v]$ 를 계산 한다 [15].

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P_L^C \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

$$\text{where } P_L^C = \begin{bmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1000 \\ 0100 \\ 0010 \end{bmatrix} \begin{bmatrix} R_L^C & t_L^C \\ 0 & 1 \end{bmatrix}$$

여기에서  $s$ 는 카메라의 스케일 인자이며  $f_u$ 와  $f_v$ 는 카메라의  $u$ 축과  $v$ 축의 초점 거리 (focal length),  $c_u$ 와  $c_v$ 는 카메라의  $u$ 축과  $v$ 축의 주-점

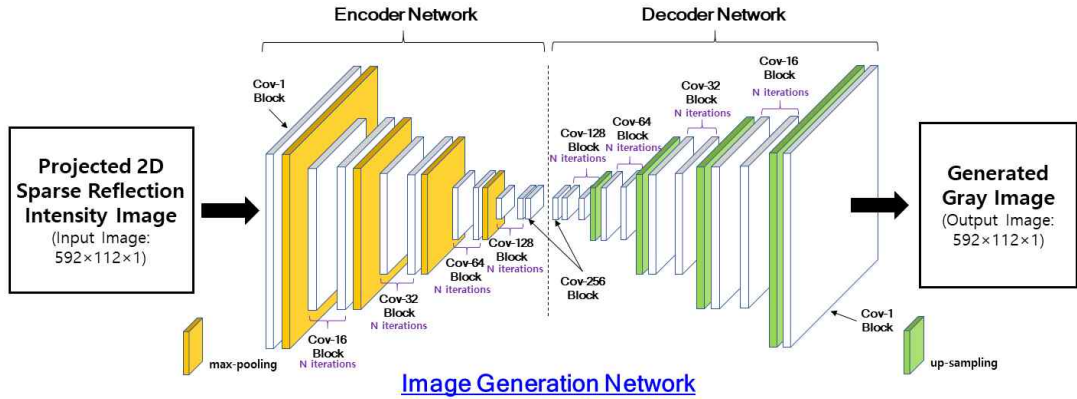


그림 2. 제안하는 흑백 영상 생성 네트워크 구조

Fig. 2 Architecture of proposed gray image generation network

(principal-point)을 나타낸다.  $3 \times 3$  행렬 인  $L^C$ 과  $1 \times 3$  행렬 인  $t_L^C$ 는 각각 라이더의 3차원 세계 좌표계를 카메라의 2차원 좌표계로 변환하는 회전변환 행렬과 평행이동 행렬을 나타낸다.

계산된 영상좌표  $[u, v]$ 에 해당되는 라이더 3차원 점-구름 좌표  $[X, Y, Z]$ 의 반사율 값을 매핑 하여 2차원 반사율 영상을 생성한다. 생성된 2차원 영상은 모든 좌표 값에 화소 값을 가지지 않는 최소한 형태를 가지게 된다.

그림 2에서 제안하는 인코더-디코더 구조를 가진 FCN 모델을 사용하는 영상 생성 네트워크 모델을 나타낸 것이다. FCN 모델에서 인코더 네트워크는 특징들을 추출하는 역할을 수행한다. 디코더 네트워크는 낮은 해상도를 가진 다양한 특징 맵들을 점진적으로 보간 연산을 수행함으로써 최종적으로 타겟 영상과 같은 해상도를 가지는 흑백 영상을 생성한다.

제안하는 네트워크 모델은 인코더와 디코더 네트워크의 깊이가 동일한 대칭적인 구조를 가지고 있다. 컨볼루션 블록은 입력 및 출력의 경우 단일 계층으로 구성되고, 그 외의 계층들은 컨볼루션 계층을  $N$  회씩 반복시켰다. 그림 2에서 Conv- $K$  block은  $K$  개의  $3 \times 3$  크기의 필터를 가진 컨볼루션 계층 (convolution layer), 배치 정규화 계층 (batch normalization layer), 그리고 활성화 함수 (activation function) 순으로 구성된다. 또한 인코더와 디코더 네트워크는 인자 2를 적용한 max-pooling과 up-sampling을 사용하였다. 모든 컨볼루션 계층은 stride 1 과 반복 padding을 사용

