

Artificial Neural Network Method Based on Convolution to Efficiently Extract the DoF Embodied in Images

Jong-Hyun Kim*

*Professor, School of Software Application, Kangnam University, Yongin, Korea

[Abstract]

In this paper, we propose a method to find the DoF(Depth of field) that is blurred in an image by focusing and out-focusing the camera through a efficient convolutional neural network. Our approach uses the RGB channel-based cross-correlation filter to efficiently classify the DoF region from the image and build data for learning in the convolutional neural network. A data pair of the training data is established between the image and the DoF weighted map. Data used for learning uses DoF weight maps extracted by cross-correlation filters, and uses the result of applying the smoothing process to increase the convergence rate in the network learning stage. The DoF weighted image obtained as the test result stably finds the DoF region in the input image. As a result, the proposed method can be used in various places such as NPR(Non-photorealistic rendering) rendering and object detection by using the DoF area as the user's ROI(Region of interest).

▶ **Key words:** Artificial neural network, Convolutional neural network, Depth of field, Image processing, Region of interest, Object detection

[요 약]

본 논문에서는 카메라의 포커싱과 아웃포커싱에 의해 이미지에서 뚜렷하게 표현되는 피사체 심도(Depth of field, DoF) 영역을 효율적인 합성곱 신경망을 통해 찾는 방법을 제안한다. 우리의 접근 방식은 RGB채널기반의 상호-상관 필터를 이용하여 DoF영역을 이미지로부터 효율적으로 분류하고, 합성곱 신경망 네트워크에 학습하기 위한 데이터를 구축하며, 이렇게 얻어진 데이터를 이용하여 이미지-DoF가중치 맵 데이터 쌍을 설정한다. 학습할 때 사용되는 데이터는 이미지와 상호-상관 필터 기반으로 추출된 DoF 가중치 맵을 이용하며, 네트워크 학습 단계에서 수렴률을 높이기 위해 스무딩을 과정을 한번 더 적용한 결과를 사용한다. 테스트 결과로 얻은 DoF 가중치 이미지는 입력 이미지에서 DoF영역을 안정적으로 찾아내며, 제안하는 방법은 DoF영역을 사용자의 ROI(Region of interest)로 활용하여 NPR렌더링, 객체 검출 등 다양한 곳에 활용이 가능하다.

▶ **주제어:** 인공 신경망, 합성곱 신경망, 피사체 심도, 이미지 프로세싱, 관심영역, 객체 검출

-
- First Author: Jong-Hyun Kim, Corresponding Author: Jong-Hyun Kim
 - *Jong-Hyun Kim (jonghyunkim@kangnam.ac.kr), School of Software Application, Kangnam University
 - Received: 2021. 02. 05, Revised: 2021. 03. 16, Accepted: 2021. 03. 22.

I. Introduction

컴퓨터 비전의 근본적인 문제인 깊이 추정은 장면에서 기하학적 관계를 보여주는 대표적인 문제이다. 이러한 객체와 환경간의 관계는 기존의 객체 인식 작업을 개선할 뿐만 아니라 3D 모델링[1], 물리 모델[2], 자율 주행[3], 비디오 감시[4], 로봇 공학[5]과 같은 응용 분야에 적용할 수 있다. 이런 깊이 맵을 추정하기 위해 스테레오 이미지 기반의 기술들이 많이 제안되었다. 하지만, 깊이 맵과 다르게 피사계 심도는 카메라의 포커싱에 따라 표현되는 대표적인 특징 중 하나이다. 이미지나 동영상에 내포된 콘텐츠가 속도를 가질 경우나 아웃포커싱 되었을 때 색상이 뿌옇게 표현되는 영역이 나타난다. 공통적으로 영역이 뿌옇게 나타나는 현상은 크게 두 가지 원인이며, 1) 속도로 인한 모션 블러와 2) 포커싱과 아웃포커싱에 대한 차이로 인해 발생하는 피사계 심도 영역이다. 본 논문에서는 이미지에 내포되어 있는 피사계 심도 영역을 인공 신경망을 통해 효율적으로 찾아낼 수 있는 프레임워크를 제시한다.

이미지나 동영상 데이터를 활용하여 영상 내 콘텐츠의 분석, 객체의 검출, 사용자의 ROI를 계산하는 과정에서 이미지의 피사계 심도는 중요한 특징 중 하나이다[6,7]. 이미지에서 DoF가 포함된 영역을 뚜렷하게 식별하는 것은 어려운 일이며, 이 같은 이유는 DoF가 포함된 이미지에서 정확하게 포커싱 영역을 추출하는데 있어서 예측 및 근사할 수 있는 정보가 부족하기 때문이다. 이미지에 내포되어 있는 장면으로부터 기하학적 관계를 분석하고자 깊이 (Depth)를 계산하는 연구들이 꾸준히 연구되고 있으며 [8,9], 이 기법은 대부분 포즈(Pose) 결정, 객체 인식 등에서 활용되고 있을 뿐, 이미지에서 특징적인 ROI 영역이라고 식별하기에는 다소 무리가 있다. 실제로 DoF는 콘텐츠의 깊이와는 다르게 사용자가 집중적으로 보고 싶은 부분을 두드러지게 하는 기법이기 때문에 깊이 기반 연구들은 본 연구에서 제안하고자 하는 DoF와는 차이가 있다.

이미지에서 포커싱과 아웃포커싱의 차이는 색상이 선명 또는 뿌옇게 표현되는 차이가 있으며 이러한 차이는 이미지 공간상에서 색상을 미분했을 때 더욱더 큰 차이를 보기에 된다. 본 논문에서는 이러한 특징을 이용하여 합성곱 신경망 기반의 새로운 네트워크 모델을 제안한다.

II. Preliminaries

컴퓨터 비전 분야에서는 이미지로부터의 깊이 값 복원은 고전적인 문제로 알려져 있으며, 이번 장에서는 본 논

문과 관련이 있는 이미지 한 장으로부터 깊이 값을 계산하는 방법과 렌즈 흐림 효과(Lens blur effects)를 이용하여 깊이 값을 다루는 접근법들에 대해서 간략하게 살펴본다.

1. Single image depth estimation

합성곱 신경망은 일반적으로 깊이 카메라로부터 얻은 깊이 맵을 이용하여 네트워크를 학습하고, 이 학습으로부터 한 장의 이미지에 대한 깊이 맵을 예측하는 틀로 많은 연구들에서 활용되고 있다[10,11,12]. 그러나 깊이 카메라는 종종 레인지(Range), 이미지 품질, 이식성, 해상도 측면에서 몇 가지 제한이 되어 있기 때문에 학습을 시키는데 있어서 특정 장면에서만 주로 작동하는 문제점이 있다 (예 : 실내 공간[2], 거리의 풍경[13], 랜드 마크[14] 등). 최근에는 깊이 값을 예측하고 감독하기 위해 Aperture 스택을 사용하였으며, 실내 공간이나 꽃 이미지에서 그 효율성을 증명하였다[15]. 그러나 깊이 값을 감독하기 위한 접근법을 사용하는 경우 특히 다양한 장면에서 깊이 모호성 (Depth ambiguities)이 생길 수 있으며, 이 방법만으로는 고품질의 결과를 얻기 어렵다.

2. Multi-images depth estimation

여러 이미지에 걸쳐 관련성을 이용하는 것이 깊이 정보를 얻는 가장 일반적인 방법이며, 다양한 분야에서는 이 방법을 통해 얇은 DoF 효과를 얻으려고 했다. 예를 들어, 모바일 디바이스의 카메라 초점으로부터 깊이 맵을 계산하는 방법[16], 스테레오 카메라 쌍을 이용하는 방법[17], 기준선을 이용한 방법[18,19], 영상의 여러 프레임 데이터를 이용한 접근법들이 존재한다[20]. 이 방법들은 전처리 단계에서 계산된 깊이 값과 포커싱된 이미지를 사용하여 얇은 DoF 효과를 합성하는 절차를 이용하는 접근법들이다.