하였다. 모든 컨볼루션 블록의 활성화 함수는 ReLU 을 사용하였고, 디코더 네트워크의 마지막 컨볼루션 블록인 Conv-1 block 에서는 tanh 활성화 함수를 사용하였다. 활성화 함수 ReLU와 tanh는 다음과 같이 정의된다.

$$ReLU = \max(0, x) \quad (2)$$

$$\tanh(x) = 2 \left( \frac{1}{1 + e^{-2x}} \right) - 1 \quad (3)$$

제안하는 네트워크의 총 파라미터 개수 ( $P_t$ )는 식 (4)에서 보인 바와 같이 컨볼루션 계층의 가중치 파라미터 개수 ( $w_t$ ), 바이어스 파라미터 개수 ( $b_t$ ) 및 배치 정규화 계층의 파라미터 개수 ( $BN_t$ )들의 합이다.

$$P_t = w_t + b_t + BN_t \quad (4)$$

여기에서  $b_t$ ,  $w_t$  및  $BN_t$ 는 다음과 같이 계산된다.

$$w_t = 3^2(K_1 + K_6^2) + 2 \times 3^2 \sum_{i=1}^6 (K_i)^2 (N-1) + 2 \times 3^2 \sum_{i=1}^5 K_i K_{i+1} \quad (5)$$

$$b_t = 2 \times N \sum_{i=1}^6 K_i \quad (6)$$

$$BN_t = 4 \left( 2N \sum_{i=2}^6 K_i + K_1 \right) \quad (7)$$

여기에서  $K_i$ 는  $i$ 번째 컨볼루션 블록에서 컨볼루

선 계층의  $3 \times 3$  크기의 필터 개수를 의미하며  $K_1 = 1$ ,  $K_i = 2^{i+2}$  ( $2 \leq i \leq 6$ ) 이다. 그리고 파라미터를 가진 총 계층 개수 ( $L_t$ )는 컨볼루션 계층과 배치 정규화 계층의 합으로 구해지며 다음과 같이 정의된다.

$$L_t = 7 + 16N \quad (8)$$

## 2. 제안된 네트워크의 학습 및 추론 과정

제안된 네트워크의 학습 과정에서, 모델의 입력으로는 3차원 라이다 점-구름 데이터에서 얻은 희소한 픽셀 정보를 가진 2차원 반사율 영상을 사용하고 학습에 필요한 타겟 영상으로는 반사율 영상에 대응하는 카메라로 취득된 흑백 영상을 이용하여 학습한다. 이때, 카메라로 취득된 흑백 영상에서 그림자가 많은 영상은 제외하고 학습을 수행한다. 라이다 데이터는 그림자와 상관없이 없으므로, 이런 영상으로 학습을 시킬 경우, 그림자가 없는 영상을 발생시킬 수 있는 장점을 가진다. 디코더 네트워크의 마지막 컨볼루션 블록의 활성화 함수는 tanh를 사용하므로 생성되는 출력 영상의 범위는  $(-1, 1)$  이다. 따라서 타겟 영상의 데이터 범위도  $(-1, 1)$ 로 변환하여 학습에 사용된다. 학습 시 손실 함수는 MSE (Mean Squared Error)를 사용하며 다음과 같이 정의 된다.

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (T_{(i,j)} - G_{(i,j)})^2 \quad (9)$$

여기에서  $T$  는 타겟 영상을,  $G$  는 생성 영상을 나타내며  $m$  과  $n$  은 각각 영상의 가로 및 세로 크기를 의미한다.

제안된 네트워크의 추론 과정에서는 3차원 라이다 점-구름 데이터로부터 라이다-카메라 투영 행렬  $P_L^C$  과 학습된 파라미터를 영상 생성 네트워크에 적용하여 흑백 영상을 생성한다. 이때 네트워크의 출력 데이터의 범위는  $(-1, 1)$  이므로 영상으로 변환하기 위해 출력 데이터 범위를  $(-1, 1)$  에서  $(0, 255)$  로 변경하여 최종적인 흑백 영상을 생성한다.