3. Lens blur effects

하나의 RGB-D 이미지에서 작동하는 렌즈 흐림 효과로는 객체 뒤에 숨겨져 있는 아웃포커싱된 영역을 정확하게 표현할 수 없다. 렌즈 블러를 근사하는 기존의 방법은 일반적으로 객체 공간과 이미지를 공간을 분류하여 사용한다. 객체 공간 방법은 광선 추적(Ray-tracing) 및 실제 카메라 모델을 살펴보고 사실적으로 얇은 DoF 이미지를 표현하는데 있어서 효과적이지만[21,22], 이러한 접근법은 대부분 복잡한 3차원 공간 정보를 필요로 하며, 그렇기 때문에 계산시간이 오래 걸린다. 이미지 공간 방법은 단일 이미지를 통해 수집(Gathering)과 확산(Spreading) 과정을 거치고, 이것으로부터 얇은 DoF 효과를 만든다

[23,24,25]. 아이폰(iPhone) 및 구글픽셀 2(Google Pixel 2)와 같은 고급 모바일 디바이스에는 DoF 이미지를 시뮬레이션하는 “Portrait mode”라는 응용프로그램이 내장되어 있다. 이 방법은 다중 뷰 또는 듀얼 픽셀 기반으로 깊이 맵을 근사한다. 기존의 방법들은 영상 내에서 깊이 값(Depth value)만을 찾으려고 했으며, 실제로 DoF를 찾으려는 연구는 아직 없었다. DoF가 없는 영상을 포커싱과 아웃포커싱 효과가 나타나도록 만들어주는 게 전부이다. 깊이 값을 평가하거나 계산하는 방법은 앞에서 언급한 것처럼 다양한 방법들이 존재하지만, DoF는 깊이 값과는 다른 특징을 가지고 때문에 이러한 접근법으로 정확한 DoF 영역을 찾아내는 것은 어려운 일이다.

$$H(x,y) = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \approx \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{\sigma^2}} \quad (2)$$

이미지에 존재하는 DoF 영역을 추출하기 위해 본 논문에서는 몇 가지 가정을 한다:

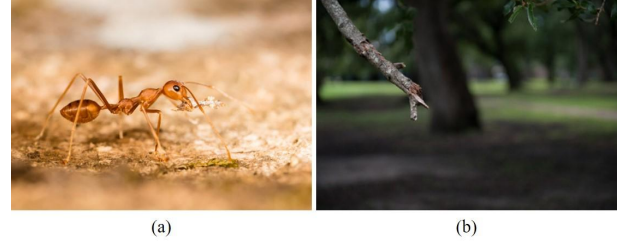


Fig. 1. Focusing and defocusing regions in photography.

III. The Proposed Scheme

1. Extraction of DoF region from images for collecting training data set

이번 장에서는 학습 단계에서 필요한 데이터를 셋을 구축할 수 있는 방법에 대해 설명한다. 일반적인 이미지 슈퍼 해상도 접근법에서는 고해상도 이미지를 다운샘플링(Down-sampling)하여 저해상도 이미지를 생성하고 이 두 이미지들 간의 관계를 이용하여 학습을 진행하지만, DoF 영역을 추출하기 위해서는 좀 더 특별한 데이터 셋이 필요하며 이를 구축하기 위한 방법을 설명한다.

본 논문에서는 이미지로부터 DoF 영역을 계산하기 위해 상호-상관 필터 G 를 이용한다. 이 필터는 두 개의 연속된 데이터들이 얼마나 연관되어 있는지를 계산하는 방법으로 영성처리, 컴퓨터 비전 등 다양한 분야에서 활용되고 있는 방법이다. 본 연구에서는 서로의 색상 값이 얼마나 유사하게 관련되어 있는지를 측정하는 방법으로 활용하며, 수식 1을 이용하여 계산한다.

$$G(x,y) = H \otimes F = \sum_{u=-1}^1 \sum_{v=-1}^1 H(u,v) F(x+u, y+v) \quad (1)$$

여기서 H 는 이미지에서 각 인접 픽셀들의 가중치로 마스크(Mask)라고 부르며, F 는 인접 픽셀들의 색상 값이다. 이 마스크는 적용 분야에 따라 다양하게 모델링되는데, 본 논문에서는 가우시안 필터링(Gaussian filtering) 기법을 이용한다(수식 2 참조). 이 수식은 가우시안 함수와 유사한 형태로 모델링된 수식이다.