## III. 실험 환경 및 결과 분석

이 절에서는 실험에 사용한 평가 데이터 구성, 학습을 위한 하이퍼-파라미터 (hyper-parameter), 평가 지표에 대해서 설명한다. 그리고, 제안된 영상

생성 네트워크의 컨볼루션 블록 개수에 따른 성능 결과를 비교·분석한다. 또한 희소한 (sparse) 라이다의 반사율 영상을 보간 하여 2차원 흑백 영상을 생성하는 기존의 방법들과 성능을 비교하여 제안 방법의 성능을 평가한다.

### 1. 평가 데이터 구성

KITTI 데이터는 주간 환경에서 다양한 지역 (일반도로, 도시 및 주거지역 등)에서 라이다 및 카메라 센서 등 다중 센서를 사용하여 취득된 주행 환경 인식을 위한 데이터이다 [15]. 제안된 방법은 라이다 데이터로부터 그림자가 많은 구간과 야간의 경우에도 그림자가 없는 동일한 품질의 영상을 생성할 수 있도록 KITTI 데이터 중 그림자가 거의 없는 도로 영상만 선택하여 학습에 사용한다.

실험을 위해 KITTI 데이터를 기반으로 동일 시간에 취득된 희소한 정보를 가진 2차원 반사율 영상과 흑백 영상을 한 쌍으로 구성하여 사용한다. 앞에서 언급한 것과 같이 2차원 반사율 영상은 영상 생성 네트워크의 입력 영상으로, 흑백 영상은 타겟 영상으로 사용된다. 실험에 사용된 평가 데이터는 총 4,308개 반사율-흑백영상 쌍으로 구성되며 학습을 위해 2,872개, 유효성 검증을 위해 718개, 그리고 테스트를 위해 718개의 반사율-흑백영상 쌍으로 분리하여 사용한다.

### 2. 학습을 위한 하이퍼-파라미터

본 연구에서 제안된 영상 생성 네트워크는 배치 (batch) 크기를 4로 설정하여 최대 2,000세대 (epochs)까지 반복 학습되었다. 학습 최적화를 위해 Adam (adaptive moment estimation) 최적화 기법 [16]을 사용하고 학습 계수  $l_r = 5 \times 10^{-4}$  와 모멘텀 파라미터  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-8}$  이 적용되었다.

### 3. 평가 지표 (Measurement Metric)

영상 생성 네트워크의 타겟 영상인 카메라로 취득된 흑백 영상과 생성된 흑백 영상 간의 화질 평가를 위해 PSNR (peak signal to noise ratio)과 SSIM (structural similarity index measure)을 사용한다 [17]. PSNR은 다음과 같이 정의된다.

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad (10)$$

표 1. 콘볼루션 블록의 반복 회수( $N$ )에 따른 층 계층의 개수 및 총 파라미터 개수

Table 1. The number of layers and parameters according to the iteration number of convolution block ( $N$ )

| The iteration number of convolution block ( $N$ ) | The number of layers ( $L_l$ ) | The number of weight parameters ( $P_l$ ) |
|---------------------------------------------------|--------------------------------|-------------------------------------------|
| 1                                                 | 23                             | 1,378,447                                 |
| 2                                                 | 39                             | 1,772,527                                 |
| 3                                                 | 55                             | 2,166,607                                 |
| 4                                                 | 71                             | 2,560,687                                 |
| 5                                                 | 87                             | 2,954,767                                 |
| 6                                                 | 103                            | 3,348,847                                 |
| 7                                                 | 119                            | 3,742,927                                 |
| 8                                                 | 135                            | 4,137,007                                 |
| 9                                                 | 151                            | 4,531,087                                 |
| 10                                                | 167                            | 4,925,167                                 |

SSIM은 인간의 시각 체계가 영상의 구조적 정보에 예민하다는 점을 기반으로 구조적 정보를 이용하여 영상의 유사도 평가하는 방법이다. SSIM은 두 영상 간의 픽셀 별 차이를 계산하는 PSNR과 달리, 인간의 시각에서 주요 내용으로 인식하는 영상의 휘도 (luminance), 명암비 (contrast), 구조 (structure)적인 차이의 조합으로 모델링하여 다음과 같이 정의된다.