- 이미지에서 표현된 DoF 영역에서는 피사계 심도를 알게 하여 초점에 맞는 피사체를 제외한 배경은 흐려지게끔 뭉개지는 효과가 나타나야 한다 (Fig. 1a 참조).
- 아웃포커싱에 의해 뿌옇게 흐려진 영역과 선명한 영역의 차이는 명확하다 (Fig. 1b 참조).

위에서 언급한 특징을 이용하여 원본이미지 I_0^{gb} 와 원본 이미지로부터 생성한 스무딩(Smoothing) 이미지 I_1^{gb} 의 색상 채널의 차이를 기반으로 DoF가중치 맵 D 를 다음과 같이 계산한다(수식 3 참조). 이 수식은 상호상관 필터로부터 유도된 수식이며, 우리는 이 수식을 활용하여 영상의 DoF를 러프하게 찾아낸다.

$$D(x,y) = \frac{\sum_{u=-m}^m \sum_{v=-m}^m H_m(u,v) O(x+u, y+v)}{m^2} \quad (3)$$

$$O(x,y) = \| I_0^{gb}(x,y) - I_1^{gb}(x,y) \| \quad (4)$$

여기서 H_m 과 m 은 각각 마스크와 마스크의 크기를 나타내며, 본 논문에서 마스크의 크기는 15로 설정했다. 마스크의 크기는 실험적으로 얻어진 값이다. O 는 앞에서 언급한 뿌연 이미지와 선명한 이미지 사이의 색상 차이를 계산한 수식이다. I_0^{gb} 는 원본 이미지로부터 얻은 색상이고, I_1^{gb} 는 가우시안 스무딩 필터를 이용하여 계산한 뿌연 이미지로부터 얻은 색상이다. 수식 4에서 보듯이 DoF 영역의 가중치는 RGB채널 값을 3차원 벡터로 표현하여, 그 벡터의 크기를 통해 계산한다.

Fig. 2b는 입력 이미지를 통해 얻은 DoF가중치 맵인 D 이다. DoF에 의해 Fig. 2a에서 아웃포커싱된 뿌연 영역의 가중치가 잘 표현되었고, 이 방법으로 네트워크 학습에 필요한 이미지 데이터 셋을 구축한다.







Fig. 2. DoF weight map D calculated using the DoF region (white : focusing, black : defocusing).

2. Convolutional neural networks with DoF weight map

우리는 앞에서 설명한 방법을 이용하여 RGB채널을 가진 입력 이미지인 $\{\delta^1, \delta^2, \dots\}$ 와 DoF가중치 맵 이미지인 $\{D^1, D^2, \dots\}$ 를 생성한다. 각 이미지를 학습 네트워크에 넣기 전에 패치단위로 분할한다. 학습 데이터가 주어지면, 우리의 목표는 예측된 값인 δ_s 와 실제 값인 D 사이의 오차를 최소화하는 매칭 함수 $f(x)$ 를 찾는 것이다. 이 과정을 수행하기 위한 목적 함수는 예측된 이미지와 실제 이미지 사이의 MSE이다. 우리의 목표는 $\delta_s = f(x)$ 값을 예측하는 모델 f 를 학습하는 것이며, 결과적으로 학습 데이터에 대한 MSE를 최소화하는 것이다 (수식 5 참조).

$$MSE = \frac{1}{2} \| D - f(x) \|^2 \quad (5)$$

본 논문에서는 위 수식을 최소화시키기 위해 SRCNN기반의 네트워크 방식을 활용한다[26]. 제안하는 방법을 SRCNN 네트워크 방식에 적용하기 위해 입력 이미지의 DoF가중치 맵에서, 이 DoF가중치 맵을 스무딩한 이미지로 변환하기 위한 가중치 매개변수 θ 를 계산해야 한다 : $\delta \rightarrow D - (\delta^* - D^*, \theta)$. 여기서 δ, D 는 입력 이미지와 그에 해당하는 DoF가중치 맵이며 (δ : , D : ) , δ^*, D^* 는 δ 를 스무딩한 입력 이미지이며, 그에 해당하는 DoF가중치 맵이다(δ^* : , D^* : ). 결과적으로 θ 는 우리가 찾고자하는 가중치 매개변수이다. 스무딩을 한 이유는 앞서서도 언급했듯이, 이미지에서 포커싱과 아웃포커싱의 차이는 선명하거나 뿌옇게 표현되는 차이가 있으며, 이

는 공간상에서 미분을 했을 때 더 큰 차이를 보이기 때문이다. 이러한 방식을 기반으로 학습을 시켰을 때 안타깝게도 우리가 기대한 결과를 얻을 수 없었으며, 학습과정에서 반복횟수를 증가시켜도 수렴하지 않는 결과를 보였다 (Fig. 3 참조). 우리는 15,000번 정도 학습시켰을 때 DoF영역이 아닌, 마치 흑백 이미지와 같은 결과로 수렴하는 학습 결과를 얻었다 (Fig. 4 참조).



Fig. 3. Training results according to the number of iteration.

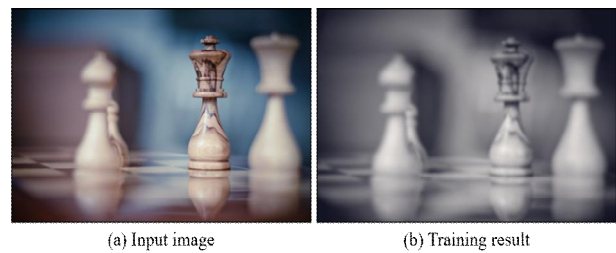


Fig. 4. Training result obtained through 15,000 iterations.

이 같은 이유는 매핑 함수 $f(x)$ 를 학습하는 과정에서 δ 와 D 의 차이가 서로 너무 다르기 때문에 원하는 방향으로 수렴하지 않는 결과가 나타났으며, 다양한 이미지에서 추가 실험을 해봤지만 여전히 같은 결과가 나타났다. 본 논문에서는 이 같은 문제를 해결하기 위해 DoF가중치 맵인 D 에 원본 이미지를 곱한 이미지인 $D' = D \cdot \delta$ 를 이용하여 네트워크의 학습 데이터로 사용하였다 (Fig. 5 참조).

Fig. 6은 SRCNN 기반으로 본 논문에서 사용한 네트워크 모델이며[26], 입력 이미지로부터 처리되는 각 단계는 1) 패치 추출, 2) 비선형 매핑, 3) 복원단계로 이루어진다.

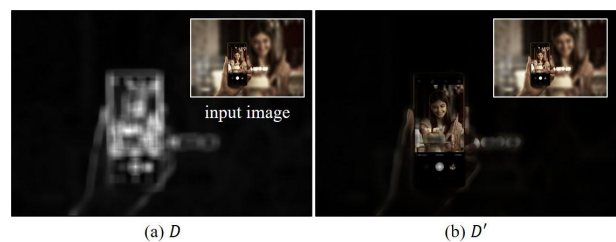


Fig. 5. Advanced weight image of DoF region.

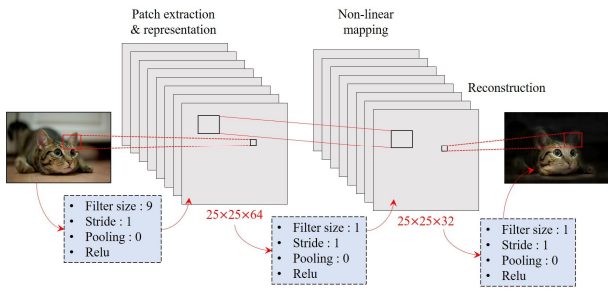


Fig. 6. The proposed network architecture.