$$SSIM = \frac{(2\mu_T\mu_G + C_1)(2\sigma_{TG} + C_2)}{(\mu_T^2 + \mu_G^2 + C_1)(\sigma_T^2 + \sigma_G^2 + C_2)} \quad (11)$$

여기에서,  $T$  는 타겟 영상을,  $G$  는 생성 영상을 나타내며  $\mu$  와  $\sigma^2$  는 각각 영상에서 평균과 분산을,  $\sigma_{TG}$  는 타겟 영상과 생성 영상 간의 공분산을 나타낸다. 양수 상수  $C_1$  및  $C_2$  는 0으로 나누어지는 것을 방지하기 위한 변수이다.

#### 4. 영상 생성 네트워크의 깊이에 따른 성능 비교

제안된 FCN 기반 흑백 영상 생성 네트워크의 깊이에 따른 성능분석을 위해 각 콘볼루션 블록의 반복 회수  $N$ 을 증가시키면서 생성된 흑백 영상의 화질을 평가한다. 표 1은 콘볼루션 블록의 반복 회수  $N$ 에 따른 층 계층의 개수 및 총 파라미터 개수를 나타낸 것이다.

표 2는 영상 생성 네트워크 모델에서 네트워크의 깊이, 즉 콘볼루션 블록의 반복 회수 변화에 따른 성능 결과를 나타낸다. 결과에 따르면, 층 계층

표 2. 영상 생성 네트워크의 깊이 (콘볼루션 블록의 반복 회수) 에 따른 성능 평가

Table 2. Performance evaluation according to the depth of image generation network (the iteration number of convolution block)

| The iteration number of convolution block ( $N$ ) | Measurement metric |              |
|---------------------------------------------------|--------------------|--------------|
|                                                   | Average PSNR [dB]  | Average SSIM |
| <u>1</u>                                          | <u>17.69</u>       | <u>0.43</u>  |
| 2                                                 | 18.80              | 0.48         |
| 3                                                 | 19.26              | 0.50         |
| 4                                                 | 19.52              | 0.51         |
| 5                                                 | 19.62              | 0.52         |
| <u>6</u>                                          | <u>19.70</u>       | <u>0.53</u>  |
| 7                                                 | 19.62              | 0.52         |
| 8                                                 | 19.52              | 0.52         |
| 9                                                 | 19.12              | 0.49         |
| 10                                                | 18.96              | 0.48         |

의 개수가 103개 (콘볼루션 블록의 반복 개수  $N=6$ ) 일 때 PSNR 과 SSIM 에서 최대 성능을 가지며 총 개수가 103개를 넘으면 너무 깊어져서 기울기 값이 사라지는 그라디언트 소실 문제 [18]가 발생되어 제대로 학습이 되지 않는 것으로 보인다. 특히, 기존의 영상 생성 연구들에서 사용된 콘볼루션 블록의 반복 사용 없는 대칭 구조의 네트워크 ( $N=1$ )에 비해 PSNR에서 평균 2.01dB, SSIM에서 평균 0.1 높은 성능을 보인다.

#### 5. 기존의 보간 방법 연구와 성능 비교

이 섹션에서는 3차원 라이더의 점-구름 데이터로부터 희소한 2차원 반사율 영상으로 변환 후, 보간 방법을 사용하여 2차원 흑백 영상을 생성하는 기존의 방법들과의 성능 비교를 수행한다. 실험에서 사용된 기존의 보간 방법은 1장에서 언급한 역거리 가중 보간법과 최근접 이웃 보간법이다.