IV. Results

본 연구의 결과들을 만들기 위해 실험한 환경은 Intel Core i7-7700K CPU, 32GB RAM, Geforce GTX 1080Ti GPU가 탑재된 컴퓨터를 이용하였다. 본 논문에서는 상호상관 필터를 이용하여 DoF 영역을 찾는 알고리즘을 제안했고, 이 기술을 인공지능을 통해 자동으로 학습하고 테스트 할 수 있는 새로운 합성곱 신경망 기법을 제안했다.



Fig. 7. Training results according to the number of iteration with our method.

Fig. 7은 본 논문에서 제안한 방법을 이용하여 학습되는 중간 결과를 보여주는 그림이다. Fig. 3은 15,000번의 반복적인 학습을 수행했음에도 불구하고 DoF와는 관계없는 흑백 이미지로 수렴했지만, 제안한 방법은 훨씬 적은 반복 횟수인 1,400번에 DoF 영역을 빠르게 찾아냈다. 이 같은 결과의 의미는 DoF 영역 추출과 별개로 수렴률이 굉장히 빠르다는 것을 나타낸다. 또한, 학습하는 과정에서 425장의 적은 데이터 셋으로 좋은 결과를 만들어 냈다. 일반적으로 수천, 수만 장의 데이터를 이용하는데 비해서 본 연구에서는 수렴률을 향상시키기 위한 전처리 방법으로 인해 적은 데이터만으로 높은 학습률을 가능하게 했다.

Fig. 8은 본 논문에서 제안하는 방법을 통해 얻어진 테스트 결과이며, 다양한 이미지들에서도 안정적으로 DoF가 중치 맵을 추출했다. 레인보우 색상 결과에서 빨간색에 가까울수록 포커싱 영역을 의미하며, 파란색에 가까울수록 아웃포커싱된 영역을 나타낸다.

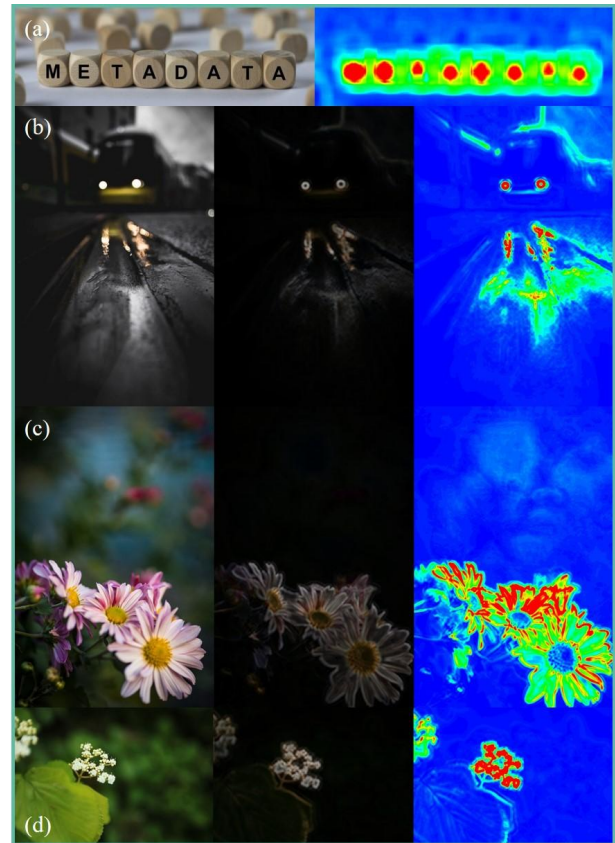


Fig. 8. Test results with our framework (rainbow color : DoF weight map).

Fig. 8a에서 입력 이미지는 나무 조각과 알파벳 영역에 포커싱이 맞춰져 있으며 상대적으로 배경 부분에는 아웃포커싱이 되었고, 이 특징을 DoF가중치 맵으로 잘 추출하였다. 특히, 알파벳 영역보다 나무 조각 외곽 부분에 가중치 값이 낮은 이유는 뿌옇게 변하는 색의 변화가 배경에서 나무 조각으로 부드럽게 변하기 때문이고, 이러한 특징까지 잘 포착했다. 색상의 변화가 거의 없는 어두운 영상에서도 DoF를 안정적으로 찾아냈으며, 꽃 이미지에서도 포커싱과 아웃포커싱 영역을 명확하게 구분하였다 (Figs. 8b~8d 참조). 제안하는 방법은 대부분의 결과에서 포커싱을 나타내는 DoF 영역을 안정적으로 잘 추출하였다. 앞에서 설명한 제안한 방법의 기술적 장점은 아래와 같다 :

- 한 장의 이미지로부터 DoF 영역을 추출할 수 있는 기술 개발
- 인공지능을 통해 DoF 영역을 학습시키기 위한 데이터 셋을 구축할 수 있는 기술 개발
- DoF 영역을 추출하기 위해 MSE(Mean squared error)의 수렴률을 향상시킬 수 있는 합성곱 신경망 기술 개발

V. Conclusions

본 논문에서는 합성곱 신경망을 이용하여 DoF 영역을 추출할 수 있는 새로운 네트워크 모델을 제안했다. 이 모델을 학습시키기 위해 필요한 메타 데이터를 상호-상관 필터를 통해 계산하고, 학습과정에서 수렴률을 개선시키기 위한 방법을 제시하여 적은 데이터와 적은 학습 반복횟수로도 빠르게 DoF 영역을 추출할 수 있게 하였다. DoF가 적용된 영상 데이터 셋이 없기 때문에 우리는 웹에서 크롤링한 영상 데이터들을 사용하였다. 제안하는 네트워크 방식은 SRCNN을 활용하고 있기 때문에 네트워크가 깊지 않으며, 향후 네트워크를 좀 더 추가하여 정확도가 높은 인공지능 모델로 확장할 계획이다. 또한, DoF의 시점 변화를 추적하여 사용자의 ROI를 좀 더 정확하게 포착할 수 있는 방법으로 확장할 예정이다.