그림 3과 표 3은 제안하는 영상 생성 네트워크 ( $N=6$ )를 적용한 2차원 흑백 영상 생성 방법과 기존의 보간 방법들과의 성능 결과를 나타낸 것이다. 그림 3에서, (a) 행은 희소한 픽셀 정보를 가진 반사율 영상을, (b) 행은 카메라로 취득한 흑백 영상을, (c) 행은 역거리 가중 보간법을 사용하여 생성된 흑백 영상을, (d) 행은 최근접 이웃 보간법을 사용하여 생성된 흑백 영상을, (e) 행과 (f)행은 제안된 흑백 영상 생성 네트워크를 사용하여 생성된 영상으로 콘볼루션 블록의 반복 회수  $N$ 이 1인 경우와 6인 경우를 각각 나타낸 것이다. 표 3에서 나타

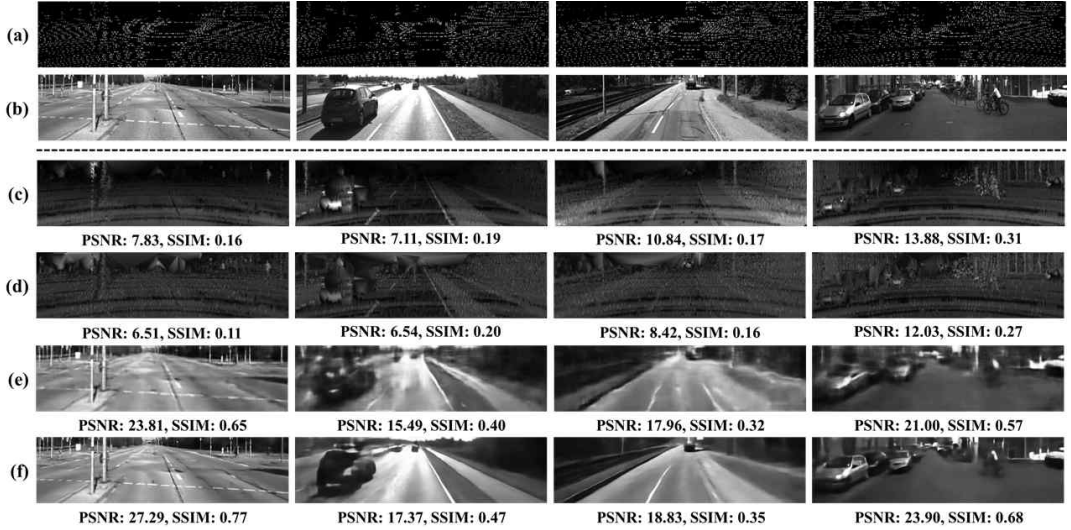


그림 3. 테스트를 위한 평가 데이터에서 영상 생성에 따른 성능 결과, (a) 반사율 영상, (b) 카메라 흑백 영상, (c) 역거리 가중 보간법으로 생성된 흑백 영상, (d) 최근접 이웃 보간법으로 생성된 흑백 영상, (e)  $N=1$  인 경우, 영상 생성 네트워크로 생성된 흑백 영상, (f)  $N=6$  인 경우, 영상 생성 네트워크로 생성된 흑백 영상

Fig. 3 Performance results of the image generation methods in the test dataset, (a) sparse lidar projected image, (b) camera based gray image, (c) inverse distance weighted interpolation based generated image, (d) nearest neighbor interpolation based generated image, (e) image generation network ( $N=1$ ) based generated image, (f) image generation network ( $N=6$ ) based generated image

표 3. 테스트를 위한 평가 데이터에서 기존 보간 방법과 제안하는 방법의 성능 결과 비교

Table 3. Performance results of the proposed method and conventional interpolation methods in the test dataset

| Image generation method                  | Measurement metric |              |
|------------------------------------------|--------------------|--------------|
|                                          | Average PSNR [dB]  | Average SSIM |
| Inverse distance weighted interpolation  | 11.14              | 0.20         |
| Nearest neighbor interpolation           | 9.34               | 0.19         |
| Proposed network based generation method | <b>19.70</b>       | <b>0.53</b>  |

난 것과 같이, 제안된 영상 생성 네트워크를 사용한 2차원 흑백 영상 생성 방법이 기존의 보간 기법을 사용하는 방법들 보다 평균 PSNR 기준으로 8.56dB 이상, 평균 SSIM 기준으로 0.33 이상 향상된 성능을 보인다.