REFERENCES

- [1] Gargallo, Pau, and Peter Sturm. "Bayesian 3D modeling from images using multiple depth maps." In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, pp. 885-891. IEEE, 2005.
- [2] Silberman, Nathan, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. "Indoor segmentation and support inference from rgb-d images." In European conference on computer vision, pp. 746-760. Springer, Berlin, Heidelberg, 2012.
- [3] Häne, Christian, Lionel Heng, Gim Hee Lee, Friedrich Fraundorfer, Paul Furgale, Torsten Sattler, and Marc Pollefeys. "3D visual perception for self-driving cars using a multi-camera system: Calibration, mapping, localization, and obstacle detection." *Image and Vision Computing*, pp. 14-27, 2017.
- [4] Müller, K., Merkle, P. and Wiegand, T., "3-D video representation using depth maps". *Proceedings of the IEEE*, 99(4), pp.643-656, 2010.
- [5] Häne, Christian, Christopher Zach, Jongwoo Lim, Ananth Ranganathan, and Marc Pollefeys. "Stereo depth map fusion for robot navigation." *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1618-1625, 2011.
- [6] Rafiee, Gholamreza and Dlay, Satnam Singh and Woo, Wai Lok, Region-of-interest extraction in low depth of field images using ensemble clustering and difference of Gaussian approaches. *Pattern Recognition*, pp. 2685-2699, 2013.
- [7] Park, Jungwoo and Kim, Changick, Extracting focused object from low depth-of-field image sequences. *Visual Communications and Image Processing*, pp. 607710, 2006.
- [8] Simonyan, Karen and Zisserman, Andrew, Data-driven estimation of cloth simulation models. *arXiv preprint arXiv:1409.1556*, 2014.
- [9] Hadsell, Raia and Sermanet, Pierre and Ben, Jan and Erkan, Ayse and Scoffier, Marco and Kavukcuoglu, Koray and Muller, Urs and LeCun, Yann, Learning long-range vision for autonomous off-road driving. *Journal of Field Robotics*, pp. 120-144, 2009.
- [10] Eigen, David, Christian Puhrsch, and Rob Fergus. "Depth map prediction from a single image using a multi-scale deep network." In *Advances in neural information processing systems*, pp. 2366-2374. 2014.
- [11] Laina, Iro, Christian Rupprecht, Vasileios Belagiannis, Federico Tombari, and Nassir Navab. "Deeper depth prediction with fully convolutional residual networks." In *2016 Fourth international conference on 3D vision (3DV)*, pp. 239-248. IEEE, 2016.
- [12] Liu, Fayao, Chunhua Shen, and Guosheng Lin. "Deep convolutional neural fields for depth estimation from a single image." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5162-5170. 2015.
- [13] Geiger, Andreas, Philip Lenz, and Raquel Urtasun. "Are we ready for autonomous driving? the kitti vision benchmark suite." In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3354-3361, 2012.
- [14] Li, Zhengqi, and Noah Snavely. "Megadepth: Learning single-view depth prediction from internet photos." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2041-2050. 2018.
- [15] Srinivasan, Pratul P., Rahul Garg, Neal Wadhwa, Ren Ng, and Jonathan T. Barron. "Aperture supervision for monocular depth estimation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6393-6401. 2018.
- [16] Suwajanakorn, Supasorn, Carlos Hernandez, and Steven M. Seitz. "Depth from focus with your mobile phone." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3497-3506. 2015.
- [17] Barron, Jonathan T., Andrew Adams, YiChang Shih, and Carlos Hernández. "Fast bilateral-space stereo for synthetic defocus." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4466-4474. 2015.
- [18] Ha, Hyowon, Sunghoon Im, Jaesik Park, Hae-Gon Jeon, and In So Kweon. "High-quality depth from uncalibrated small motion clip." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5413-5421. 2016.
- [19] Joshi, Neel, and C. Lawrence Zitnick. "Micro-baseline stereo." *Technical Report MSR-TR-2014-73*, p. 8, 2014.
- [20] Klose, Felix, Oliver Wang, Jean-Charles Bazin, Marcus Magnor, and Alexander Sorkine-Hornung. "Sampling based scene-space video processing." *ACM Transactions on Graphics (TOG)* vol, 34, no. 4, pp. 67, 2015.
- [21] Haeberli, Paul, and Kurt Akeley. "The accumulation buffer:

- hardware support for high-quality rendering." ACM SIGGRAPH computer graphics vol. 24, no. 4, pp. 309-318, 1990.
- [22] Lee, Sungkil, Elmar Eisemann, and Hans-Peter Seidel. "Real-time lens blur effects and focus control." ACM Transactions on Graphics (TOG) vol. 29, no. 4, pp. 1-7, 2010.
- [23] Kraus, Martin, and Magnus Strengert. "Depth-of-field rendering by pyramidal image processing." In Computer graphics forum, vol. 26, no. 3, pp. 645-654. Oxford, UK: Blackwell Publishing Ltd, 2007.
- [24] Lee, Sungkil, Gerard Jounghyun Kim, and Seungmoon Choi. "Real-time depth-of-field rendering using anisotropically filtered mipmap interpolation." IEEE Transactions on Visualization and Computer Graphics vol. 15, no. 3, pp. 453-464, 2009.
- [25] Yang, Yang, Haiting Lin, Zhan Yu, Sylvain Paris, and Jingyi Yu. "Virtual dslr: High quality dynamic depth-of-field synthesis on mobile platforms." Electronic Imaging, no. 18, pp. 1-9, 2016.
- [26] Dong, Chao and Loy, Chen Change and He, Kaiming and Tang, Xiaoou, Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, pp. 295-307, 2015.

Authors



Jong-Hyun Kim received the B.A. degree in the department of digital contents at Sejong University in 2008. He received M.S. and Ph.D. degrees in the department of computer science and engineering at Korea University,

in 2010 and 2016. Prof. Kim is an associate professor in the school of software application in Kangnam University. His current research interests include fluid animation and virtual reality.