#### IV. 결 론

본 연구에서 제안된 흑백 영상 생성 네트워크가 라이다 데이터로부터 카메라로 촬영된 영상과 유사한 영상을 생성할 수 있음을 보였다. 3차원 라이다의 반사율 세기 정보 기반 흑백 영상 생성을 위해 대칭 구조의 인코더-디코더 구조를 가진 FCN 모델을 적용하는 방법을 제안하였다. 특히, 라이다의 3차원 점-구름 데이터의 반사율 세기 정보를 투영하여 2차원 반사율 영상으로 변환하면 조밀하지 않는 희소한 픽셀 정보를 가진 영상으로 변환되므로 컨볼루션 블록을 반복 적용함으로써 효율적으로 영상을 생성시키는 네트워크를 제안하고 실험을 통해 증명하였다.

라이더 센서는 햇빛과 그림자에 영향을 받지 않기 때문에, 제안된 방법을 사용하면 야간의 경우에도 동일하게 주간 영상을 생성할 수 있을 것으로 기대되며 가까운 미래에 카메라 센서를 대체할 수 있을 것으로 기대된다.

## References

- [1] Y. Park, M. Lee, S. Hur, "Obstacle Classification Method Based on Single 2D LIDAR Database," IEMEJ. Embed. Sys. Appl., Vol. 10, No. 3, pp.179-188, 2015 (in Korean).
- [2] S. K. Kwon, E. Hyun, J. Lee, J. Lee, S. H. Son, "A Novel Human Detection Scheme Using a Human Characteristics Function in a Low Resolution 2D LIDAR," Journal of IEMEJ. Embed. Sys. Appl., Vol. 11, No. 5, pp.267-276, 2016 (in Korean).
- [3] D. S. Hall, "Color lidar scanner," U.S. Patent 8,675,181, Mar. 18, 2014.
- [4] T. D. Barfoot, C. McManus, S. Anderson, H. Dong, E. Beerepoot, C. H. Tong, P. Furgale, J. D. Gammell, J. Enright, "Into Darkness: Visual Navigation Based on a Lidar-intensity-image Pipeline," Proceedings of International Symposium Robotics Research, Vol. 114, pp. 487-504, 2016.
- [5] I. Ashraf, S. Hur, Y. Park, "An Investigation of Interpolation Techniques to Generate 2d Intensity Image from Lidar Data," Journal of IEEE Access, Vol. 5, pp. 8250-8260, 2017.
- [6] A. Asvadi, L. Garrote, C. Premebida, P. Peixoto, U. J. Nunes, "Realtime Deep Convnet-based Vehicle Detection Using 3d-lidar Reflection Intensity Data," Proceedings of Iberian Robotics Conference, Vol. 694, pp. 475-486, 2017.
- [7] V. Badrinarayanan, A. Kendall, R. Cipolla, "Segnet: A Deep Convolutional Encoder-decoder Architecture for Image Segmentation," Journal of IEEE Transaction Pattern Anal. Mach. Intel., Vol. 39, No. 12, pp. 2481-2495, 2017.
- [8] R. Yasrab, N. Gu, X. Zhang, "An Encoder-decoder Based Convolution Neural Network (cnn) for Future Advanced Driver Assistance System (adas)," Journal of Applied Sciences, Vol. 7, No. 4, p. 312, 2017.
- [9] D. Varga, T. Szirnyi, "Fully Automatic Image Colorization Based on Convolutional Neural Network," Proceedings of IEEE International Conference Pattern Recognit., pp. 3691-3696, 2016.
- [10] F. Baldassarre, D. G. Mor'in, L. Rod'es-Guirao, "Deep Koalarization: Image Colorization Using Cnns and Inception-resnet-v2," CoRR, Vol. abs/1712.03400, pp. 1-12, 2017.
- [11] P. Isola, J. Zhu, T. Zhou, A. A. Efros, "Image-to-image Translation with Conditional Adversarial Networks," CoRR, Vol. abs/1611.07004, pp. 1-17, 2016.
- [12] W. Chen, J. Hays, "Sketchygan: Towards Diverse and Realistic Sketch to Image Synthesis," Proceedings of IEEE Conference Comput. Vis. Pattern Recognit., pp. 9416-9425, 2018.
- [13] P. L. Surez, A. D. Sappa, B. X. Vintimilla, "Infrared Image Colorization Based on a Triplet Dcgan Architecture," Proceedings of IEEE Conference Comput. Vis. Pattern Recognit., pp. 212-217, 2017.
- [14] U. Qaynm, Q. Ahsan, Z. Mahmood, M. A. Chcmdary, "Thermal Colorization Using Deep Neural Network," Proceedings of IEEE Interenational Conference Applied Science Tech., pp. 325-329, 2018.
- [15] A. Geiger, P. Lenz, C. Stiller, R. Urtasun, "Vision Meets Robotics: The Kitti Dataset," Journal of International J. Robotics Research, Vol. 32, No. 11, pp. 1231-1237, 2013.
- [16] D. P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," CoRR, Vol. abs/1412.6980, pp. 1-15, 2014.
- [17] A. Hore, D. Ziou, "Image Quality Metrics: Psnr vs. ssim," Proceedings of IEEE Conference Pattern Recognit., pp. 2366-2369, 2010.
- [18] S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," Journal of International J. Uncert., Fuzzi. and Knowledge-Based Syst., Vol. 6, No. 2, pp. 107-116, 1998.

**Hyun-Koo Kim (김 현 구)**

He received the B.S. and M.S. degrees from the department of information and communication engineering, Yeungnam University, Korea, in 2009 and 2011, respectively. He is currently pursuing the Ph.D. degree in the department of information and communication engineering with Yeungnam University, Korea. From 2012 to 2013, he researched in MINES ParisTech (Ecole des mines de Paris), France. From 2013 to 2017, he was a senior researcher in Gyeongbuk Institute of IT Convergence Industry Technology, Korea. His current research interests include deep-learning based signal processing, sensor fusion, object detection, and image generation for autonomous vehicle system.

Email: kim-hk@ynu.ac.kr

**Kook-Yeol Yoo (유 국 열)**

He received the B.S. degree in Electronics Engineering from Kyungpook National University in 1991, and the M.S. and Ph.D. degrees in EE from KAIST (Korea Advanced Institute of Science and Technology), in 1993 and 1998, respectively. From 1998 to 2001, he was a senior engineer of the Samsung electronics, and was engaged in the R&D on visual communication systems. He is with the School of EECS at Yeungnam University since 2001. Now he is a professor. He is a director of KISTI and KICS societies. His research interests are video compression, computer vision, machine learning, and deep learning.

Email: kyoo@ynu.ac.kr

**Ju H. Park (박 주 현)**

He received the Ph.D. degree in the department of electronics and electrical engineering from POSTECH, Korea in 1997. From May 1997 to February 2000, he was a Research Associate in ERC-ARC, POSTECH. In March 2000, he joined Yeungnam University, Korea, where he is currently the Chuma Chair Professor. His research interests include robust control and filtering, neural networks, complex networks, multi-agent systems, and chaotic systems. He serves as Editor of Int. J. Control, Automation and Systems. He is also a Subject Editor/Associate Editor/Editorial Board member for several international journals, including IET Control Theory Appl., Appl. Math. Computation, Journal of The Franklin Institute, Nonlinear Dynamics, IEEE Trans. Fuzzy Systems and so on. He is a recipient of Highly Cited Researcher Award listed by Clarivate Analytics since 2015. He is a Fellow of the Korean Academy of Science and Technology (KAST).

Email: jessie@ynu.ac.kr

**Ho-Youl Jung (정 호 열)**

He received the Ph.D. degree in the Department of Electronics Engineering from the Institut National des Sciences Appliquées de Lyon (INSA de Lyon), France, in 1988. He is currently a Professor in the Department of Information and Communications Engineering, Yeungnam University, Korea. Both his teaching and research interests include digital signal processing, image signal processing, deep-learning, advanced driver assistance system, and autonomous vehicle.

Email: hoyoul@yu.ac.kr